The Department of Integrative Medical Biology
Umeå University, SE-901 87 Umeå, Sweden
www.imb.umu.se

ISBN 978-91-7601-53-8
ISSN 0346-6612
New Series No 1835

Fredrik Bergström

The Neural Substrates of Non-Conscious Working Memory

Umeå University 2016

The Neural Substrates of Non-Conscious Working Memory

Fredrik Bergström

The Department of Integrative Medical Biology
Physiology Section
Umeå 2016

The Neural Substrates of Non-Conscious Working Memory

Fredrik Bergström

The Department of Integrative Medical Biology,
Physiology Section.
Umeå 2016

Responsible publisher under Swedish law: the Dean of the Medical Faculty

ISBN: 978-91-7601-53-8
ISSN: 0346-6612
New Series No: 1835
Ev. info om Omslag/sättning/omslagsbild:
Elektronisk version tillgänglig på http://umu.diva-portal.org/
Tryck/Printed by: Print & Media, Umeå University
Umeå, Sweden 2016

In memory of my friend, lover, and the most beautiful mind I have known.
Forever remembered, cherished, and admired.

Aida Alipour (1981 - 2016)

Mina visa ord (för-) blir
"nothing worth having
comes easily"

Puss

Aida

Table of Contents

Acknowledgements

First and foremost I want to thank my two and a half supervisors for their collective knowledge, guidance, and support, Johan Eriksson for his reflective and analytical approach, Lars Nyberg for his piercing and insightful comments, and Greger Orädd for his advice during the brief period as supervisor.

Secondly, I want to thank Benoni Edin for stimulating conversations and valuable book recommendations. One of those books was by Joaquín Fuster, who although minimal personal correspondence, has greatly influenced my conceptual understanding of the brain through his impressive body of work.

Thirdly I want to thank all my current and former colleagues for their support, especially Micael Andersson for technical support, Anders Lundqvist for fun with statistics, Göran Westling for building me stereoscopes, Mikael Stiernstedt for tips and tricks with Illustrator and Photoshop, Karolina Kauppi, Sara Pudas, Linnea Karlsson Wirebring, Per Nordmark, Lenita Lindgren, Urban Ekman, Sara Stillesjö, Amar Awad for creating a stimulating environment and random hilarities, the MR staff and administrators for narrowing my focus of attention.

Last but not least, I want to thank my friends and family for their support and help in alleviating any social isolation, especially Mauritz Börjeson and Jagannath Tammeleth for keeping in touch despite long distances and busy schedules. Special thanks to Aida Alipour for making life colorful.

List of papers

This doctorate dissertation is based on the following three studies:

I. Bergström, F., & Eriksson, J. (2014). Maintenance of non-consciously presented information engages the prefrontal cortex. *Frontiers in Human Neuroscience*, *8*(938), 1–10. doi:10.3389/fnhum.2014.00938, PMID: 25484862

II. Bergström, F., & Eriksson, J. (2015). The conjunction of non-consciously perceived object identity and spatial position can be retained during a visual short-term memory task. *Frontiers in Psychology*, *6*(1460), 1–9. doi:10.3389/fpsyg.2015.01470, PMID: 26483726

III. Bergström, F., & Eriksson, J. (2016). Neural evidence of non-conscious short-term memory. *Manuscript.*

Other papers

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive Architecture of Working Memory. *Neuron*, *88*(1), 33–46. doi:10.1016/j.neuron.2015.09.020, PMID: 26447571

Abstract

Background: Despite our distinct impression to the contrary, we are only conscious of a fraction of all the neural activity underlying our thoughts and behavior. Most neural processes occur non-consciously, and in parallel with our conscious experience. However, it is still unclear what the limits of non-conscious processes are in terms of higher cognitive functions. Many recent studies have shown that increasingly more advanced functions can operate non-consciously, but non-conscious information is still thought to be fleeting and undetectable within 500 milliseconds. Here we used various techniques to render information non-conscious, together with functional magnetic resonance imaging (fMRI), to investigate if non-consciously presented information can be retained for several seconds, what the neural substrates of such retention are, and if it is consistent with working memory maintenance.

Results: In *Study I* we used an attentional blink paradigm to render stimuli (single letters) non-conscious, and a variable delay period (5 – 15 s) prior to memory test. It was found that non-conscious memory performance was above chance after all delay durations, and showed no signs of decline over time. Univariate fMRI analysis showed that the durable retention was associated with sustained BOLD signal change in the prefrontal cortex and cerebellum during the delay period. In *Study II* we used continuous flash suppression (CFS) to render stimuli (faces and tools) non-conscious, and a variable delay period (5 or 15 s) prior to memory test. The durable retention of up to 15 s was replicated, and it was found that stimuli identity and spatial position was retained until prospective use. In *Study III* we used CFS to render tools non-conscious, and a variable delay period (5 – 15 s) prior to memory test. It was found that memory performance was not better than chance. However, by using multi-voxel pattern analysis it was nonetheless possible to detect the presence vs. absence of non-conscious stimuli in the frontal cortex, and their spatial position (left vs. right) in the occipital cortex during the delay.

Conclusions: Overall these findings suggest that non-consciously presented information (identity and/or position) can be retained for several seconds, and is associated with BOLD signal in frontal and posterior regions. These findings are consistent with working memory maintenance of non-consciously presented information, and thereby constrain models of working memory and theories of consciousness.

Keywords: non-conscious, working memory, neural substrates, visual perception, consciousness, functional magnetic resonance imaging (fMRI)

Abbreviations

AB	attentional blink
BOLD	blood-oxygenation-level-dependent (signal)
CFS	continuous flash suppression
EEG	electroencephalography
fMRI	functional magnetic resonance imaging
GNW	global neuronal workspace (theory)
HOT	higher-order thought (theory)
ISI	inter stimulus interval
ITI	inter trial interval
LGN	lateral geniculate nucleus
MEG	magnetoencephalography
MVPA	multivoxel pattern analysis
PAS	perceptual awareness scale
tDCS	transcranial direct-current stimulation
TMS	transcranial magnetic stimulation

Introduction

We are memories. Our memories define who and what we are, that is, everything from what species we are to our individual personalities. As we spring into existence we have already inherited memories from our parents and ancestors all the way back to the beginning of life on earth. Our genetic inheritance interacts with our environment, and form new memories throughout our existence. The vast majority of the memories that make up our uniqueness as human beings is at any given time *non-conscious* (i.e., not consciously experienced), and latently stored as widely-distributed and interconnected webs of neurons. Together they constitute the grey and squishy bundle of nerve-cells commonly known as the brain.

Normally, subsets of these interconnected neurons will become active and fire signals between each other, which enables us to go about our everyday lives. Inexplicably, these activated networks of neurons also enable us to *consciously experience* the content of our current memories as we perceive our external and internal environment – sensory experiences, emotions, goals and motivations, sense of agency, etc. – as we navigate our way through the adventures of life. How our brain gives rise to our subjective experiences remains a mystery. Philosophers call this mystery “the explanatory gap”. This gap is what some philosophers and cognitive neuroscientists are trying to explain. One way to become informed about what is special with consciousness, is to get a better understanding of what is *not* special with it, by investigating the limits of non-conscious cognition.

There is a special kind of short-term memory that is intimately linked with conscious experience. It is called *working memory* – the temporary retention of information for prospective use – and is immensely important for us. Working memory is used when we, for example, perform mental arithmetic, are trying to solve difficult problems, or doing contingency planning for the future. Not surprisingly, working memory capacity is therefore a good predictor of general intelligence and success in life. Indeed, most of our advanced behavior is associated with conscious experience and working memory. However, most of our neural activity occurs non-consciously, and in parallel with our conscious experience. The limits of our non-conscious cognition in terms of higher cognitive functions remain unclear. Empirical studies have shown that increasingly more advanced cognitive functions can operate non-consciously, but it is still believed that non-consciously perceived information is fleeting and undetectable within 500 milliseconds. The questions we have tried to find answers to within the context of this dissertation are if non-conscious presented information can be retained for several seconds, what the neural substrates of such retention are, and if it is consistent with working memory maintenance.

Background

Building blocks of cognition

This section is meant to briefly define the functional properties of the brain that are relevant for this dissertation, and give a brief overview of the functional organization of the brain. Importantly, the cognitive functions will be defined at a psychological level, as well as anchored in neural terms.

Memory

Our memories define what and who we are, without them, we would literally not exist. Memory is therefore the most fundamental property of any organism, and forms the basis of all cognition. Indeed, the word *cognition*, which today is used as a blanket term for all cognitive functions, comes from the Latin word *cognoscere* meaning "*to learn*" or "*to know*". Broadly defined, *memory is an organism's capacity to retain information about itself and its external environment*, such as in tissue scarring, the immune system, or the brain. However, here we will be focusing on memories in the brain.

The science of memory has a long history dating back to the ancient Greek philosophers (Hergenhan, 2001; Radvansky, 2014). Plato thought humans had immortal souls that bathed in pure and complete knowledge, which only could be attained by introspection and reason, rendering sensory experience redundant. He therefore thought humans were born with all their memories. Aristotle, instead, thought humans were born as *tabulae rasae* ("blank slates"), and that memories were imprints (as on wax tablets) caused by sensory experience to later be recalled. Contrary to Plato, he therefore thought that we acquire all our memories through experience. Aristotle also postulated what were to become the bases of all learning theory, namely, *the laws of association*. These laws state that when we think of something we also tend to think of things that are: experienced along with it (*the law of contiguity*), similar to it (*the law of similarity*), and its opposite (*the law of contrast*). Additionally, he posited that associations generally become stronger the more they are experienced (*the law of frequency*), but can sometimes become strong after only one experience. Aristotle's ideas on memory, as we shall see, were not all too far off, except about memories being stored in the heart instead of the brain.

Ebbinghaus (1885/1913) was the first experimental psychologist to systematically study memory and learning. He studied himself while trying to learn lists of nonsense syllables with varying list lengths, retention intervals, and learning conditions. From his meticulous efforts he was able to estimate how long it took, and how frequent practice was needed to learn certain information (*the learning curve*), the optimal spacing between practice sessions, the time it takes to forget information (*the forgetting curve*), that overlearning could extend the forgetting curve, and that forgotten memories no longer consciously accessible nevertheless could have a non-conscious influence on behavior and learning.

In the brain, memories are stored in the structure of interconnected and distributed networks of neurons. New neural memories are formed by Hayek-Hebbian principles (“*neurons that fire together wire together*”) to create new or modify existing neural networks (Hayek, 1952; Hebb, 1949). These neural connections are formed throughout the neocortex with important modulatory input from subcortical limbic regions, such as the hippocampi and amygdalae. Thus, similarly to what Aristotle postulated, connections between neurons (i.e., memories) are formed by way of associations that become stronger with frequency. However, an emotional event can cause life-long associations by way of modulatory input from the amygdalae.

Contrary to the thoughts of Plato and Aristotle, organisms can be said to consist of innate memories (that we are born with), and acquired memories. Innate memories are called *phyletic memories* (“memory between species and organisms”), while memories formed by experience are called *ontogenetic memory* (“memory within the life span of an organism”). Phyletic memory consists of the genetically determined functional and structural layout of the brain (all things that are similar across individuals), and serves as the base on which ontogenetic memories grow (things that are different between individuals). Phyletic and ontogenetic memories are intertwined, and there is no way to clearly separate them. The evolutionarily older primary sensory and motor regions seem to be relatively more phyletic, while the relatively newer cortical regions of association seem to be more ontogenetic. However, even primary sensory cortices need some sensory experience during critical periods to develop functionally. In the broadest sense, we therefore consist entirely of memories (our ancestor’s and our own).

Memory, therefore, is a global property of the brain, and cannot be assigned to any specific part of the brain. However, memory can be differentiated based on its current state (active or inactive) over time. The difference between active and inactive memory states has been used to distinguish between short-term (active) and long-term (inactive) memory, as well as conscious (active) and non-conscious (inactive) states. Traditionally, short-term memory was therefore assumed to be an active conscious state, while latent long-term memory was assumed to be an inactive non-conscious state (until reactivated). Short-term memory (or retention) is henceforth used as an umbrella term to refer to any short-lived memory (i.e., sensory memory, working memory, and short-lived latent neural changes). Sensory memory is a brief high capacity memory, and working memory is a durable low capacity memory dependent on prospective use. Memory can also be spatially differentiated in terms of its specific content. That is, different memory content can be found in different areas of the brain. The spatial distinction between content and the temporal distinction among states largely underlies the categorization of memory (Fuster, 1995; Tulving, 1972). The various categorizations of short-term and long-term memory will be more closely reviewed below in the section on memory and conscious experience.

Perception & action

Simplified, the frontal cortex can be said to contain motor memory, while the posterior parts of the cortex contain perceptual memory, but the two are intertwined and the distinction is merely conceptual (Fuster, 1995, 2003a, 2015). In the posterior cortex, perceptual knowledge is organized hierarchically with concrete unimodal elementary sensory features in the primary sensory cortices, and progressively more abstract polymodal sensory, semantic, and conceptual knowledge towards the central regions of the posterior cortex (Binder & Desai, 2011; Fuster, 1997, 2003a). Perception is not a passive process, nor does it reflect reality as it really is. As suggested by Berkeley in 1709, Helmholtz in 1925, and Hayek (1952) perception is the active process of interpretation and categorization of sensory information guided by memory (Fuster, 2003a; Palmer, 1999). Essentially, we perceive the world as we remember it. That is, new sensory information reactivates (phylogenic and ontogenetic) memory to process and interpret it. In the process, new sensory information is stored as extensions to old perceptual memory, which in turn will interpret incoming sensory information accordingly. To perceive, therefore, is to remember or “re-cognize” (Fuster, 1995).

Motor knowledge follows a similar hierarchical gradient such as simple movement-related information in the primary motor cortex, goal-related information in premotor cortex, plans and contingencies in the prefrontal cortex (Badre & D’Esposito, 2009; Fuster, 2015; Koechlin & Summerfield, 2007). It has been suggested that the prefrontal cortex is crucial for temporal integration of the past and the future, which is an essential component of working memory (Fuster & Bressler, 2015; Ingvar, 1985). However, the lowest levels of the motor hierarchy are not in the cortex, but in the basal ganglia, cerebellum, brain stem, and spinal cord, where well-learned sensory-motor interactions and reflexive actions are stored. Action is the activation of (phylogenic and ontogenetic) executive and/or simpler motor memory.

The perceptual and motor hierarchies are interconnected at all hierarchical levels in the cortex, and together with subcortical modulation, form a perception-action cycle (Fuster & Bressler, 2015; Fuster, 2003a, 2004). The perception-action cycle makes the brain a causally deterministic self-guiding system, without need of an executive homunculus or free will to govern it. It is through the perception-action cycle an organism traverses its environment, by gathering sensory information as input, transforming it to actionable output, as it continuously learns and adapts to the environment. The perception-action cycle is also crucial for sustaining recurrent activity of internal representations without external input, for example during working memory. However, perception as well as action, relies on attention to be able to select what to perceive and what action to initiate. Indeed, perception and action is the interaction between memory and attention.

Attention

In our everyday lives, aspects of attention can be described by the "cocktail party effect" (Broadbent, 1958). That is, in a room full of voices it is possible to selectively focus on and understand one voice at a time (and switch as you please), while all other voices become garbled background noise. However, if one of the background voices addresses you by name you are likely to automatically switch focus to that voice. This scene captures many of the distinguishing features of attention. Throughout history there has been many influential cognitive models of attention born out of the cocktail party effect (Broadbent, 1958; Driver, 2001; Neisser, 1967; Treisman & Gelade, 1980; Treisman, 1960). However, I prefer the following one, which is stated more closely to neurophysiology, but is nonetheless consistent with neuropsychological findings.

Attention is another fundamental property of the nervous system, and can be defined as *the selective allocation of neural resources*. In neural terms this translates to excitation of some neural networks and inhibition of other competing or irrelevant networks (Desimone & Duncan, 1995; Franconeri, Alvarez, & Cavanagh, 2013; Fuster, 2003a). The efficient allocation of attentional or neural resources is crucial for normal functioning because of its limited capacity (Marois & Ivanoff, 2005). The *focus (or foci) of attention* is the neural networks with most resources allocated to them, that is, most neural activity at a specific point in time. The information inside the focus of attention is usually consciously experienced. Outside the focus of attention there is less neural activity, that is, less attended networks, as well as, unattended networks that are (more or less) completely inhibited. Less attended information can be consciously or non-consciously processed, while unattended information is not processed at all (Chun & Marois, 2002). *Exogenous attention* is when the allocation of attentional resources are caused by external stimuli, usually something salient (e.g., a loud bang or one's name) that reflexively captures ones attention. *Endogenous attention* refers to when attentional resources are allocated by internal processes such as an intention or will to achieve a goal or task. However, endogenous and exogenous attention does not capture all aspects of attention. Other subcortical structures involving motivation, reward, and emotion also modulates attention in the cortex (Awh et al., 2012; Fuster, 2003a; Petersen & Posner, 2012).

Sustained attention is likely driven by recurrent neural activity at some stage in the perception–action cycle. Attention does not only operate on perception of external stimuli, but can also be directed internally (Nobre, Coull, Maquet, & Frith, 2004). When internally directed, sustained attention maintains activity in perceptual networks without the need for external stimuli. Examples of internally directed attention are working memory and mental imagery (e.g., visualizing or imagining something from memory). Working memory is the interaction between attention and memory, without external input, for its prospective use.

Other concepts that have been associated with attention are arousal, vigilance, and alertness (Petersen & Posner, 2012; Posner & Petersen, 1990). These properties are strongly associated with subcortical structures such as the reticular formation (Moruzzi & Horace, 1949) and the cholinergic system (Perry, Walker, Grace, & Perry, 1999). Although arousal clearly interacts with attention (Fuster, 1958) I prefer to separate them. Whereas attention concerns the allocation of neural resources, arousal concerns the available quantity of neural resources, rather than the allocation per se.

The conception of attention as neural competition, excitation and inhibition, among neural networks (Desimone & Duncan, 1995; Franconeri et al., 2013; Fuster, 2003a) has the merit of working equally well within all systems, and at all levels of the nervous system. For example, neural competition from on/off-centers in retinal ganglion cells to distributed networks in visual associative cortex, and distributed networks in the prefrontal cortex to reciprocal spinal innervation of extensor and flexor muscles in limbs (Fuster, 2003a). It is also consistent with the load theory of attention, which holds that attentional resources are allocated by an exogenous perceptual selection mechanisms, and an endogenous attentional control mechanism (Lavie, Hirst, de Fockert, & Viding, 2004; Lavie & Tsal, 1994; Lavie, 1995). Further empirical support comes from cell recordings in primates (Chelazzi, Miller, Duncan, & Desimone, 1993; Moran & Desimone, 1985), and human neuroimaging (Rees, Frith, & Lavie, 1997; Rees, Rees, Russell, & Frith, 2012).

Attention, like memory, is thus a global property that exists in the whole brain, not only in a specific brain region, and can be differentiated based on informational content. For example, lesions to the right parietal cortex can lead to unilateral neglect, that is, a dysfunction in spatial attention. Patients with spatial neglect do not attend to their left side in space, and can for example miss shaving the left side of their face or leave the food on the left side of the plate (Corbetta & Shulman, 2002, 2011). Discrete lesions in the visual cortex will make it impossible to attend to (and perceive or remember) that specific visual content.

When we apply the various attentional terms to the cocktail party context we get the following. We can endogenously focus attention between single voices as we please, and thus consciously experience and understand its content. Although the garbled noise of voices in the background is barely noted, it is still consciously and/or non-consciously processed to some extent. When our name is non-consciously processed it can induce an exogenous shift to attend the voice that uttered your name, and thereby make it conscious.

Consciousness

It is still unclear whether consciousness is a function such as memory, perception, action, and attention, or if it is an epiphenomenon (i.e., without function - a fancy side effect of being). To make matters worse, there is no proper definition of consciousness that avoids circular reasoning (Block, 1995; Chalmers, 1996), and it is an ambiguous concept that can mean different things among laymen, philosophers, and scientists (Zeman, 2001). However, conscious experience is here assumed to be a property of the brain, and I will try to clarify what it refers to in this dissertation.

A common conceptual distinction is made between *states or levels* of consciousness, and the *content* of consciousness. The former refers to different global states of general arousal (e.g., coma, sleep, and wakefulness), while the latter refers to the subjective experience of being, that is, what it is like to experience something (e.g., the redness of red or the sensation of pain; Nagel, 1974). This is, however, merely a conceptual distinction, as the two are intimately linked (Hohwy, 2009). The content of consciousness varies, and can for example be perceptual (e.g., visual, auditory), executive (e.g., intentions, sense of agency), or metacognitive (i.e., cognitive states about other cognitive states) in nature. Traditionally, states or levels of consciousness have been depicted two-dimensionally as a function of the content of consciousness (Laureys, 2005). However, it has recently been suggested that global states of consciousness more aptly should be conceptualized as multi-dimensional states with content- and functionality-related dimensions (Bayne, Hohwy, & Owen, 2016). I tend to agree, and although a certain global state of arousal may be necessary to experience conscious content, it should not be conflated with it. In sum, the global state of arousal determines the amount of neural resources available, attention is the allocation of resources among competing networks, and conscious experience is the phenomenon that somehow ensues when networks receive enough neural resources.

This dissertation will mainly focus on the content of consciousness, and the use of *conscious*, *consciousness*, *phenomenal*, or *experience* will henceforth refer to when a subjective experience accompanies certain neural processes. Likewise, *non-conscious* will refer to the absence of subjective experiences (despite certain neural processes). The neural substrates of conscious experience are not known, and what the scientific study of consciousness is trying to find out.

The scientific study of consciousness

The scientific study of conscious experience began in the late 1800's with Fechner's psychophysics, Wundt's and Titchener's introspective approaches, and William James writings, but disappeared in the 1920's as behaviorism rejected all things subjective and unobservable (Hergenhan, 2001; Revonsuo, 2010). It was not until the late 1900's that consciousness began to be considered a respectable topic for scientific and philosophical research again (Baars, 1988; Chalmers, 1996; Crick & Koch, 1998; Nagel, 1974).

The scientific study of consciousness have since then mainly focused on finding the functional and neural correlates of consciousness (Crick & Koch, 2003; Koch, Massimini, Boly, & Tononi, 2016). The former tries to determine what functions, if any, are uniquely associated with consciousness, and to dissociate those functions (e.g., perception, action, attention, and memory) from conscious experience. The latter is concerned with finding the specific neural substrates underlying the conscious experience of certain content.

In both cases it has been common to use a simple subtraction logic when designing experiments. Such experiments usually have three conditions: (i) baseline trials without target stimulus, and therefore no conscious experience of the absent stimulus, (ii) trials with target stimulus, but without conscious experience of the stimulus because of some manipulation, and (iii) trials with target stimulus and conscious experience of the stimulus. Subtracting (ii) from (iii) will reveal correlates of conscious experience. Although successful, it has been pointed out that subtraction experiments nonetheless can be confounded by other functional properties that closely correlate with conscious experience, such as attentional control or the reporting of perceptual experience (Aru, Bachmann, Singer, & Melloni, 2011; Naghavi & Nyberg, 2005). Researchers are thus continuously trying to develop new ways to separate the neural correlates of conscious experience from possible confounds (e.g., no-report paradigms; Tsuchiya, Wilke, Frässle, & Lamme, 2015). Subtracting (i) from (ii) will reveal correlates of non-conscious processing. If certain functions and neuronal mechanisms are found to be associated with non-conscious processing, they can be inferred not to be neural substrates and functions unique to consciousness. It is therefore not only relevant to explore the functional and neural boundaries of non-conscious cognition for its own sake, but also to become informed about conscious experience.

Non-conscious cognition

The idea of non-conscious perception was perhaps first expressed by Leibniz in 1765, which he called *petites perceptions* (little perceptions). Leibniz thought these petites perceptions were too small to be experienced, but if enough of them were aggregated they would eventually pass a *limen* (i.e., threshold), and become conscious (Hergenhan, 2001; Schacter, 1987). The idea of a limen proved to be very influential in experimental psychology as investigations into subliminal (under the threshold or simply non-conscious) perception began in the late 1800's, and remains relevant today. The first experiments found support for non-conscious perception (Pierce & Jastrow, 1884; Sidis, 1898; Stroh, Shaw, & Washburn, 1908), and similar results followed until Eriksen (1960) criticized the use of subjective measures of conscious experience. The result was wide-spread skepticism about the existence of non-conscious perception. In the 1980's several studies using objective measures and process-dissociation procedures found evidence of non-conscious semantic priming. However, most of these studies were methodologically flawed (Holender, 1986). After further methodological and technical advances there was a consensus about the existence of non-conscious lexical and orthographical processing in the 1990's, and the controversy shifted from the existence towards the extent or limits of non-conscious cognition (for a historical review see Kouider & Dehaene, 2007). This was about the same time that serious efforts toward a scientific study of consciousness, and the search for neural correlates of consciousness began (Baars, 1988; Chalmers, 1996; Crick & Koch, 1998; Nagel, 1974). Since the 1990's it has been shown that non-consciously presented information can be processed at all levels throughout the visual system (Dehaene & Changeux, 2011; Dehaene, Charles, King, & Marti, 2014; Rees, Kreiman, & Koch, 2002). This suggests that nowhere in the visual system is neural activity in itself sufficient for having visual experiences.

Even more interesting is that studies have continued to push the limits of non-conscious cognition, and found that non-conscious information can engage increasingly more complex and flexible executive functions (some associated with prefrontal and parietal cortex). This suggests that neural activity in the prefrontal and parietal cortices might also not be sufficient in itself for conscious experience. For example, it has been shown that the monetary value of non-consciously presented images of coins can be processed in reward related brain regions (Pessiglione et al., 2007) and modulate working memory performance (Zedelius, Veling, & Aarts, 2011). Non-consciously presented information can engage various cognitive control and inhibitory mechanisms, and related frontal cortical regions (Lau & Passingham, 2007; Reuss, Kiesel, & Kunde, 2015; Reuss, Kiesel, Kunde, & Hommel, 2011; van Gaal, Lamme, Fahrenfort, & Ridderinkhof, 2011; van Gaal, Ridderinkhof, Scholte, & Lamme, 2010). Error-monitoring has been shown to occur non-consciously by measuring post-error slowing in reaction times despite participants not being conscious of making the errors (Logan & Crump, 2010; Pavone, Marzi, & Girelli, 2009). Simpler arithmetic tasks such as adding two single digits (Ric & Muller, 2012), and adding or subtracting three single digits (Sklar et al., 2012) seems to be performed non-consciously. Indeed, it seems like most if not all cognitive processes can be engaged non-consciously to some extent. The difference between conscious and non-conscious processes does not seem to be one of kind, but rather of degree. That is, non-conscious neural and behavioral effects tend to be small compared to its conscious counterpart, but to what extent that depends on conservative experimental conditions is unclear.

Neural correlates of consciousness

It is still somewhat unclear what the neural correlates of consciousness are, but there has been considerable progress, especially in the visual system. At the lowest level of the visual system, in the retinae, different wave-lengths of light are received from the external environment, and transformed into neural signals, which propagates through the visual system, and ultimately, somehow, causes the vivid visual experiences we have every day. Activity in the retinae does not seem to correlate with conscious perception. Patients without functioning retinae experience phosphenes (flashes of light) when transcranial magnetic stimulation (TMS) is used on the visual cortex (Koch, 2004). Furthermore, the retinal "blind spots" where nerve fibers pass through the retinae are not visually experienced as blind spots (Koch, 2004). The retinae are therefore not necessary for conscious perception. The next stop in the visual system is the lateral geniculate nucleus (LGN) in the thalamus, which is a relay station between the retinae and the primary visual cortex. The LGN is the earliest stage of visual processing that correlates with conscious perception. During binocular rivalry BOLD signal change in eye-specific regions increase when that eye's input is conscious, and decrease when suppressed (Haynes, Deichmann, & Rees, 2005; Wunderlich, Schneider, & Kastner, 2005). The LGN does not seem to be necessary since TMS to the visual cortex still induce phosphenes. However, it is possible that TMS induced cortical activity is fed back to the LGN.

The LGN relays 90% of the visual nerve projections to the primary visual cortex, where finer details are processed, such as line orientation and length, motion direction, color, contrast, spatial frequency, and ocular dominance (Tong, 2003). Neural activity in the primary visual cortex is not sufficient for visual experience because there are several cases of non-conscious visual processing in the primary visual cortex (Haynes & Rees, 2005; He & MacLeod, 2001; Jiang, Zhou, & He, 2007; Sergent, Baillet, & Dehaene, 2005; Zou, He, & Zhang, 2016). Lesions in the primary visual cortex disrupt visual experience, and can lead to cases of *blindsight*, in which patients perform better than chance on forced-choice discriminations, without visual experience of the stimuli (Cowey, 2010; Leopold, 2012). In addition, patients with lesions in higher visual areas, but intact primary visual cortex, do not have visual experiences (Horton & Hoyt, 1991).

It is still a matter of debate whether the primary visual cortex is directly necessary for visual experience, or indirectly necessary as an information gateway to higher visual regions (Crick & Koch, 1995; Silvanto, 2014; Tong, 2003). Neural activity in the primary visual cortex of primates does not seem to correlate with conscious perception, in contrast to activity in higher visual regions that do correlate with visual experience (Leopold, 2012; Logothetis, 1998). BOLD signal change in the human primary visual cortex sometimes correlate with conscious perception (Harrison & Tong, 2009; Lee, Blake, & Heeger, 2005; Polonsky, Blake, Braun, & Heeger, 2000; Ress & Heeger, 2003), but not when controlling for attention (Bahrami, Lavie, & Rees, 2007; Lee, Blake, & Heeger, 2007; Tse, Martinez-Conde, Schlegel, & Macknik, 2005; Watanabe et al., 2011). Other studies suggest that the primary visual cortex might be directly necessary for visual experience when information is fed back from higher visual regions (Koivisto, Mäntylä, & Silvanto, 2010; Silvanto, Cowey, Lavie, & Walsh, 2005). It is therefore still unclear in what capacity the primary visual cortex is necessary for visual experience.

Visual information is relayed from the primary visual cortex to higher visual regions where information, such as color, motion, spatial position, faces, and other aspects are processed in relatively specialized regions. (for review see Grill-Spector & Malach, 2004). The higher visual regions follow a ventral stream that is important for object recognition in the inferior temporal cortex (Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2012), and a dorsal stream toward the parietal cortex that processes spatial information in relation to visually guided action and navigation (Kravitz, Saleem, Baker, & Mishkin, 2011; Whitwell, Milner, & Goodale, 2014). Cell-recordings in primates, and human neuroimaging studies converge on the finding that activity/signal in specialized visual regions correlate with visual experience of the same content (Koch, Massimini, Boly, & Tononi, 2016; Logothetis, 1998; Lumer, Friston, & Rees, 1998; Moutoussis, Keliris, Kourtzi, & Logothetis, 2005; Rees, Kreiman, & Koch, 2002; Rees, 2007; Zeki, Watson, & Frackowiak, 1993). Lesions to specialized visual regions eliminates the possibility of visual experiences of such content (Barton, 2011). In addition, activity in higher visual cortex correlates with visual experience in the absence of external stimuli, in cases as visual imagery (Kreiman, Koch, & Fried, 2000; O'Craven & Kanwisher, 2000), hallucinations (Ffytche et al., 1998), and illusions (Zeki et al., 1993). For example, disrupting the region that processes motion with TMS will disrupt the illusion of motion (Ruzzoli et al., 2011). The higher visual cortex is therefore necessary for visual experience, but not sufficient. Because it is now known that non-conscious visual processing can occur throughout the higher visual regions (Dehaene et al., 1994; Dehaene & Changeux, 2011; Fang & He, 2005; Marois, Yi, & Chun, 2004; Moutoussis & Zeki, 2002).

Additionally, most previous studies consistently found that neural activity and BOLD signal change in prefrontal and parietal cortex correlated with conscious perception (Dehaene & Changeux, 2011; Naghavi & Nyberg, 2005; Rees et al., 2002; Rees, 2007; Tononi & Koch, 2008), which has been the foundation of some prominent theories of consciousness (for more details see "implications for theories of consciousness" in the discussion). However, the latest efforts in trying to disentangle the neural correlates of consciousness from other closely related processes (Naghavi & Nyberg, 2005) have found that BOLD signal in the prefrontal cortex likely is associated with attention (Eriksson, Larsson, & Nyberg, 2008; Tse et al., 2005), and reporting the presence of visual experience (Frässle, Sommer, Jansen, Naber, & Einhauser, 2014) rather than visual experience per se. Frässle et al. (2014), elegantly, used the eye's nystagmus reflex and pupil size as a proxy for explicit reports of conscious experience in a binocular rivalry paradigm. One of two competing stimuli (dark red or light green bars moving in opposite directions) were presented to each eye, the nystagmus reflex correlated with the motion, and the pupil size with the brightness, of the consciously experienced stimulus. Using fMRI they showed that the BOLD signal in frontal cortex did not correlate with conscious perception without explicit reports, and concluded that frontal cortex is related to introspection and action rather than sensory experience. It therefore seems like the frontal cortex might not be necessary for conscious sensory experience.

Attention and conscious experience

Attention and conscious experience are intimately linked, and it was previously thought to be the same function (William, 1890). However, it is now widely acknowledged that attention and conscious experience are two different properties of the brain. The debate now revolves around their relationship, and whether attention and conscious experience can be dissociated from each other. Attention is often used synonymously with endogenous attention, but I will separate attention from endogenous and exogenous attention. Here attention will refer to resource allocation by any mechanism. The latter two being different mechanisms for allocation of a shared neural resource. Some argue that endogenous attention is necessary, but not sufficient, for conscious experience (Cohen, Cavanagh, Chun, & Nakayama, 2012; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006), while others argue that endogenous attention is neither necessary nor sufficient for conscious experience (Koch & Tsuchiya, 2007; Lamme, 2004).

It is clear that attention modulates visual experience. Covert exogenous attention affects the visual experience of stimuli in terms of enhanced contrast (Carrasco, Ling, & Read, 2004), spatial resolution and acuity (Gobell & Carrasco, 2005), and color (saturation, but not hue; Fuller & Carrasco, 2006). Similarly, covert endogenous attention can modulate the conscious experience of brightness (Tse, 2005), contrast (Liu, Abrams, & Carrasco, 2009), and spatial resolution (Abrams, Barbot, & Carrasco, 2010). However, it is still unclear whether endogenous and exogenous attention can have different effects on consciousness.

Exogenous attention can be engaged by non-conscious information. Non-consciously presented nude pictures of the opposite sex will automatically shift attention towards them, and for males also away, if the nude pictures are of the same sex (Jiang, Costello, Fang, Huang, & He, 2006). Endogenous attention to cues presented in blindsight patients' "blind spot" speeds up their responses (Kentridge, Heywood, & Weiskrantz, 2004; Kentridge, Nijboer, & Heywood, 2008). It has also been shown that cuing exogenous spatial attention toward non-consciously presented stimuli can increase non-conscious priming effects (Marzouki, Grainger, & Theeuwes, 2007).

Conscious endogenous attention toward non-conscious Gabor gratings flanked by distractors (i.e., "crowding") in the peripheral visual field does not necessarily make the orientation of the gratings lines conscious (He, Cavanagh, & Intriligator, 1996). Endogenous focal attention to conscious object-features automatically modulates non-conscious stimuli outside the focus of attention that shares the same features (Kanai, Tsuchiya, & Verstraten, 2006; Melcher, Papathomas, & Vidnyánszky, 2005). It has also been shown that endogenous attention and conscious experience can have opposite effects on the duration of afterimages (van Boxtel, Tsuchiya, & Koch, 2010). However, that conscious endogenous attention can modulate non-conscious stimuli without making them conscious, is not the same as non-conscious endogenous attention. Although attention was not of primary interest, cases of non-consciously engaged cognitive control/inhibition are examples of non-conscious endogenous motor attention (Lau & Passingham, 2007; Reuss et al., 2011; van Gaal et al., 2010). In such experiments the task is to perform one of two actions depending on which of them are cued. The non-consciously presented cue activate (attend) parts of the cognitive control system, and affects motor output if the non-conscious cue is incongruent with the performed task, and therefore constitutes non-conscious endogenous motor attention.

More controversially, Koch & Tsuchiya (2007) have argued that there are cases of conscious experience without endogenous attention, such as the gist of quickly presented scenes (Fei-Fei, Iyer, Koch, & Perona, 2007; Mack & Rock, 1998), and discrimination of stimuli in the periphery during dual tasks (Li, VanRullen, Koch, & Perona, 2002; Reddy, Reddy, & Koch, 2006). However, discrimination performance has been shown to decrease if endogenous attention is taxed enough when trying to determine the gist of scenes (Cohen, Alvarez, & Nakayama, 2011; Mack & Clarke, 2011; Marois et al., 2004; Yi, Woodman, Widders, Marois, & Chun, 2004), and other stimuli during dual tasks (Evans & Treisman, 2005; Joseph, Chun, & Nakayama, 1997; Landau & Bentin, 2008; Walker, Stafford, & Davis, 2008). Cohen et al., (2012) therefore argue that endogenous attention is necessary for consciousness, but that most tasks do not fully engage all attentional resources, and differences in task difficulty can explain why conscious experience occur to varying extents while endogenous attention is directed at other tasks. In essence these experiments are in line with the load theory of attention (Lavie et al., 2004), and show that if there are attentional resources to spare, they can be exogenously allocated, and if enough is allocated it becomes conscious. Otherwise the information is non-consciously processed (Chun & Marois, 2002), or not processed at all. The studies also show that the focus of attention is not necessary for conscious experience, since information outside the focus can be experienced in some circumstances.

In sum, attention is necessary (but not sufficient) for, and can modulate, conscious and non-conscious processes. Exogenous and, importantly, endogenous attentional allocation can be engaged by non-conscious information. The fact that endogenous attention can be engaged by non-conscious information is a prerequisite for non-conscious working memory.

Memory and conscious experience

From Ebbinghaus' work in 1885 to 1960 many believed there was only one unitary memory system. However, William James (1890) argued that there was two memory systems, which he called primary and secondary memory. According to James, the primary memory lasted only tens of seconds, while the secondary memory could last for a life time. It was not until 1960 that new findings began to support shorter memories (Palmer, 1999). Those shorter memories were later to be called *sensory memory*, and *short-term* or *working memory* instead of primary memory, in contrast to *long-term memory* instead of secondary memory. In the context of this dissertation I will be using short-term memory as an umbrella term for any memory that lasts a short time, and not synonymously with working memory.

Sensory memory and working memory have traditionally been conceived as inherently conscious, in contrast to long-term memory, which was thought to be non-conscious while not being actively recollected. Long-term memory can be divided into three epochs; encoding, storage, and retrieval. Studies of long-term memory have resulted in various categories. These categories have traditionally been based on whether the participant/patient could consciously recollect the encoding event during the retrieval epoch or not. The taxonomy based on conscious recollection contained two main categories: *explicit* and *implicit* memories. Explicit memory initially included *episodic* and *semantic memory*, while implicit memory includes *priming*, *conditioning*, and *procedural memory* (Squire & Zola-Morgan, 1988). However, Tulving (2002) has suggested that semantic memory more aptly should be considered an implicit memory since conscious recollection of the encoding event is not necessary. Although implicit memory is seemingly synonymous with non-conscious memory, an important distinction is made between the usage of implicit and non-conscious here forth. Contrary to explicit memory, implicit memory does not require conscious recollection during retrieval (Graf &

Schacter, 1985; Schacter, 1987). Implicit memory can therefore be consciously or non-consciously encoded. However, for what I call non-conscious memory, it is necessary that the information to be remembered is non-consciously encoded or perceived.

Not all these memory systems are relevant for this dissertation, and I will therefore only elaborate on a subset of them, namely, iconic memory, working memory, episodic/semantic memory, and priming. Although we intend to investigate non-conscious short-term memory, we need to consider the possibility that long-term memory mechanisms can underlie memory retention within short-term memory tasks as well.

Priming

Repetition priming (henceforth priming) is a well-studied phenomenon, and is regarded as an implicit memory (for historical review see Schacter, 1987). Priming is defined as *a change in task performance (e.g., accuracy or reaction time) of a stimulus as a result of prior experience to an identical or similar stimulus*. Priming is indexed by comparing the difference between two conditions where (i) a stimulus (i.e., a prime) was presented prior to an identical (or similar) target stimulus, and (ii) where the prime and target were different. The difference between matching and mismatching prime and target is the so-called *priming effect*. As an implicit memory, priming is something we rarely notice outside of the laboratory, but most certainly something that happens most of the time. Priming can occur when information is consciously or non-consciously encoded. I will refer to the former case as consciously encoded priming, and the latter as non-consciously encoded, or simply non-conscious, priming. These two types of priming, as we shall see, differ significantly in memory strength, but it is primarily non-conscious priming that is of interest in this dissertation.

Priming is considered a separate kind of memory based on dissociations between priming and other memories, such as, episodic/semantic memory, in neurophysiological lesions (Graf, Squire, & Mandler, 1984; Warrington & Weiskrantz, 1974), as well as, psychological studies (for review see Tulving & Schacter, 1990). Cell-recordings in primates and human neuroimaging have found that priming is tightly linked to a phenomenon called repetition suppression (Henson, 2003; Schacter, Dobbins, & Schnyer, 2004). Repetition suppression is when repeated exposure to the same stimulus results in a decrease in neural activity or BOLD signal the second time the stimulus is processed. There are several suggested models that try to explain how priming (a facilitation in performance) can lead to a decrease in neural activity/BOLD signal (Grill-Spector, Henson, & Martin, 2006). However, there are also cases of repetition enhancement (Henson, 2000; Tartaglia, Mongillo, & Brunel, 2015). Complicating matters more, it has been argued that repetition suppression is not caused by priming, but a form of antipriming. Antipriming is when the first stimulus (prime) is not followed by an identical target stimulus, but instead, a different stimulus with overlapping features (Marsolek, 2008). Antipriming is essentially the opposite of priming, that is, a decrease in performance. When comparing baseline (i.e., no or little priming) trials with primed and antiprimed trials there was no difference between baseline and primed, but an increased BOLD signal for antiprimed trials (Marsolek et al., 2010). It was therefore argued that repetition suppression results from comparing primed with antiprimed trials, since priming experiments usually lack baseline trials.

Repetition suppression has been used to map out the neural correlates of priming in the brain, and by using various material in priming experiments it has been found that different areas in the brain are involved with priming of different material (Henson, 2003; Schacter, Wig, & Stevens, 2007). For example, visual priming depends on the visual cortex (Koutstaal et al., 2001), semantic priming on the temporal cortex (Rossell, Price, & Nobre, 2003), affective priming on the amygdala (Dannlowski et al., 2007), and contextual priming on the hippocampi (Greene, Gross, Elsinger, & Rao, 2007). These material-dependent findings have been confirmed by dissociations between priming of different material, such that, lesions or atrophy (e.g., from dementia) in specific brain regions causes content-specific priming deficits, while sparing priming of other content (Schacter & Buckner, 1998).

Consciously encoded perceptual priming (e.g., visual images) are known for being very robust, and can exist many years after just a single exposure (Cave, 1997; Mitchell, 2006; van Turennout, Ellmore, & Martin, 2000). Perceptual priming is therefore mainly thought to be latent long-term changes in neural networks. However, the definition of priming is quite broad, and could equally well correspond to short-lived latent neural changes, residual or persistent neural activity. Indeed, (conscious) working memory tasks can elicit very strong priming effects (e.g., ~ 250 milliseconds) in reaction times that likely depend on persistent activity rather than latent neural changes (e.g., in Study II & III of the present dissertation), and popular models of semantic priming is posited to be residual activity because it only last for seconds (which qualifies as iconic memory, see below). Thus, priming seems to occur in almost any brain region depending on material, consist of any possible neural memory mechanism, and thus partly or wholly overlap with most other kinds of memory. It has been said that a *"useful concept in science frequently is one whose definition not only makes very clear what it includes, but also what it excludes"* (Tulving, p. 384, 1972). I therefore believe that priming, as currently defined, is too inclusive to be a useful concept as a specific kind of memory, and more useful as a general memory phenomenon.

Non-consciously encoded priming is somewhat weaker than its conscious counterpart. Historically, as mentioned earlier, it was assumed that non-conscious retention was the result of priming, and that these non-conscious priming effects ceased to be detectable within 500 ms (Dehaene & Changeux, 2011; Draine & Greenwald, 1998; Greenwald, Draine, & Abrams, 1996; Mattler, 2005). However, there are examples of non-consciously encoded perceptual priming effects lasting for 15 - 20 min (Bar & Biederman, 1998, 1999), and neural repetition suppression effects without behavioral effects after 47 min (Gaillard et al., 2007). It is possible that the discrepancy in non-conscious priming longevity mirrors the difference between consciously encoded perceptual (years) and semantic (seconds) priming to some extent. Given that non-conscious perceptual priming can last minutes in some cases it is necessary to consider it as a potential explanation for short-term retention of several seconds as well.

Episodic & semantic memory

Episodic/semantic memory is something we make use of all the time in our everyday lives, and probably what we most commonly think of when we hear the word memory. Episodic memory is called “explicit” because we are consciously aware of the encoding event when we recall it. Semantic memory is “implicit” because it is not necessarily accompanied by conscious recollection of the encoding event, even if there might be conscious experience of knowing the semantic information. I will nonetheless treat them together here because they share many similarities.

Although episodic and semantic memory have several things in common, there is a clear distinction to be made between them (Tulving, 1972, 1983). Episodic memory refers to our memory of events that have been experienced personally, at a specific place, and at a specific point in time. Semantic memory refers to our knowledge of the world, such as, facts or language. They are, however, intimately linked in usage and cortical organization. Furthermore, they are highly dependent on subcortical structures in the medial temporal lobe, primarily the hippocampi structures. Lesions to the medial temporal lobes can cause complete episodic and severe semantic anterograde amnesia (i.e., the inability to form new long-term memories), and to some extent retrograde amnesia (i.e., the inability to recall previous long-term memories), but leaves intelligence, short-term, and implicit memory largely intact, as demonstrated by the famous amnesia patients H.M. (Corkin, 2002; Scoville & Milner, 1957), and K.C. (Rosenbaum et al., 2005). H.M.’s lesions caused partial retrograde amnesia in semantic and episodic memory, but K.C.’s lesions caused complete episodic retrograde amnesia, while sparing semantic memories. H.M.’s episodic retrograde amnesia only seemed to stretch 5 years back from the operation initially, but it might have fully degraded over time (Corkin, 2002; Rosenbaum et al., 2005). Interestingly, it was later shown that H.M., and K.C. had learned new semantic information, albeit extremely slowly across many repetitions, without medial temporal lobes (Hayman, Macdonald, & Tulving, 1993; O'Kane, Kensinger, & Corkin, 2004; Tulving, Hayman, & Macdonald, 1991).

Human neuroimaging studies have found that activity in the hippocampi during encoding predicts subsequent recollection performance (Danker, Tompary, & Davachi, 2016; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Nyberg, McIntosh, Houle, Nilsson, & Tulving, 1996). The medial temporal structures are therefore clearly essential for normal episodic/semantic memory function, but so is the neocortex.

Episodic memories are stored in widely distributed networks across the neocortex that connects various sensory information, time, and place throughout the neocortex (Cabeza & Nyberg, 2000; Fuster, 1995, 2009; Habib, Nyberg, & Tulving, 2003; Nyberg et al., 1997; Nyberg, Habib, Mcintosh, & Tulving, 2000; Nyberg, McIntosh, Cabeza, et al., 1996). Different parts of the neocortex therefore contribute to different aspects or contents of episodic memories. For example, visual information is stored in the visual cortex, auditory information in the auditory cortex, spatial information in the parietal cortex, and temporal information relies on the prefrontal cortex. Semantic memory is mainly stored left lateralized in the temporal, parietal, but also in the frontal cortex (Binder, Desai, Graves, & Conant, 2009; Binder & Desai, 2011; Cabeza & Nyberg, 2000; Fuster, 1995, 2009; Mårtensson et al., 2012). Severe lesions to specific cortical regions therefore renders specific content (e.g., faces, names, when or where events took place) of those specific regions unusable, but does not affect other aspects of episodic/semantic memory.

The widely distributed networks in the neocortex form their connections relatively slowly, and it is largely believed that the hippocampi, which form new connections relatively quickly, holds the cortical networks together until they become strong enough for independence. However, there is still considerable debate about the specifics of the interaction between the medial temporal lobe and neocortex (Nadel & Moscovitch, 1997; Squire, Kosslyn, Zola-Morgan, Haist, & Musen, 1992; Winocur & Moscovitch, 2011).

As previously mentioned, it was traditionally assumed that episodic (and by some) semantic memory always was consciously recollected, and that the hippocampi, which these structures heavily depended on, also were linked to consciousness (Squire & Zola-Morgan, 1988; Squire & Dede, 2015; Tulving, 2002). However, this view was challenged by a series of recent studies from Henke's lab that showed hippocampi-based retention of non-consciously encoded semantic associations (Degonda et al., 2005; Duss et al., 2014; Duss, Oggier, Reber, & Henke, 2011; Henke, Mondadori, et al., 2003; Henke, Treyer, et al., 2003; Reber, Luechinger, Boesiger, & Henke, 2012; Reber & Henke, 2011). These findings are consistent with semantic memory being an implicit memory (Tulving, 2002).

For example, it was found that non-conscious semantic associative learning engaged the hippocampi, and interacted with conscious memory (Degonda et al., 2005). Reber et al. (2012) showed that BOLD signal change in the hippocampi during non-conscious encoding predicted memory performance 1 minute later. In a study with a patient group with hippocampi lesions (amnesiacs), and healthy controls, Duss et al. (2014) found that only healthy controls could retain word pair associations (semantic memory), but both groups could retain a single non-conscious word (semantic priming). These findings show that hippocampi structures are necessary for semantic associative learning with or without conscious experience. Taken together, these, alongside other findings, have been used to criticize the traditional long-term memory taxonomy based on conscious recollection (Henke, 2010).

Semantic memory can thus be encoded by non-consciously presented information. However, the same cannot be said for episodic memory. Partly because conscious recollection is a part of the definition of episodic memory, and it seems unlikely to have a conscious recollection of something non-consciously encoded. Although, if a non-conscious memory fulfills all criteria of episodic memory except conscious recollection, I am inclined to think that the definition of episodic memory needs a revision. As far as I know, nothing like a non-conscious episodic memory has been shown to exist. However, that does not exclude its existence, and I will therefore use non-conscious hippocampi-based memory to refer to non-conscious semantic memory and/or memory of "episodic" information. Non-conscious hippocampi-based memory therefore needs to be considered as a potential explanation to short-term retention.

Sensory memory

Experiments prior to 1960 usually presented participants with brief (100 milliseconds) arrays of up to 20 letters, and found that participants only could remember a few letters (full-report paradigms). However, Sperling (1960) later added a tone after an array of three rows with four letters each presented for 50 milliseconds. The tone was presented after stimuli offset, and used to cue which row the participants had to report letters from (high, medium, low). Sperling found that participants could correctly recall about four letters from any cued row (partial-report paradigm). His findings show that it was possible to recall much more than a few letters, but only for a brief time, otherwise the memory had faded about the time it took to report four letters. Furthermore, it was found that this sensory memory had a capacity up to 16 out of 18 letters, but gradually decayed rapidly to an asymptote of four letters within 500 ms to 2 s after stimuli offset (depending on brightness), and that if other stimuli were subsequently presented at the same spatial position the memory would be overwritten (Averbach & Coriell, 1961; Averbach & Sperling, 1961). This brief visual sensory memory was later named *iconic memory* (Neisser, 1967).

Iconic memory was defined as the *retention of information* commonly measured by the partial-report paradigms (Averbach & Sperling, 1961; Coltheart, 1980; Neisser, 1967).

Iconic memory was traditionally thought to consist of three components, visible persistence (i.e., persisting conscious experience), neural persistence (i.e., residual neural activity), and information persistence (i.e., iconic memory). It was furthermore assumed that these three components described the same phenomenon at different levels. However, it has aptly been argued, that visual persistence, which shows inverse relationships with stimuli duration and luminance (if the duration is brief), can be dissociated from iconic memory, which does not show inverse relationships (Coltheart, 1980). Coltheart thinks it is likely that photoreceptor persistence in the retina can explain the inverse relationship with luminance, and residual activity in the LGN might explain the inverse relationship with stimuli duration, but residual cortical activity is likely also necessary. It is less clear how the neural substrates of iconic memory differ from visible persistence, since both must rely on the visual system. Nevertheless, since visible persistence can be dissociated from iconic memory, it seems plausible that there can be such a thing as non-conscious iconic memory. However, if cued prior to its complete disappearance, the information can be maintained within the limited capacity of working memory until probed. It might be possible that techniques can be used to render iconic memory non-conscious from the beginning without losing all the information.

In a study Sligte, Scholte, & Lamme (2008) proposed the existence of a new *fragile visual short-term memory* based on a dissociation from iconic memory and working memory, which by some have been criticized for actually being working memory (Makovski, 2012). However, I will briefly explain why I think they mistakenly confused “iconic memory” with visible persistence, and “fragile memory” with iconic memory. Sligte et al. (2008) are essentially using a modified partial-report paradigm with a retro-cue after 1 s, which they argue is the reason why it is not iconic memory (because of iconic memory’s rapid decay). Firstly, however, they are by (Sperling’s, Neisser’s, and Coltheart’s) definition, measuring iconic memory (i.e., the informational persistence), thus any memory effect is by definition an iconic memory effect. Secondly, they have a black or grey background after stimuli offset; according to Averbach & Sperling (1961) a black context can make iconic memory persist for more than 2 seconds. Thirdly, there are other differences that could affect memory retention, such as stimuli duration (five times longer than Sperling’s duration), the delayed match-to-sample task (recognition instead of free recall), and less complex stimuli (orientation of rectangles instead of letters), could facilitate performance. Thirdly, they show that fragile memory is not affected by isoluminant (same luminance but different colors) stimuli and masks of uniform bright color, but is erased by masks with irrelevant stimuli at same spatial position, which is all consistent with iconic memory (Averbach & Coriell, 1961; Coltheart, 1980). Fourthly, their "iconic memory" is affected by isoluminant stimuli and bright masks of uniform color, which is to be expected of visible persistence (Coltheart, 1980). The novelty, I think, is that they, in their fourth experiment, show that "fragile memory" (i.e., iconic memory) can last up to 4 seconds (with gradually reducing capacity over time) in some circumstances. I will therefore interpret subsequent studies of "fragile memory" as showing that iconic memory performance correlates with BOLD signal change in visual cortex (V4; Sligte, Scholte, & Lamme, 2009), does not depend on dorsolateral prefrontal cortex (Sligte, Wokke, Tesselaar, Scholte, & Lamme, 2011), and is erased only if irrelevant masks are of same spatial location and object type (Pinto, Sligte, Shapiro, & Lamme, 2013).

It therefore seems like the neural substrate of iconic memory is limited to residual neural activity in higher visual cortical areas, but it is unclear how it differs from visible persistence. Another implication is that backward masking might erase visible persistence, but not necessarily iconic memory, which means that non-conscious iconic memory is plausible, and might even explain non-conscious semantic priming effects. It might be the case, but has yet to be investigated, that non-consciously presented information can be retained in a rapidly decaying high capacity iconic memory that can last for a few seconds.

Working memory

Working memory is something we use daily when we reason, make decisions, and perform tasks in general. It is for example used to keep a phone number in mind until the number have been dialed, when doing arithmetic in the head, or when planning a future decision. It is known for having a relatively small capacity limit (Brady, Konkle, & Alvarez, 2011; Keisuke Fukuda, Awh, & Vogel, 2010), and is a reliable indicator of intelligence (Fukuda, Vogel, & Mayr, 2010; Unsworth, Fukuda, Awh, & Vogel, 2014). There might not be one definition of working memory that all agree on, but I will use the following: Working memory is *"...the temporary retention of information - sensory or other – for the performance of a prospective act to solve a problem or attain a goal"* (Fuster, p. 144, 2015). That the retention is *for* a prospective action is crucial, as it is the defining feature that separates working memory from sensory memory. Neurophysiological investigations have revealed that working memory is an emergent property from the interaction between long-term (phylogenic and ontogenetic) memory and sustained endogenous attention (for reviews see Eriksson, Vogel, Lansner, Bergström, & Nyberg, 2015; Fuster, 1995; Postle, 2006; Sreenivasan, Curtis, & D'Esposito, 2014). That is, the temporary activation of neural networks until a prospective act is complete. Similarly to long-term memory and perception, different brain regions represent different aspects of working memory, and partly depend on the specific task and material used (for meta analyses see Fuster, 2009; Nee et al., 2013; Wager & Smith, 2003). It is well established that the prefrontal cortex is important for working memory from studies with cell-recordings in primates (Funahashi, Bruce, & Goldman-Rakic, 1989; Fuster & Alexander, 1971; Goldman-Rakic, 1995), human neuroimaging (Courtney, Ungerleider, Keil, & Haxby, 1997; Postle & D'Esposito, 1999), lesions (D'Esposito & Postle, 1999), electrical, or magnetic stimulations (Brunoni & Vanderhasselt, 2014; Feredoes, Heinen, Weiskopf, Ruff, & Driver, 2011).

The role of the prefrontal cortex seems to be maintaining task-related information (Pochon et al., 2001; Rowe et al., 2007; Sakai & Passingham, 2003), and protect the remembered information from distractions (Feredoes et al., 2011; Sakai, Rowe, & Passingham, 2002). The prefrontal cortex seems to be involved no matter what kind of information that is being retained, but with slight variations in cortical location (Nee et al., 2013; Wager & Smith, 2003). For example, the frontal eye fields (also involved in coordinating eye movement) are involved in maintenance of visuospatial information (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998), while the left inferior frontal cortex (also involved in speech production) is involved in maintenance of verbal information (Paulesu, Frith, & Frackowiak, 1993), but the most anterior prefrontal cortex seems less dependent of material.

The posterior cortex is also important, but is much more dependent on the information held in working memory. That is, the visual cortex will be active when visual information is being maintained, the auditory cortex for sound, parietal cortex for spatial information, etc. (Druzgal & D'Esposito, 2001; Emrich, Riggall, Larocque, & Postle, 2013; Fuster & Jervey, 1981; Fuster, 2009; Harrison & Tong, 2009; Miller & Desimone, 1994; Miyashita & Chang, 1988; Postle, Stern, Rosen, & Corkin, 2000; Pratte & Tong, 2014). Locations of lesions therefore differentially affect working memory depending on location/content (Owen et al., 1996; Pisella, Berberovic, & Mattingley, 2004). Additionally, the parietal cortex seems involved in more executive aspects of working memory (Collette et al., 2005; Koenigs, Barbey, Postle, & Grafman, 2009), and have been associated with working memory capacity (Riggall & Postle, 2012), but the latter could also reflect spatial strategies to simultaneously maintain many items.

These prefrontal and posterior brain areas seem to causally depend on each other for a normally functioning working memory. For example, cortical cooling of either prefrontal or stimulus-specific regions in primates cause a decrease in neural activity in the other region, together with a drop in working memory performance (Fuster, Bauer, & Jervey, 1985). Other brain areas relevant for working memory are the hippocampi (Libby, Hannula, & Ranganath, 2014; Ranganath & D'Esposito, 2001), basal ganglia (Cools & D'Esposito, 2011; Ekman et al., 2012; Wager & Smith, 2003), and cerebellum (Desmond, Chen, & Shieh, 2005; Desmond & Fiez, 1998; Stoodley & Schmahmann, 2009).

It should be noted that all the neurophysiological findings mentioned above report persistent neural activity during working memory maintenance. However, recent computational models and empirical findings suggest that the “persistent” part of the neural activity might partly or sometimes be an artifact created from averaging data over trials, and that short-lived latent neural changes instead might play an important role to maintain the memory trace between activity (Barak & Tsodyks, 2014; D’Esposito & Postle, 2015; Eriksson et al., 2015; Lundqvist, Herman, & Lansner, 2011; Mongillo, Barak, & Tsodyks, 2008; O’Reilly, Braver, & Cohen, 1999; Sprague, Ester, & Serences, 2016). These models might produce a more stable and economical working memory maintenance. By using trial-by-trial analyses, it has been argued that the intensive neural activity that occurs during perception encodes the information in short-lived latent (structural) changes, which is then retained by sporadic bursts of neural activity after stimulus offset, and again intensifies during the use of the information (Lundqvist et al., 2015).

Sprague et al. (2016) found that the fidelity of working memory content decreased when more information was maintained, but that a retro-cue indicating that only a subset of the maintained information was relevant retroactively increased the fidelity of the relevant information. In addition, the retroactive increase in fidelity correlated with memory performance. The recovery in fidelity is inconsistent with pure spiking models because they predict that lost fidelity cannot be recovered. It was therefore concluded that some latent neural change must have facilitated the recovery in fidelity/neural activity. Lewis-Peacock, Drysdale, Oberauer, & Postle (2012) used multivoxel pattern classification analysis of to decode the BOLD signal from prioritized and deprioritized information during a working memory task. Interestingly, they could only decode the prioritized, but not the deprioritized information. The deprioritized information was nonetheless remembered when later probed (on half of the trials). These results was later replicated with electroencephalography (EEG; LaRocque, Lewis-Peacock, Drysdale, Oberauer, & Postle, 2013). It might therefore be the case that working memory more accurately is an emergent property based on the interaction between long-term (phylogenic and ontogenetic), short-lived (ontogenetic) memory, and sustained endogenous attention.

Briefly summarizing and integrating these neurophysiological findings about working memory with the framework espoused in the building blocks of cognition section leaves us with the following view. Presented stimuli are processed in stimulus-specific sensory areas when perceived. Stimulus-specific information is then retained by persistent neural activity caused by recurrent loops between stimulus-specific areas, and prefrontal action-related areas within the perception-action cycle. Persistent neural activity in posterior areas mainly relate to the information to be remembered, in the prefrontal cortex it relates to the maintenance of the prospective task (e.g., goal and plan), and coordination of attentional resources (endogenous attentional control). The activity gradually shifts towards the primary motor cortex as the plan and motor action is implemented, and after task completion the neural activity subsides.

William James (1890) made a distinction between primary and secondary memory based on phenomenological grounds. Primary memory was consciously experienced, while secondary memory was not. The relationship between conscious experience and working memory carried over to Atkinson & Shiffrin's (1968, 1971) model, which contained a sensory register (iconic memory), a short-term store that was conscious (working memory), and a long term store. Psychological evidence for multiple memory stores led Baddeley & Hitch (1974) to postulate a multi-store model, with a “central executive” that controls and coordinates information in three short-term stores (visuospatial sketchpad, phonological loop, and later episodic buffer; Baddeley, 1983, 2003, 2012). Baddeley's (1992, 2000, 2003) model initially assumed that all working memory operations, input and output, was conscious, but later introduced an episodic buffer that was similar to Baars (1988, 2005) global workspace, and integrated information from other stores into conscious episodes. Baars & Franklin (2003) thought all active working memory operations were consciously experienced.

These early *systems-based* working memory models posited separate short-term and long-term systems or stores based on neurophysiological evidence that lesions can affect long-term, but not short-term memory, and vice versa (Scoville & Milner, 1957; Shallice & Warrington, 1970). However, later neurophysiological findings showed that the brain regions involved in perception, working memory, and long-term memory of a certain informational content overlap (Fuster, 2003; Lewis-Peacock & Postle, 2008). Contemporary (*state-based*) models therefore posit that the distinction between short-term and long-term memory depends on the state of the memory (active or latent). There are several state-based models that differ in the details about the number of states possible, their capacity, and neural substrates, but most seem to agree that information in the focus of attention is consciously experienced, and consists of persistent neural activity (Cowan, 1995; Fuster, 1995; Jonides et al., 2008; McElree, 2001; Oberauer, 2002).

Cowan (1995, 2008) and Fuster (1995, 2003a) conceives an intermediate state of an active portion of long-term memory outside the focus of attention, which is non-conscious, and presumably consist of less neural activity. McElree (2001; 2006) thinks there is only latent non-conscious memory with varying synaptic weights outside a very narrow focus of attention. Oberauer (2001, 2002, 2005) posits a very narrow focus of attention, a state of direct access or broad focus of attention (like Cowan’s focus of attention), and active long-term memory. However, Oberauer’s model is neutral about conscious experience. Support for these postulated differences in states comes from different psychological experiments, and neuroimaging findings are still indecisive as there has been support for both McElree's (Oztekin, Davachi, & McElree, 2010; Oztekin, McElree, Staresina, & Davachi, 2009), and Oberauer's (Nee & Jonides, 2011, 2013) model.

Despite the historically intimate relationship with consciousness, current working memory models leave some theoretical room for non-conscious states even if that aspect seldom is in focus. Furthermore, empirical studies have shown that non-consciously presented information can engage endogenous attention in the form of cognitive control (Lau & Passingham, 2007; Reuss et al., 2015, 2011; van Gaal et al., 2010). Thus it has been shown that executive functions needed for working memory maintenance can operate on non-consciously presented information. It is thus plausible that working memory can operate on non-consciously presented information as well.

Hassin, Bargh, Engell, & McCulloch (2009) have made a case for implicit working memory, that is, non-conscious working memory operations on consciously presented information. They sequentially presented five disks (full or empty) in a grid matrix. The five disks could be in (i) a predetermined sequence/pattern, (ii) a "broken pattern" where the first four disks followed the predetermined pattern, but the last disk did not, and (iii) "control patterns" where the first three disks were randomly placed, and the last two disks followed the pattern. Participants did not know about the pattern, and was supposed to decide if the disks were full or empty. As predicted they found that reaction times were faster for predetermined patterns compared to broken patterns, and concluded that it was implicit working memory. However, it is difficult to know if this really was implicit working memory or implicit sequence learning based on basal ganglia (or both). The reaction times are also difficult to interpret since a closer look shows that reaction times for determined patterns was the same as for control patterns, but the broken pattern was slower than the other two patterns. Thus, there was no facilitated reaction time as predicted, only a slower reaction time from the broken patterns.

There have been two early attempts at exploring working memory maintenance of non-consciously presented information (Bergström, 2011, Bachelor's thesis, not peer-reviewed; Soto, Mäntylä, & Silvanto, 2011). I used an attentional blink paradigm similar to Study I (see method section) to render target stimuli (single letters) non-conscious followed by a variable delay period (150 milliseconds, 3 seconds, and 7 seconds), and found that participants performed better than chance after all delay durations, challenging the assumption that non-conscious retention is extremely short-lived and fickle.

Soto, Mäntylä, & Silvanto (2011) used a backward masking paradigm to render a target stimuli (Gabor gratings) with different orientations non-conscious. After a short delay (2 or 5 s with distractors) a test cue appeared. The test cue was also a Gabor grating, but with a different orientation than the target stimulus. The participants had to guess if the target stimuli orientation was clock-wise or counter clock-wise in respect to the test cue. They found above-chance performance despite no reported visual experience. These findings were interpreted as non-conscious working memory because the test cue's orientation never matched the target stimulus (excluding repetition priming), and since the distracters did not erase the non-conscious memory trace (excluding iconic memory). Their findings were compelling and consistent with (Bergström, 2011). Together these findings provided interesting results for continued investigations into the existence of non-conscious short-term retention.

Aim

The overall aim of the empirical studies of this dissertation was to investigate short-term retention of non-consciously presented information to better understand its functional properties (i.e., longevity and content), neural substrate (using fMRI), and relate the findings to current models of memory (e.g., working memory), and theories of consciousness. More specifically, the aims of the three studies were to:

Study I

- Find support for durable retention (up to 15 seconds) based on subjective criterion of experience
- Estimate the longevity of such short-term retention
- Investigate the neural substrates of the retention

Study II

- Replicate longevity findings with new method to render stimuli non-conscious based on objective criterion of experience
- Investigate the specificity of the retained content (identity & position)

Study III

- Further investigate the neural substrates of the retention in relation to:
- (i) prefrontal task-related cortical regions
- (ii) posterior stimulus-related cortical regions

Materials and methods

This section will give an overview of the varying techniques and methodologies used in the three studies. It will also provide some of the reasoning and underlying assumptions behind using them.

Participants

For all studies healthy young participants were recruited from the Umeå University campus area. All participants had normal, or corrected to normal, vision, gave written informed consent, and was paid for their participation. All studies that used fMRI were approved by the ethics committee of the University Hospital of Northern Sweden. For more details concerning inclusion/exclusion see the appended articles.

Study I contained two experiments. Experiment I analyzed 24 participants. Experiment II consisted of two sessions (pre-fMRI and fMRI session), and analyzed 27 participants. Study II contained two experiments. Experiment I analyzed 15 participants, and Experiment II 16 participants. Study III consisted of one experiment with two sessions. The pre-fMRI session analyzed 25 participants, and the fMRI session 26 participants.

Inferring the absence of conscious experience

There are several ways of measuring conscious experience, and infer non-conscious perception. These different measures can be categorized into two main approaches; subjective and objective approaches. Subjective approaches define conscious experience as reported conscious experience. If participants therefore report "no conscious experience" of the target stimuli, any subsequent behavioral or neuroimaging effects are attributed to non-conscious processes. The main advantage with using a subjective approach is that you are more directly measuring the phenomenon of interest, since conscious experience is subjective by nature. Another advantage is the possibility to index different degrees of conscious experience and content. However, subjective measures risk not being completely *exhaustive*, that is, they might not detect all relevant conscious experience, and can therefore overestimate non-conscious effects (Merikle & Joordens, 1997; Reingold & Merikle, 1988; Snodgrass & Shevrin, 2006).

The objective approach relies on objective task performance instead of subjective reports, and define conscious experience as above-chance performance on a direct (i.e., instructed) task (e.g., discrimination accuracy). The absence of conscious experience is therefore inferred when the direct task performance is at chance, and only then can indirect measures (e.g., reaction time or BOLD signal change) be attributed to non-conscious processes. The main advantage with objective measures is that any detectable non-conscious effect is very unlikely to be contaminated with weak conscious experience. However, objective measures might not *exclusively* measure conscious effects, and therefore risk underestimating non-conscious effects (Lau, 2008; Reingold & Merikle, 1988).

Historically, subjective measures were first used to investigate non-conscious processing (Pierce & Jastrow, 1884; Sidis, 1898; Stroh et al., 1908), but later suffered severe criticism (Eriksen, 1960), which gave rise to the use of objective measures. However, after methodological advances (and the shift from behaviorism to cognitive science) it is now generally accepted that both are valid approaches to measure the presence or absence of conscious experience. Indeed, it has been argued that an objective approach can be seen as a more conservative version of the subjective approach since they both produce similar and reliable results (Dehaene & Changeux, 2011; Merikle, Smilek, & Eastwood, 2001). There is currently no consensus on which measure is the most exhaustive and exclusive (Boly et al., 2013), and they can be seen as complementary.

Sandberg, Timmermans, Overgaard, & Cleeremans (2010) compared two subjective and one objective measure of conscious experience. Firstly, the *perceptual awareness scale* (PAS), in which participants are instructed to report their perceptual experience of the target stimuli ranging from no experience to clear experience (Ramsøy & Overgaard, 2004). Secondly, *confidence ratings*, where participants are instructed to report how confident they were of seeing or guessing correct target stimulus ranging from no confidence to high confidence (Cheesman & Merikle, 1986). Thirdly, an objective measure called *post-decision wagering* where participants are instructed to bet small monetary sums on their task accuracy ranging from a smaller to a larger monetary sum (Persaud, McLeod, & Cowey, 2007). They found that the PAS was more exhaustive, and thereby more conservative, than confidence ratings and post-decision wagering (Sandberg et al., 2010). In addition, Sandberg, Del Pin, Bibby, & Overgaard (2014) found that the PAS seems to be more exclusive than exclusion tasks (Debner & Jacoby, 1994). Another study found that confidence ratings were more exhaustive than post-decision wagering because the latter was affected by loss aversion (Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008).

Going forward we chose to use the PAS to measure conscious experience. We did so for three reasons. Firstly, since we intended to investigate non-conscious short-term memory, a fragile phenomenon that most assumed did not exist, we thought it prudent to begin with a liberal approach. If the liberal approach would yield memory effects we could try to replicate it with more conservative measures. Secondly, we chose subjective measures because they best catch conscious experience. Thirdly, we chose the PAS because it seemed the most conservative and reliable of the subjective measures. After settling on how to measure conscious experience, we had to decide what technique we would use to make the target stimuli non-conscious.

Trial procedures and stimuli

There are various techniques available to render information non-conscious. These techniques all have in common that they, in different ways and to varying degrees, manipulate attention to make the target stimulus seemingly invisible. They usually manipulate either exogenous or endogenous attention. Techniques that manipulate endogenous attention tend to lead to more extensive non-conscious processing (Kouider & Dehaene, 2007). However, there is no technique that is superior to all other, they all have different strengths and weaknesses (for a review see Kim & Blake, 2005), which makes them more or less suitable for different purposes.

Study I

For the purpose of investigating the retention of non-consciously perceived information we wanted a technique that maximized non-conscious perceptual processing, yet was reliable enough to work on several consecutive trials even when the participants were aware of the manipulation. Here the underlying assumption was that stronger non-conscious perceptual processing would translate into longer and more robust retention. To that end, we decided to use an *attentional blink* paradigm for Study I (Figure 1). The attentional blink phenomenon is known to exhibit relatively strong non-conscious effects (Kouider & Dehaene, 2007; Luck, Vogel, & Shapiro, 1996; Marois, Yi, & Chun, 2004; Sergent, Baillet, & Dehaene, 2005).

To achieve an attentional blink two targets are presented within succession in a rapid serial visual stream of irrelevant stimuli (distractors). When endogenous attention is used to identify the first target (T1) there is subsequently an attentional blink (or gap) such that the second target (T2) is rendered non-conscious (Duncan, Ward, & Shapiro, 1994; Marois, Chun, & Gore, 2000; Raymond, Shapiro, & Arnell, 1992). However, T2 will only become non-conscious when there is a time-lag in the order of a few hundred milliseconds between T1 and T2 (short lag). If the time-lag is too short or too long there will be a reduced attentional blink. For baseline comparisons we also included trials without T2 (replaced by another distractor).

Previous literature suggested that more cognitively demanding T1 tasks generate stronger attentional blink effects (Martens & Wyble, 2010). We therefore made some modifications to the traditional attentional blink paradigm, in which T1 usually is a simple identification task (e.g., even or uneven number). In our attentional blink paradigm, T1 was a simple math task (e.g., 5+3 in red color) to be solved immediately, and remembered until the end of the trial. The T1 math task enabled us to present T2 (a single letter; A, S, D, or F; flanked by two irrelevant digits) for longer durations than traditional attentional blink paradigms without being consciously experienced. Here we assumed that longer presentation durations would lead to more durable retention of the non-consciously presented stimuli. However, since the attentional blink worked differently on different participants, the presentation durations were adjusted during the experiment based on how many T2 had been visually experienced. The algorithm decreased or increased the presentation duration so that the participant only would experience approximately 50% of the short-lag trials.

To estimate the longevity of non-conscious retention in Experiment I we had a variable (distractor-filled) delay of 5, 10, or 15 s between T2 presentation and T2 response (Figure 1A). From these data points our plan was to extrapolate a potential decline in performance over time, and thereby estimate when it would reach chance level (if it had not already done so within the three time points). In some trials we used long time lags between T1 and T2 to make T2 visible (figure 1B). These long-lag trials had different lag length such that T2 was presented early, intermediately, or late in the trial. The participants' could therefore not know whether they had missed an early T2 or if it was yet to come, and therefore had to attend the rapid serial visual presentation until its end.

After the delay, participants were shown a screen with all four letters, during which time they reported the letter that had been presented previously that trial. If participants did not consciously experience T2 they were instructed to guess by using the first alternative that came to mind or their gut-feeling. Each finger was mapped to a specific letter, making it possible to create stimulus-response bindings (Henson, Eckstein, Waszak, Frings, & Horner, 2014), which might further boost retention. Next, the participants estimated their perceptual experience of T2 using a four-point perceptual awareness scale (PAS; 1. no experience, 2. vague experience, 3. almost clear experience, 4. clear experience; Sandberg et al., 2010). Lastly, the participants' had to report the solution to T1.

Experiment II: Pre-fMRI session had some minor changes to the protocol. The PAS was reduced to a three-point scale (1. no experience, 2. vague experience, 3. clear or almost clear experience). The new three-point scale was judged to be easier to use without sacrificing the critical distinction between "no experience" and "vague experience". We also increased the T1 math task difficulty slightly to facilitate stronger attentional blink effects. The pre-fMRI session was primarily used to screen for unsuitable participants (e.g., non-blinkers that are immune to the attentional blink; Martens & Wyble, 2010), and calibrate the individual presentation durations, prior to the fMRI session (though presentation durations were also adjusted continuously during the fMRI session). The fMRI session had some additional changes (figure 1C). Inter trial intervals (ITIs) and inter stimulus intervals (ISIs) were jittered to reduce correlations between components in the statistical model. The delay period only had distractors for a short time after T2 to maintain the attentional blink phenomenon, but otherwise consisted of a dot to be passively viewed. Removing the distractors was thought to make the stimulus-representation less noisy in terms of BOLD signal. However, in hindsight, it might have been preferable to leave them in for methodological reasons, and to strongly activate the prefrontal cortex in its role to protect the memoranda against distraction (Feredoes et al., 2011).

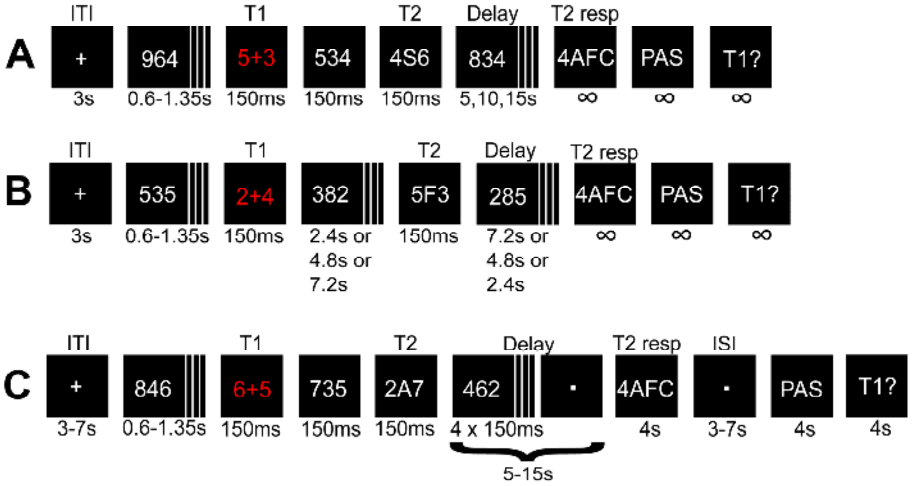

Figure 1. Illustration depicting the procedure for experiment 1 and 2. A math task (T1) and a letter (T2) flanked by distracters were presented in a rapid serial visual presentation (RSVP). The solution to T1, and T2 identity, were held in memory during a variable delay until probed for responses. **(A)** Short- and **(B)** long-lag trials in experiment 1 and the pre-fMRI session. In **(B)** the variable delays before and after T2 presentation were adjusted such that T2 appeared early, in the middle, or at the end of the RSVP. **(C)** Illustration of a short-lag trial in the fMRI session.

Study II

For Study II we chose to use a different technique to render stimuli non-conscious, namely, continuous flash suppression (CFS; Tsuchiya & Koch, 2005). We changed from the attentional blink to CFS for several reasons. The main strength with CFS is its strong suppression. The suppression leads to fewer ambiguities as to whether something was experienced or not, and stimuli can be presented for longer durations (more than a minute; Tsuchiya & Koch, 2005). Its main weakness is that the suppression might be too strong, and therefore limit non-conscious effects. We reasoned that we might be able to offset some of the lost effect by using relatively long presentation durations (seconds instead of milliseconds), and that we could use an objective instead of subjective criterion of experience.

CFS is based on the phenomenon bistable perception. Bistable perception is when the conscious experience switches between two mutually exclusive perceptual states, as is the case when perceiving ambiguous figures or during binocular rivalry when two completely different static images are simultaneously presented to each eye.

During CFS the visual input from the left side of the monitor was presented in isolation to the left eye, and vice versa for the right eye, by using a mirror stereoscope. A non-salient target stimulus is presented to the non-dominant eye, while colorful moving geometrical figures named mondrians (after the Dutch painter Piet Mondrian) are presented to the dominant eye. Because the mondrians are much more salient and dynamic, they will consequently capture exogenous attention, and thereby suppress the target stimulus from consciousness. In addition to the presentation condition with suppressed (non-conscious) target stimuli, there were also baseline trials without target stimuli where mondrians suppressed a blank grey background, and conscious trials where the target stimulus was presented to both eyes (without mondrians). We used CFS in combination with a common short-term memory task called delayed match-to-sample (DMS) task (figure 2). In the DMS task participants are shown a target stimulus to be remembered for a short delay-period until they are shown a probe. The probe can either match the memorandum or not, and the participants were instructed to make a match/non-match discrimination. A detection task followed where participants reported if there was any stimulus at all or not, and, similarly to Study I, a three-point PAS.

In Experiment I the participants were shown emotional faces in different spatial quadrants to be remembered for 5 or 15 s delays (figure 2A). The face sample either matched in face identity, emotion, and spatial position, or none of them. However, the participants were only instructed to remember the spatial position. We reasoned that spatial position had the highest likelihood of being non-consciously retained, but matching emotion and identity might facilitate retention, making it easier to detect any non-conscious effects.

Given the success of Experiment I (which could not separate the distinct information components), we intended to further probe the specificity of the information retained non-consciously in Experiment II. In Experiment II the participants were therefore shown tools in different spatial quadrants to be remembered for a 5 s delay (figure 2B). We decided to use tools instead of faces because their distinct shapes/features should facilitate better within-category discrimination, and previous studies indicate that tools might be processed to a larger extent during CFS (Almeida, Mahon, Nakayama, & Caramazza, 2008; Fang & He, 2005). The participants were instructed to remember the identity and position. The probe either matched identity and position, only identity, only position, or neither. By having a probe that could match only one information component, we could later determine whether one of the components was driving the memory performance. Lastly, we decided to take away the detection task to make the trials shorter since there was no differential effect between the discrimination and detection task in Experiment I.

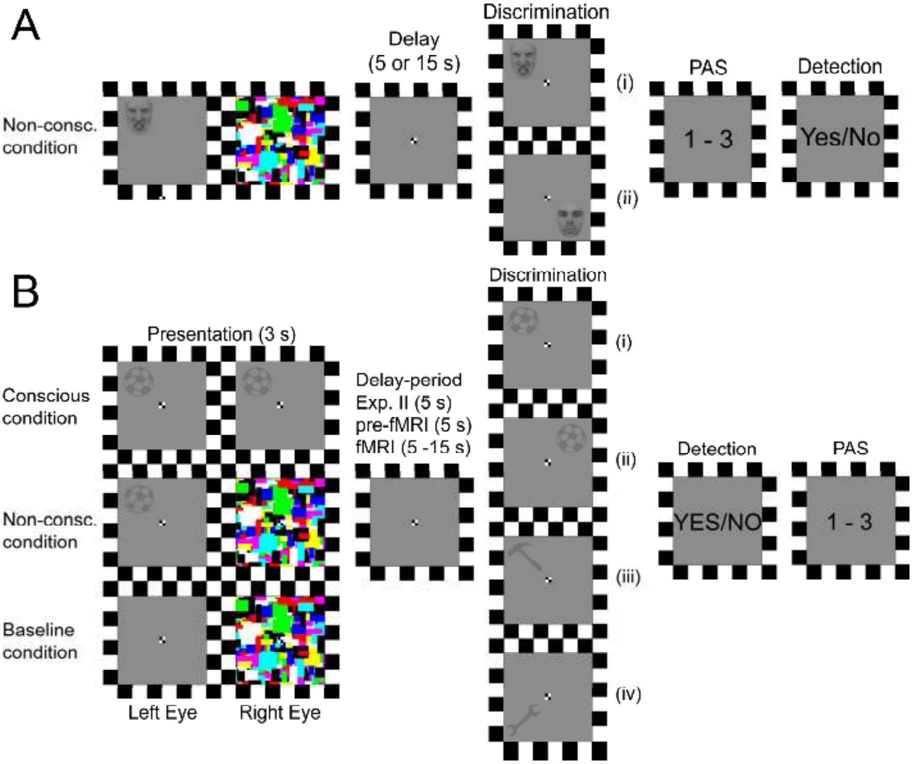

Figure 2. Illustration depicting the procedures of **(A)** Experiment I in Study II, and **(B)** Experiment II in Study II, and Study III. There was no detection task in Experiment II in Study II. PAS = perceptual awareness scale.

Study III

For Study III we used a very similar experimental protocol to Experiment II in Study II (figure 2B). Here we wanted to investigate the neural correlates underlying the non-conscious retention in Study II. The only differences were that we added the detection task, and the fMRI session had a delay duration ranging from 5 – 15 s and jittered ITI (figure 2).

Functional magnetic resonance imaging (fMRI)

Magnetic resonance imaging (MRI) and functional magnetic resonance imaging (fMRI) are relatively new neuroimaging techniques. After the basis for all modern MRI, nuclear magnetic resonance, was discovered in the 1940s (Bloch, 1946; Purcell, Torrey, & Pound, 1946) many non-biological MRI studies followed. It was not until the 1970s that the first biological MR image was produced (Lauterbur, 1973), and after considerable improvement in MR image acquisition speed (Mansfield, 1977) that MRI became popular to image the anatomical structure of the brain. The birth of fMRI occurred in the 1990s when it was discovered that changes in blood oxygenation could be measured with MRI (Ogawa, Lee, Kay, & Tank, 1990; Ogawa et al., 1992). Today fMRI is one of the most popular and widely used techniques to investigate brain function. Its popularity stems mainly from being a non-invasive neuroimaging technique with high spatial resolution.

MRI makes use of very strong magnetic fields to create images of biological tissue. It is now common for MR scanners to have static magnetic fields of 3 Tesla or higher. For comparison, the magnets used to lift cars at scrap yards have approximately 1 Tesla. The static magnetic field aligns the atomic nuclei in the body to the magnetic field in a state of equilibrium (usually hydrogen atoms in water and fat because of its prevalence in the body). Radiofrequency coils then transmits electromagnetic waves into the body, which excites the atomic nuclei. The excited atomic nuclei absorb the electromagnetic energy from the radiofrequency coils causing some of them to change from a low-energy state to a high-energy state, and thereby change the equilibrium. When the radiofrequency coil stops transmitting electromagnetic waves, the atomic nuclei reverts to equilibrium (relaxation). As the atomic nuclei revert to low-energy states again they release the absorbed energy, which is received by the radiofrequency coils. The received electromagnetic energy is the raw MR signal. However, the raw MR signal has no spatial information. Gradient coils therefore superimpose magnetic gradients on the static magnetic field to modify its strength along the x, y, and z axis, making it possible to create three-dimensional images (for a primer on fMRI see Huettel, Song, & McCarthy, 2004).

By changing the parameters of excitation and relaxation, it is possible to relate the MR signal to different physiological properties, and thereby create different images for different purposes. The so called T1-weighted images are used to create high-resolution images of the brains anatomical structure. Because different brain tissues (grey matter, white matter, and cerebral spinal fluid) have different densities of hydrogen atoms, the corresponding MR signal becomes different. The difference in MR signal is then used to distinguish between different tissues when creating the structural images.

The so called T2*-weighted images are used to create low resolution images of the brains blood flow as an indirect measure of neural activity. More precisely, the T2*-weighted imaging measures the concentration of oxygenated (to deoxygenated) hemoglobin over time (which changes with blood flow), and is therefore referred to as the blood-oxygenation-level dependent (BOLD) signal (Kwong et al., 1992; Ogawa et al., 1990; Ogawa et al., 1992). The paramagnetic property of deoxygenated hemoglobin disrupts the MR signal. Thus, a low concentration of oxygen leads to a weak BOLD signal, while a high concentration of oxygen leads to a high BOLD signal. When a task is performed, the corresponding neurons will fire, the firing neurons require energy in the form of glucose and oxygen, which is delivered through the blood stream, and gives rise to the BOLD signal.

However, the blood flow is relatively slow compared to neuronal firing. After an initial dip of 1 - 2 s (Buxton, 2001) the BOLD signal peaks 4 - 6 s after a brief stimulus presentation, and then returns to baseline after 12 -20 s ending with a late undershoot (Boynton, Engel, Glover, & Heeger, 1996; Poldrack, Mumford, & Nichols, 2011). The BOLD signal's maximum observed amplitude is approximately 5% in primary sensory regions, and 0.1 - 0.5% for other regions (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001).

Although the BOLD signal is an indirect measure of neuronal activity, and their exact relationship remains unclear, there is nonetheless a reliable correlation between the BOLD signal and neuronal activity, which mostly relates to the neurons synaptic (input) rather than spiking (output) activity (Goense & Logothetis, 2008; Logothetis et al., 2001; Logothetis & Wandell, 2004). There are, however, some caveats regarding the BOLD signal (Logothetis, 2008). Firstly, it is rather coarse, as such a typical fMRI voxel (smallest spatial unit) of 55 mm^3 contains about 5.5 million neurons. Secondly, the BOLD signal cannot distinguish between top-down and bottom-up signals. Thirdly, since excitatory and inhibitory neurons both cause regional blood flow, the BOLD signal likely conflates the two.

Univariate analysis and multi-voxel pattern analysis (MVPA)

After the fMRI data is collected it needs to go through a fair amount of preprocessing (slice timing correction, motion correction, normalization, and smoothing) prior to statistical analyses (for more on preprocessing see Huettel et al., 2004; Poldrack et al., 2011). The traditional way to analyze fMRI data is referred to as univariate (or "uni-voxel") analysis, and essentially means that each voxel is analyzed independently from the other voxels. Instead, multivariate (or "multi-voxel") analyses use the combined information from more than one voxel.

In short, the univariate analysis makes use of the general linear model (GLM) as a framework to do a multiple regression. The modelled BOLD signal time-series for each cognitive task are used to predict the collected fMRI data (the actual BOLD signal time series for all voxels) one voxel at the time. How good each modelled response fits the actual BOLD signal time-series in a voxel will then be calculated as parameter estimates (or beta values). For example, if a participant is randomly presented with visual or auditory stimuli over time, voxels in the visual cortex will have high beta values when related to the modelled time-series for the visual stimuli, and low beta values when related to the modelled time-series for the auditory stimuli, and vice versa for voxels in the auditory cortex. The beta values are converted to useful statistics, for example, T values or Z values. These statistical values can then be used to infer whether the beta values of voxels are significantly different from zero, or beta values of another task (e.g., by subtracting the statistical values related to a visual task from an auditory task or a passive baseline task). Thus, making it possible to infer which voxels are associated with a certain task relative to another task (Lindquist, 2008; Smith, 2004).

There are many different kinds of multi-voxel pattern analyses (see Haxby, Connolly, & Guntupalli, 2014), but I will only focus on the particular one used in this dissertation (Study III), which uses machine learning algorithms that are trained to detect differences in fMRI patterns (amongst many voxels) between experimental conditions. Although a univariate analysis will be able to distinguish between two different visual stimuli presentations (e.g., faces vs. houses), it might not be able to distinguish between two different visual stimuli within the same category (e.g., two different faces). However, a multi-voxel pattern analysis might be able to make such a distinction when drawing on the BOLD signal pattern of all available voxels.

Firstly, one needs to select which voxels that should be used in the MVPA, which is called “feature selection”. The feature selection is usually a subset of all voxels in the brain based on a priori hypothesis or functional localizer. Selecting many irrelevant voxels will introduce unnecessary noise, and make it harder to detect pattern differences. For example, if the aim is to decode which face the participant is looking at, one might select voxels that are associated with face perception in general. Secondly, the data is organized such that the BOLD signal of the selected voxels corresponds to the right point in time when a certain face was presented. Thirdly, the chosen machine learning algorithm is then fed a portion of the data to train and learn to predict which face the participant is looking at based on the BOLD signal pattern of the voxels. Fourthly, once the algorithm is trained, it is fed new data for generalized testing, to determine if the trained algorithm can predict the correct face significantly better than chance. The classification performance is usually tested for significance with non-parametric permutation testing (for more on MVPA see Haynes & Rees, 2006; Haynes, 2015; Norman, Polyn, Detre, & Haxby, 2006; Tong & Pratte, 2012).

Results

Study I

The historically close relationship between working memory, long-term memory, and conscious experience has inspired theories of consciousness to posit that durable maintenance of information is a uniquely conscious function (Dehaene et al., 2014; Dehaene & Naccache, 2001). It was generally assumed that the retention of non-consciously presented information was caused by priming, and disappeared within 500 milliseconds (Dehaene & Changeux, 2011; Greenwald et al., 1996). However, we had reason to believe that non-consciously presented information could be retained for a few seconds (Bergström, 2011; Soto et al., 2011).

Overall, the experiments in Study I showed that non-consciously presented information could be retained up to 15 seconds without signs of decline. The long delay periods enabled us to isolate the sluggish BOLD signal specifically related to the delay period rather than stimuli presentations and discriminations. This non-conscious retention was associated with sustained BOLD signal change in the prefrontal cortex and cerebellum (Figure 3B). Contrary to expectations from (conscious) working memory maintenance, the sustained BOLD signal in the prefrontal cortex did not correlate with memory performance (Pessoa, Gutierrez, Bandettini, & Ungerleider, 2002; Wager, Spicer, Insler, & Smith, 2013).

The non-consciously presented letters were associated with BOLD signal change in the visual area thought to process letters (Heinzel et al., 2008; Park, Hebrank, Polk, & Park, 2012), and prefrontal and parietal regions previously related to letter perception (James & Gauthier, 2006). These areas largely overlapped with the more pervasive BOLD signal related to conscious letter perception (figure 3A). Surprisingly, there was no BOLD signal change related to the memory test of non-conscious information (figure 3C). Reaction times for memory tests of conscious information were quicker than non-conscious (and baseline) reaction times.

We lowered the statistical threshold ($p \leq .05$) to see if there was any subthreshold hints of hippocampi involvement during presentation and memory test, but there were none[1]. This makes it less likely that hippocampi-based mechanism played a large role in the non-conscious retention.

In sum, we found that non-consciously presented information could be retained for much longer than previously assumed despite irrelevant distractors. This non-conscious retention was associated with sustained BOLD signal change in prefrontal and cerebellar regions. These initial findings are consistent with working memory, and can have implications on working memory models and theories of consciousness.

---

[1] The hippocampi control analysis was not reported in the published article.

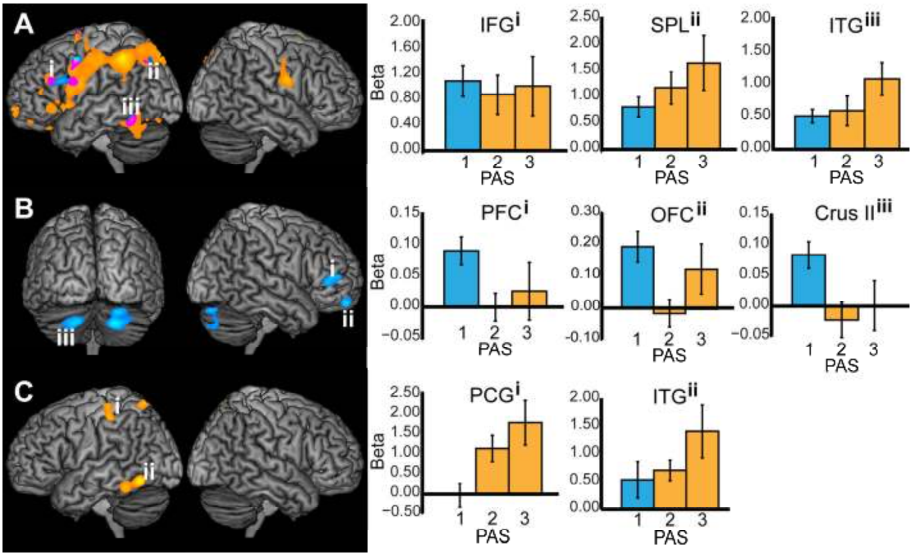

Figure 3. All working memory epochs with T2-seen > T2-absent trials in orange, T2-unseen > T2-absent trials in blue, and the overlap in purple (p ≤ .001, uncorrected for multiple comparisons, k = 20). The Y-axis: Beta values, X-axis: PAS = perceptual awareness scale, bar colors correspond to the PAS classification of T2-seen and T2-unseen trials, error bars: standard error of the mean. (A) Stimulus presentation (IFG = inferior frontal gyrus, SPL = superior parietal lobule, ITG = inferior temporal gyrus). (B) Delay period displayed at p ≤ .005 for illustrative purposes (PFC = prefrontal cortex, OFC = orbitofrontal cortex, Crus II in the cerebellum. (C) Response (PCG = postcentral gyrus, ITG = inferior temporal gyrus).

Study II

Given the results from Study I with a subjective criterion of conscious experience, we wanted to replicate those findings with a new method and material. We therefore changed technique from attentional blink to CFS. The CFS suppression leads to fewer ambiguities as to whether something was experienced or not, and stimuli can be presented for longer durations. CFS might also make it possible to use an objective criterion of experience. The purpose of Study II was to replicate the longevity findings from Study I with CFS, and to investigate what informational content that could be non-consciously retained.

Experiment I of Study II showed that spatial position (of emotional faces) could be retained up to 15 seconds, replicating the longevity of Study I with new material and CFS. In Experiment II we wanted to investigate the specificity of the retained information. That is, if tool identity and spatial position could be simultaneously retained rather than only one of them. The memory-test probe was therefore not limited to full match or non-match (as in Experiment I), but could also partially match the target stimuli in regard to identity or position. We could therefore analyze the hit rate, and false alarm rates for partial matches, and non-matches separately (figure 4). It was found that the hit rate was larger than the other false alarm rates, and no difference between the false alarm rates. Had for example only spatial position been retained, the hit and spatial false alarm rates would have been equal, and larger than the others. These results show that it indeed was the conjunction of identity and position that was retained, and used during discrimination. The non-conscious effects found in Study II were on direct measures when reports indicated no experience, and therefore also based on subjective criteria. The reaction times did not show signs of non-conscious priming, and similarly to Study I the conscious reaction times were quicker than non-conscious (and baseline) reaction times.

In sum, we replicated Study I in terms of non-conscious retention longevity (15 seconds) with CFS and new material. Furthermore, we showed that the conjunction of stimuli identity and positon can be non-consciously retained, and that reaction times showed no priming effects. These results support the notion that non-consciously presented information is more durable than previously assumed, and indicated that the experimental procedure was stable enough to use in the fMRI scanner.

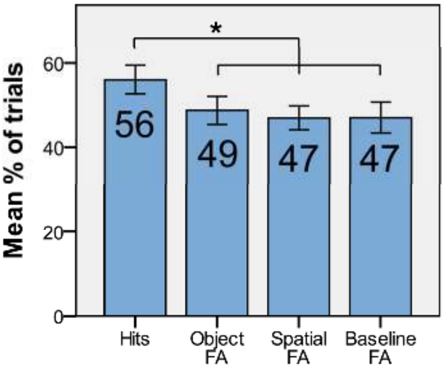

Figure 4. The average proportion of non-conscious hits (sample matched identity and position), object-FA (sample matched object identity), spatial-FA (sample matched spatial position), and baseline-FA (sample did not match) with standard error for non-conscious trials. *p < 0.05, one-tailed.

Study III

A hallmark of working memory maintenance is persistent neural activity in prefrontal task-related, and posterior stimulus-related regions. In Study I we found sustained BOLD in prefrontal regions (and cerebellum), but not in stimulus-specific sensory regions. In Study III we therefore used the same paradigm as Study II with fMRI to see if we could extend our previous findings with sustained BOLD signal in regions related to stimuli identity and position.

In the pre-fMRI session the non-conscious memory performance (discrimination and detection tasks) was above chance after 5 second delay periods. Surprisingly, discrimination and detection performance dropped to chance for the same participants after a delay of 5 – 15 seconds in the fMRI session. Additionally, the univariate fMRI analysis did not detect any BOLD signal change during the delay period. However, multivoxel pattern classification analysis (MVPA) can be more sensitive than behavioral measures (Haynes & Rees, 2005) and univariate analysis (Emrich et al., 2013; Riggall & Postle, 2012). We therefore used MVPA to further investigate the trial epochs (presentation, delay, and discrimination). We used the conscious trial epochs to select relevant subsets of voxels to include in the MVPA for the same epoch. That is, a conscious delay-based subset to decode non-conscious delay-related information, etc. Although our main interest was in the delay period, we began with stimuli presentations for comparison. It should be easier to decode something during presentations because of the constant visual input.

The classification algorithms successfully decoded non-conscious stimuli presence vs. absence using a subset of voxels from the whole brain. Spatial information (left vs. right visual field) was decoded using a subset of voxels in the parietal and occipital cortex. Next we tried to do the same for the delay period. The algorithms failed to decode presence vs. absence from a subset of whole-brain voxels. However, they successfully decoded presence vs. absence when using a subset of voxels in the frontal cortex, and spatial position (left vs right) from a subset of voxels in the occipital cortex. To control that the successful decoding of BOLD signal during the delay period was not caused by residual BOLD signal from the presentation, we used the same voxels that successfully decoded BOLD signal during the delay period, and tried to decode BOLD signal during the presentation epoch. Critically, the algorithms failed to decode presence vs. absence, and spatial position when using the same voxels as the delay period during the presentation epoch. This suggests that the successfully decoded BOLD signal during delay is related to memory, and not perception.

The univariate analysis detected BOLD signal change in the right anterior insula, and right inferior frontal cortex during the non-conscious discrimination tests. In addition, there was more pronounced BOLD signal change when the memory test probe was not identical (non-match) to the memoranda compared to when the probe and memoranda was identical (match). Non-matching probes also reveal higher BOLD signal change relative to baseline trials. These findings were similar for both conscious and non-conscious trials, and partially overlapped (figure 5).

Control analyses showed that non-conscious univariate and MVPA results cannot be explained by participants' accidentally mislabeled non-conscious trials where they "experienced something" (PAS > 1) as "no experience" (PAS = 1). We determined that by dividing baseline trials in two groups, and contaminated one group with conscious trials. The amount of contaminated trials was based on the proportion of conscious trials that were accidentally reported as "no experience" (1 trial per participant). Critically, neither univariate analysis nor MVPA could detect the contamination. In a similar control analysis as Study I ($p \leq .05$), we found BOLD signal change in the right hippocampus during stimuli presentation, but not discrimination[2].

In sum, we found sustained stimulus-unspecific BOLD signal in frontal cortex, and stimulus-specific spatial information in the occipital cortex with an objective criterion of experience (memory performance at chance). These findings are in line with Study I & II, and show strong support for working memory maintenance of non-consciously presented information.

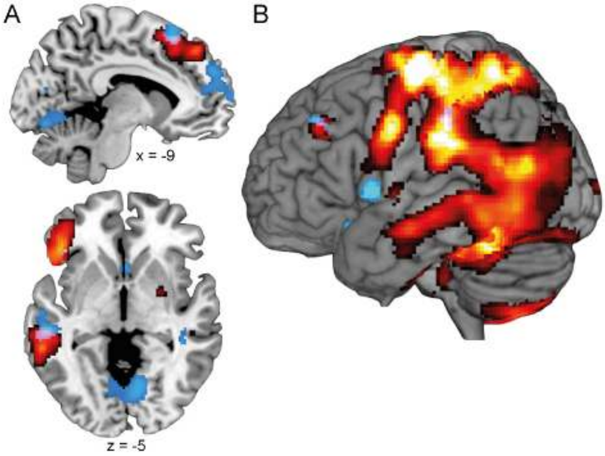

Figure 5. $p \leq .001$, uncorrected for multiple comparisons, $k = 20$. **A**. BOLD signal change during conscious (hot colors) and non-conscious (cold colors) trials for memory probe recognition when the probe did not match the sample compared with sample-probe match (non-match > match). Overlap (purple) is evident in medial frontal (upper) and middle temporal (lower) regions. **B**. BOLD signal change during conscious (hot) and non-conscious (cold) trials for non-matching probes compared with baseline trials. Overlap (purple) is evident in middle frontal and supramarginal gyrus.

---

[2] The hippocampi control analysis was not reported in the manuscript.

Discussion

The purpose of this dissertation was to investigate the functional properties and neural substrates of non-conscious short-term retention, in order to determine what kind of memory it is, and possible implications to models of memory and theories of consciousness. Studies I – III show that non-consciously presented information can be retained for up to 15 seconds, sometimes with distractors, across different techniques (attentional blink and CFS) and material (letters, spatial position, and the conjunction of tools and position) to be remembered, and was based on neural and/or behavioral evidence using subjective and objective criterions of experience. In addition to sustained BOLD signal in the frontal cortex during the delay period, Study I found it in the cerebellum, and Study III in the occipital cortex (related to spatial position). I will here discuss these findings in relation to other literature, and possible theoretical and practical implications.

What kind of non-conscious memory is it?

Overall, the three studies corroborate and extend previous findings (Bergström, 2011; Soto et al., 2011), suggesting that non-consciously presented object identity and/or spatial position can be retained for several seconds, sometimes with distractors. However, this short-term retention could in principle be explained by many different kinds of memory, which here will be considered in turn.

Working memory

Conscious working memory maintenance is known to be relatively robust against distractions, and can be maintained for as long as the information is attended. Delay periods of several seconds are therefore not a problem as demonstrated by the high working memory performance in the conscious conditions of the studies in this dissertation. In addition to converging evidence from Study I & III showing that short-term retention of non-consciously presented information was associated with sustained BOLD signal in the frontal cortex, Study I found it in the cerebellum, and Study III found BOLD signal related to spatial positon in the occipital cortex. It is possible that the absence of significant BOLD signal in visual regions in Study I can be explained by the relatively insensitive univariate analysis, and that a MVPA would have detected presence vs. absence or stimulus-specific features similar to Study III. Together, these findings are consistent with predictions based on (conscious) working memory, which commonly find task-related BOLD signal in the prefrontal cortex (Fuster, 2015; Pochon et al., 2001; Rowe et al., 2007; Sakai & Passingham, 2003), and stimulus-related BOLD signal in posterior regions during maintenance (Druzgal & D'Esposito, 2001; Emrich et al., 2013; Fuster, 2009; Harrison & Tong, 2009; Postle et al., 2000; Pratte & Tong, 2014). Furthermore, the cerebellum is anatomically connected to the prefrontal cortex (Bostan, Dum, & Strick, 2013; Kelly & Strick, 2003), and have been related to verbal working memory (Desmond et al., 2005; Desmond & Fiez, 1998; Stoodley & Schmahmann, 2009). The findings of Study III are especially compelling because in addition to stimuli-unspecific information in the frontal cortex, there was also stimulus-specific spatial information in the occipital cortex - a hallmark of working memory - with an objective criterion of experience. Although it may seem unbelievable to decode non-conscious information from BOLD signal despite behavioral performance being at chance, it is in line with previous research. Haynes & Rees (2005) successfully decoded the orientation of fine lines in the primary visual cortex during non-conscious perception when discrimination was at chance.

Furthermore, there is indirect behavioral support for non-conscious working memory maintenance when direct performance was at chance (Pan, Lin, Zhao, & Soto, 2013). Pan, Lin, Zhao, & Soto (2013) used backward masking to render faces non-conscious, and CFS during the delay period to suppress either an identical or different face, while timing how long it took for the suppressed face to break through suppression (and become conscious). Participants then had to guess if the face that broke through suppression was the same or different as the non-consciously presented face. Importantly, they found that faces that matched the non-consciously presented faces broke through suppression quicker compared to mismatching faces, but memory performance was not better than chance. Critically, a follow-up experiment showed that if there was no memory test at the end of the trial, there was no difference in break-through time between matching and mismatching trials. This study provides compelling evidence for non-conscious working memory for two reasons. First, they show that the non-conscious retention depends on prospective action, which is a crucial part of the definition for working memory. Secondly, they used an objective criterion for non-conscious memory (i.e., memory performance was at chance).

Showing that non-conscious short-term retention can depend on prospective use is an important contribution and complement, which lends weight to a working memory interpretation of Studies I - III. Studies I & III carry important contributions because unlike other studies, the BOLD signal can be related specifically to the delay period with relatively high spatial and content specificity. That is, they specifically focus on the actual retention processes related to the non-conscious information, which is crucial to determine if the retention state is active or latent.

Similarly, Dutta, Shah, Silvanto, & Soto (2014) used a similar masking paradigm as Soto et al. (2011), but with a 1.5 s delay and no distractors, to investigate its neural substrate with fMRI and transcranial direct current stimulation (tDCS). They found that BOLD signal change in prefrontal, parietal, and temporal cortex correlated with memory performance, and that tDCS on the prefrontal cortex modulated performance. Although their findings are in line with Study I & III there are some caveats. Firstly, the sluggishness of the BOLD signal together with their short 1.5 s delay makes it impossible to separate the BOLD signal from presentation, delay, and discrimination. Secondly, the tDCS was applied for 15 min before the experiment, which similarly means that the modulated prefrontal involvement could be related to processes during any of the three trial epochs. Since it is impossible to know if their findings are related to encoding or retrieval rather than retention, it is impossible to conclude anything about how the information was retained, only that it was, and that the prefrontal cortex played a causal role.

King, Pescetelli, & Dehaene (2016; submitted, but not yet peer-reviewed) also used a similar masking paradigm as Soto et al. (2011), but with 800 milliseconds delay and no distractors. They used magnetoencephalography (MEG) to study neural activity during non-conscious retention. MEG has a much higher temporal (but lower spatial) resolution than fMRI, which permitted them to separate neural activity from presentation, delay, and discrimination despite short delay periods. Using MVPA (on whole-brain data) it was found that stimuli presence vs. absence could be decoded from presentation, delay, and discrimination, but stimulus-specific information (i.e., target angle) could only be decoded from presentation and discrimination, not the delay period.

The short delay period of 800 milliseconds opens up the possibility that neural activity represents residual activity from perception (iconic memory) rather than maintained activity as in working memory. The authors argue that the presence of non-conscious information during the delay indicates task-relevant maintenance. However, the target angle could not be decoded, and stimuli presence potentially contains target angle (task-relevant feature), spatial frequency and contrast level (task-irrelevant features). Since the task-irrelevant features could not be decoded from the delay period of consciously experienced stimuli, it could be argued that they are unlikely to persevere during non-conscious retention. However, after experiencing the stimuli participants likely focus attention on the relevant feature, and might thereby inhibit the irrelevant features more compared to when the stimuli was non-conscious. Consistent with Study I & III, the MEG pattern suggests that the non-consciously retained information is supported by global activity, not only by local activity in visual cortex, which is consistent with a working memory interpretation.

Studies I & III and King et al. (2016) seems to imply that non-consciously presented information can be maintained as sustained BOLD/neural signal. However, given the relatively weak signals and that BOLD and MEG signals are more sensitive to synaptic processing rather than neural spikes, it is unclear how persistent the underling neural spiking rate was, and the signal could reflect semi-persistent bursts of activity and/or metabolically demanding synaptic events rather than constant neural spiking, but nevertheless reflect short-term retention (Lundqvist et al., 2015; Mongillo et al., 2008; Shafi et al., 2007). This would mean that at its most extreme, the stimuli representation might be retained by sub-threshold membrane potentials without any neural spiking. However, for such sustained changes in membrane potentials there must be sustained input coming from somewhere else, which has to have neural spikes. In the case of Studies I & III, it would likely come in the form of sustained endogenous attention from the frontal cortex. For simplicity I will use persistent as in "persistent and/or on average persistent" to refer to all of these possible states ranging from non-spiking neural activity to constant spiking since my measures are on average BOLD signal, and because these active states will be considered together in contrast to latent states.

Studies supporting latent short-term retention have failed to find evidence of sustained BOLD/neural signal during successful (consciously presented) short-term retention of deprioritized information (LaRocque et al., 2013; Lewis-Peacock et al., 2012; Nelissen, Stokes, Nobre, & Rushworth, 2013). It has therefore been suggested that latent neural changes underlies deprioritized short-term retention. It remains difficult to know whether these deprioritized stimuli received unmeasurably small sporadic bursts of activity, very small amounts of continuous activity, or if they truly were completely inhibited (i.e., no neural activity at all). However, if the latter is assumed true, then it puts additional constraint on working memory models. Based on these findings it has been argued that latent neural changes underlies the working memory state called "activated long-term memory" (Larocque, Lewis-Peacock, & Postle, 2014).

It might seem surprising or contradictory that non-consciously presented information that never has been inside the focus of attention nevertheless can be maintained as persistent neural activity, while information previously conscious in the focus is retained as latent neural changes. How can these two different findings be reconciled within one working memory framework? A possible, but speculative, solution might be found in the definition of working memory – information retained *for* prospective use. I hypothesize that the persistent neural activity (in any form) is dependent on prospective action, while the latent neural change does not. A consequence of this line of reasoning is that any form of latent neural retention is not strictly working memory, but short-term memory, because it will decay in a time-interval dependent on other factors than prospective use.

It is currently hard to evaluate this hypothesis since the deprioritized stimuli in (LaRocque et al., 2013; Lewis-Peacock et al., 2012) was not currently necessary for a prospective action, but could become cued as necessary later in the trial with a 50% probability. Similarly, Nelissen, Stokes, Nobre, & Rushworth (2013) also failed to decode the deprioritized stimuli, but in their case the stimuli never became relevant. In contradiction, Peters, Roelfsema, & Goebel (2012) were able to successfully decode unattended stimuli in a visual search task, but interestingly, they were always cued as relevant later in the trials. However, the interpretation gets slightly more complicated as the deprioritized stimuli appeared in the search stream on half of the trials, which could have biased decoding performance. Behaviorally, Oberauer's state of activated long-term memory was initially based on the intrusion effect from information that no longer was task relevant, and never became relevant (Oberauer, 2001). In later studies the information was passive, and therefore not task relevant per se, but was still tested later on an unknown amount of trials (Oberauer, 2002, 2005). It seems clear that the context of how relevant the information outside the focus of attention is does affect its state as measured by behavioral studies (Zokaei, Ning, Manohar, Feredoes, & Husain, 2014). However, it is not possible to know the neural states during these behavioral studies.

Importantly, in Study I – III and other similar studies, the non-consciously presented stimuli was always relevant for prospective use, which might explain why non-conscious information outside the focus of attention seem to be maintained as persistent neural activity. The idea that persistent neural activity is dependent on prospective use might therefore reconcile the two conflicting cases where non-consciously perceived information can be maintained as persistent neural activity, and that consciously perceived but deprioritized information can be retained as short-lived latent neural changes. The length until deprioritized information might become relevant again and working memory load could also be factors that contribute to the neural state information is retained in. Study III had minimal working memory load, which might lend more neural resources to the non-conscious maintenance. However, Study I had a higher (but still small) load because the solution to the math task (T1) was simultaneously held in working memory consciously. It is interesting to note that many believe that attentional blink is the consequence of T1 being encoded into working memory (Bowman & Wyble, 2007). The letter (T2) is thus non-consciously encoded in working memory while participants are holding T1 in focus of attention while solving it. This suggests some level of parallel processing in terms of multiple prospective actions and encoding.

In sum, the findings of studies I – III and the current literature converge on the finding that non-consciously presented information can be maintained in working memory as persistent neural activity. However, we still need to consider the possibility of influences from other kinds of memory.

Iconic memory

Iconic memory can survive distractors in the same spatial location as long as they are of a different kind than the memoranda (Pinto et al., 2013), which was the case in Study I (letters vs. numbers). However, Soto et al. (2011) showed short-term retention despite conscious and non-conscious distractors of same type and spatial position as the memoranda, which should have overwritten iconic memory. Similarly, iconic memory is not dependent on prospective use (Pan et al., 2013). Although iconic memory is commonly thought to be a very brief (< 1 second) high capacity memory, it can last up to at least 4 seconds under certain conditions (with diminished capacity) when information is consciously presented (Averbach & Sperling, 1961; Sligte et al., 2008). In the name of prudence, I will assume that non-consciously presented information can last as long as consciously presented information. That means that iconic memory could have influenced the results of Experiment II in Study II, and the pre-fMRI session in Study III since the delay period was 5 seconds without distractors. Similarly, it could explain the retention after 1.5 seconds (Dutta et al., 2014) and 800 milliseconds (King et al., 2016). However, it seems unlikely that residual neural activity from non-consciously presented information would persist between 5 and 15 seconds. Even if it could, iconic memory would only predict BOLD signal/neural activity in visual regions during retention (Sligte et al., 2009), and is therefore inconsistent with the frontal involvement (Sligte et al., 2011) found in Study I & III and global oscillations in (King et al., 2016). Taken together, iconic memory is unlikely to have influenced the non-conscious short-term retention, with a cautious exception for the experiments with 5 second (Exp. II and pre-fMRI) and 1.5 second delay periods (Dutta et al., 2014).

Priming

Contrary to iconic memory, non-consciously encoded perceptual priming have been shown to last several minutes in some cases (Bar & Biederman, 1998, 1999; Gaillard et al., 2007). Similarly to iconic memory, non-conscious priming might be affected by distractors that share spatial position and object features at the same processing level, but is potentially quite robust against other distractors (Henson, 2003; Marsolek, 2008). The other studies used memory probes that never were identical to the memoranda, which make priming an unlikely explanation (Dutta et al., 2014; King et al., 2016; Soto et al., 2011). Priming can occur when stimuli are similar (not only identical) to the second repetition, usually in terms of semantic/conceptual relatedness or similar perceptual features in identification tasks. However, in these tasks they were to judge if the memorandum was clock-wise or counter clock-wise to the memory probe, thus the similarity or closeness to the probe cannot facilitate task performance.

In Study II we found that the conjunction of object identity and spatial position could be retained. Initially, we reasoned that such a feat would set working memory apart from priming. However, high-level visual object recognition has, contrary to common belief, been suggested to be position-dependent (Kravitz, Kriegeskorte, & Baker, 2010; Kravitz, Vinson, & Baker, 2008). If high-level visual object recognition automatically processes the object's spatial position, then it should be possible for the conjunction of identity and position to be retained by non-conscious priming mechanisms. However, Study II and III did not find any signs of non-conscious priming effects in reaction times (Study I could not index priming because all four alternative letters were always shown during discrimination).

In Study III, there was higher BOLD signal change related to non-conscious discrimination when comparing memory probe non-match to match, and to baseline, which partly overlapped with the same comparison for conscious discriminations. These findings seems at odds with priming (consistent with the absence of priming indexed by reaction time) since there is no repetition suppression effect in the traditional sense (Henson, 2003). However, it is consistent with an antipriming interpretation of repetition suppression (Marsolek et al., 2010). It is not clear how to interpret the non-match findings in terms of priming. As mentioned previously, priming is too inclusive as a kind of memory, and can include latent structural neural changes, and residual activity, but I would also add persistent activity.

The conscious conditions in Study II & III shows strong priming effects in terms of reaction times (~ 250 milliseconds), and antipriming effects in BOLD signal (non-match > match). A facilitation of ~250 milliseconds seems to indicate more than only structural changes, since that is enough time for a sensory input to be sent from the retinae, processed in the cortex, and transformed to motor output. It therefore seems likelier that such strong priming effects mainly stems from active maintenance and preparation for response. If these effects are seen in a conscious working memory task, then finding them in the non-conscious condition cannot exclude working memory mechanisms by themselves. Certainly not when there are other indicators consistent with non-conscious working memory maintenance, which themselves are compatible with the antiprimed effects in BOLD signal.

Taken together, there is no indication that the non-conscious short-term retention is completely caused by non-conscious priming. However, there is a slight possibility that some priming mechanisms other than those overlapping with working memory mechanisms could have influenced the antiprimed repetition suppression effect to some extent.

Hippocampi-based memory

Non-consciously encoded hippocampi-based memory have been shown to last for minutes, with other stimuli presented between encoding and retrieval, and should therefore not have a problem with distractors (Degonda et al., 2005; Reber et al., 2012). Contrary to findings in Study I – III, hippocampi-based long-term memory predicts the absence of delay-related BOLD/neural signal since it consists of latent (structural) neural changes. However, the presence of sustained BOLD signal during the delay period does not exclude influences from the hippocampi. We therefore investigated the presentation and discrimination epochs for signs of hippocampi-related encoding and/or retrieval respectively. In Study I there was no indication of hippocampi involvement during presentation or discrimination when lowering the threshold. In Study III, weak BOLD signal change appeared in the right hippocampus during stimuli-presentation. The absence of significant BOLD signal change in the hippocampi cannot exclude hippocampi-based influences (might be too weak to detect), but is likely not the whole story given the delay-related findings. Especially, when the medial temporal lobes including hippocampi are important for working memory tasks involving novel associations among items or features and long delay periods (Jonides et al., 2008; Ranganath & Blumenfeld, 2005; Ranganath & D'Esposito, 2001). It could therefore be the case that non-conscious hippocampi-based encoding of the conjunction of object identity and spatial position (for that specific trial) facilitates or necessitates the working memory maintenance of the associated features, and might be weakly active during the delay period as well. It can therefore not be assumed that there are no hippocampi-based influences, as latent neural changes or as working-memory related maintenance.

Interim summary

Overall, Studies I – III are consistent with the current literature on non-conscious short-term retention (Bergström, 2011; Dutta et al., 2014; King et al., 2016; Pan et al., 2013; Soto et al., 2011; Soto & Silvanto, 2014), which together makes a compelling case that non-consciously presented information can be maintained in working memory as persistent neural activity. It seems unlikely that iconic memory contributed to the short term retention. Although we cannot completely exclude influences from hippocampi-based memory and/or priming, we do not need to do so, because working memory, hippocampi-based memory, and priming are not mutually exclusive. Indeed, there are probably few if any memory tasks that isolate a single kind of memory (Tulving, 2002). Critically, however, these findings imply the existence of non-conscious working memory.

Critique against non-conscious memory

There has been some critique against the findings related to non-conscious working memory in Studies I & II and in (Dutta et al., 2014; Soto et al., 2011), but concerns any kind of non-conscious memory (Samaha, 2015; Stein, Kaiser, & Hesselmann, 2016). This criticism can be described as two main concerns: (i) that subjective measures of experience might be biased towards underreporting, and (ii) that participants, after non-conscious perception, made a premature guess, and held the guess consciously in working memory. According to (i) and (ii), the findings can therefore be explained by conscious working memory rather than non-conscious working memory. In addition, Stein et al. (2016) directed more specific critique against the way d' was used by (Dutta et al., 2014; Soto et al., 2011), for their reply see (Soto & Silvanto, 2016).

The possibility of bias when relying on subjective measures is well known, and applies to non-conscious cognition more generally. As described in the above section "Inferring the absence of conscious experience", it is possible that participants report "no experience" when they in fact had a "vague experience" either systematically or accidentally. Subjective measures are liberal, and risk overestimating the non-conscious effect. However, objective measures are conservative, and run the opposite risk of underestimating or failing to find an existing non-conscious effects. Furthermore, both measures tend to converge on similar results, and might best be viewed as complementary.

We therefore reasoned that it was most sensible to start with subjective measures since a null effect likely meant that there was nothing to find, while a significant effect could be further investigated with more conservative measures. Studies I & II should therefore be interpreted with that caveat in mind rather than being dismissed altogether. However, Study III showed sustained BOLD signal despite subjective reports of "no experience", objective direct measures (discrimination and detection) being at chance, and indirect measures (reaction times) not being different from each other or baseline. Additionally, control analyses in Study III showed that neither the univariate nor the MVPA results can be explained by accidental mislabeling of PAS. Study III, therefore, not only shows strong support for non-conscious working memory with an objective criterion of experience, but also strengthens the validity of studies I & II by partly replicating (and extending) their results of frontal involvement by maintenance of spatial information.

If, as Stein et al., (2016) suggests, participants had decided what to guess prematurely, and held the conscious guess in working memory during the delay, then reaction times should be similar for conscious and non-conscious trials. However, control analyses for Study I – III showed that discrimination reaction times were much faster for conscious compared to non-conscious trials. Furthermore, such conscious guesses would apply equally to baseline trials, and can therefore not explain the BOLD signal difference between non-conscious and baseline trials in Study I & III.

There is another concern that has been raised against non-conscious high-level processing in general. The concern is that non-conscious high-level features are driven by conscious experience of low-level features. For example, maybe only semantic information can be processed non-consciously when color and/or shape are processed consciously. This line of argument goes for both subjective and objective criteria of experience, and depends on what processing level the subjective or objective task is measuring. For example, if the task is to discriminate between faces, it is possible that face discrimination performance is at chance, while still consciously experiencing the shape and spatial position of the faces. If the subjective or objective measures only measures the higher level it is impossible to know if the lower level is conscious or not. There is empirical support for this hypothesis with the attentional blink (Elliott, Baird, & Giesbrecht, 2016) and CFS (Gelbard-Sagiv, Faivre, Mudrik, & Koch, 2016).

It is not entirely clear that this is a criticism per se. It seems possible to argue that if the research question concerns a specific processing level, it does not matter if lower levels are experienced or not. However, I will address this issue here anyway. In Study I – III the participants were always instructed to use the PAS scale on all levels of the stimuli. That is, "no experience" was only to be used if there was no experience what so ever. If something undefined was experienced they were to use "vague experience", and if they experienced enough to identify it "almost or clear experience" should be used.

However, using the PAS as instructed would arguably not be enough during the attentional blink paradigm (Study I), since the lowest level that sets letters apart from numbers is still pretty high. That is, if participants consciously experienced the letters' low-level features, such as color and part of the shape, the features would be too similar to the distractors' (numbers) low-level features for the participants to know the difference. It is therefore possible that the participants consciously perceived some low-level features (without knowing it), and that it enabled or facilitated the non-conscious processing of higher-levels necessary to discriminate between letter vs. number, different letters, and maybe the retention. This phenomenon might even be the reason why paradigms like attentional blink and inattentional blindness show relatively strong non-conscious effects. The CFS paradigms are less susceptible to this kind of criticism assuming the PAS instructions were followed, because experience of low-level features should be easy to discern from the mondrians, and reported as vague experiences. Furthermore, in Study III the performance on the detection (stimuli presence vs. absence) task was also at chance. This means that the participants were not able to detect any low-level features by objective measures either. The non-conscious BOLD signal can therefore not be attributed to conscious low-level processing based on an objective criterion (detection performance being at chance).

Overall, the critique raised against a non-conscious working memory interpretation does not seem to be able to explain the results of the studies in this dissertation or similar studies.

Is non-conscious memory weak and fragile?

The non-conscious behavioral and neural effects were statistically significant but small, which raises a couple of questions. Why are the effects small? Do such small effects have any theoretical and/or practical relevance? What do the small effects tell us about non-conscious working memory? By showing small but statistically significant effects, we can infer that non-conscious working memory exists. This is itself an important theoretical implication, and as discussed in later sections, can have implications on models of working memory and theories of consciousness. The practical significance will also be discussed further in a later section. If the measured effect of non-conscious memory is the full extent of its influence on behavior it might seem rather insignificant. What possible significance can such a weak and fragile memory effect have in our everyday lives all things considered? I will briefly discuss whether the observed non-conscious memory effects are fragile, and whether the measured effects are the full extent of its influences on behavior.

The non-conscious effects were small, but replicable over experiments, except for the fMRI session in Study III, which dropped to chance level. Why the performance dropped to chance is unclear. The mirror stereoscope used for CFS was slightly different in the fMRI scanner, which might have caused a slight rotation of the stimuli, but I do not see how that could affect retention. The delay period was longer (5 – 15 seconds) in the fMRI scanner compared to 5 seconds in the pre-fMRI session, which could have been a factor, but considering the 15 second retention in Study II that should not have been a problem by itself. In addition, the fMRI scanner environment was different, participants had to lie down in darkness, tightly squeezed, with rhythmic noise, for a long time without much rest/or movement, which causes drowsiness and less than optimal focus. The latter (maybe combined with the former changes) might have contributed to the drop in performance.

Given these overall small non-conscious effects it is tempting to assume that only a few number of trials contain memory traces, and therefore produce small mean effects. However, it is not possible to know which memory epoch(s) that caused the small non-conscious effects, and the performance drop in Study III. The bottleneck could be weak perception, maintenance, and/or use of information. That is, maintenance could be really reliable without much loss in fidelity, but if the perceived representation is weak and noisy, that is what will be maintained. Similarly, if perception and maintenance are relatively robust, that will not show behaviorally unless this non-conscious information is actually put to use. Given that the BOLD signal during the delay period did not correlate with memory performance in Study I, and we found maintenance related BOLD signal despite memory performance being at chance in Study III, one possibility is that the use of information is a smaller bottleneck than maintenance. It could therefore be the case that non-conscious memory traces were present on most trials, but we cannot know if that is the case. The fact that non-conscious effects are small compared to conscious effects are likely due to the techniques that render information non-conscious. Ideally, the relevant information would be presented consciously, but used non-consciously, for maximum effect. However, such circumstance creates difficulty in controlling that the information is not used consciously. A solution is to present information non-consciously, that is, through some manipulation the information processing is degraded to such an extent that it is not experienced. Therein lies the problem. The only way to reliably measure effects of non-conscious processing is to severely limit them, and therefore inherently underestimate them. In the studies of this dissertation we have tried to isolate the non-conscious memory effect under relatively conservative experimental conditions, and they should be interpreted within that context. The full extent of non-conscious working memory can therefore not be inferred from these findings alone, and could be much larger than measured.

There are clinical cases where non-conscious processing does not appear to be as severely limited as in experiments with healthy participants. Blindsight patients have damage to their primary visual cortex, and report that they are blind in the affected visual field. They can nevertheless perform very well at certain discrimination tasks where the information is presented within their blind visual field. For some discrimination tasks the performance can be close to 100% correct when 50% is chance, while being lower in other tasks (Cowey, 2010). It is therefore possible that non-conscious working memory (and non-conscious cognition in general) has a much greater impact on our behavior than we are able to reliably measure in most experiments.

Implications for working memory models

I will here discuss the implications of non-conscious working memory maintenance as persistent neural activity on contemporary (state-based) working memory models. Based on all the data espoused through this dissertation, there are, somewhat simplified, four possible memory states: (i) latent (non-conscious) long-term memory with varying memory strength, (ii) latent (non-conscious) short-term memory with varying memory strength, (iii) non-conscious residual activity, (iv) conscious residual activity, (v) non-conscious persistent activity, and (vi) conscious persistent activity. The allocation of neural resources (attention) is what determines if a memory is in a latent or active state. What causes an active memory to become conscious is still a matter of debate, but the amount of attention (i.e., neural activity) is likely an important part in the neocortex. Whether a memory is active for prospective use is what distinguishes working memory from other active memory such as sensory memory (table 1).

Table 1. Taxonomy of memory states

Index	Memory	Neural substrates	Conscious	*For* prospective use
(i)	LTM	Latent change	No	no
(ii)	STM	Latent change	No	no
(iii)	SM	Residual activity	No	no
(iv)	SM	Residual activity	Yes	no
(v)	WM	Persistent activity	No	yes
(vi)	WM	Persistent activity	Yes	yes

Note: LTM = long-term memory, STM = short-term memory, SM = sensory memory, WM = working memory.

The existence of the (i) and (vi) state is not controversial, and therefore taken for granted here. Studies I – III, together with a growing amount of support (Bergström, 2011; Dutta et al., 2014; King et al., 2016; Pan et al., 2013; Soto et al., 2011; Soto & Silvanto, 2014), suggests that the (v) state exists, and other studies suggest that the (ii) state exists (LaRocque et al., 2013; Larocque et al., 2014; Lewis-Peacock et al., 2012; Nelissen et al., 2013). Given these constraints, a model of working memory should consist of at least one state resembling (v) and another resembling (vi). That is not to say that (i) and (ii) is not crucial for working memory. Indeed, working memory is the interaction between sustained attention and (i & ii).

As previously stated, most working memory models assume that the most privileged state information can be in, is the focus of attention, where information is maintained by persistent neural activity, and is consciously experienced. McElree's (2001; 2006) 1-state model posits a conscious focus of attention, and anything outside the focus is non-conscious latent long-term memory with varying strengths depending on how recent it was activated. It therefore assumes that conscious experience is necessary for working memory. Such 1-state models are inconsistent with findings of non-conscious persistent activity because they do not allow for the (v) state. In contrast, 2-state models usually allow for a non-conscious state as the focus of attention only captures a subset of all working memory content. Cowan's model (1995, 2001, 2008) may be compatible with these constraints because it assumes a non-conscious portion of activated long-term memory in addition to the conscious focus of attention. However, the activated long-term memory is usually thought to decay if not attended again, and thus resembles (iii), but it is unclear if such attendance can be non-conscious as in (v). For clarity, it might be reasonable to distinguish these two cases as two separate states. Fuster's model (1995, 2003a, 2015) is consistent with these restraints, as it explicitly postulates that conscious experience is not necessary for working memory, and therefore predicts states like (v).

Oberauer's 3-state model is difficult to evaluate in terms of conscious states because it remains neutral on that matter. However, I will try to make reasonable inferences about the different states. The focus of attention is presumably conscious, but it is less clear if what he calls the "state of direct access" or more recently "the broad focus of attention" is conscious or not. Since this broad focus is supposed to be similar to Cowan's focus of attention, it seems reasonable to assume that this state also is conscious. Oberauer also posits an activated long-term memory that likely is non-conscious, but arguably, might not depend on prospective use (Oberauer, 2001), and might consist of latent synaptic changes (LaRocque et al., 2013; Larocque et al., 2014; Lewis-Peacock et al., 2012; Nelissen et al., 2013). It therefore seems like the activated long term memory state is more like (ii) than (iii), which makes this model inconsistent with findings supporting the existence of (v). However, as mentioned earlier, it is difficult to evaluate because the deprioritized information are sometimes never relevant and other times only relevant on a proportion of the trials. A more rigorous evaluation of how the probability of prospective use of deprioritized information affects the psychological and neural state of such information will surely clarify the ambiguity, and may even lead to more working memory states.

Implications for theories of consciousness

Historically, working memory and conscious experience has been tightly linked. Investigations into the neural correlates of consciousness showed that neural activity/BOLD signal in higher visual regions, and the prefrontal and parietal cortex correlated with conscious perception. The frontoparietal regions involved in conscious perception were also involved in other functions such as working memory and attention (Naghavi & Nyberg, 2005). These findings have inspired prominent theories of consciousness such as the *Global Neuronal Workspace* (GNW; Dehaene & Changeux, 2011; Dehaene et al., 2014; Dehaene & Naccache, 2001), and the *Higher-Order Thought* (HOT; Lau & Rosenthal, 2011a, 2011b) theories. According to the GNW theory, non-conscious information is processed in local modules and is relatively short-lived (< 500 milliseconds). For that information to become consciously experienced and maintained in working memory, the information needs to be globally broadcast via the frontoparietal network.

However, the findings of Studies I - III seem to be inconsistent with the GNW theory. Firstly, that non-consciously presented information can be retained for longer than assumed (up to 15 seconds). Secondly, that this non-conscious short-term retention seems to be supported by relatively global persistent fronto-occipital (Study III) and fronto-cerebellar (Study I) neural activity. Thirdly, that conscious experience is not necessary for working memory maintenance. Fourthly, that non-consciously presented information can partly activate the cognitive control network during non-conscious discrimination in Study III, which adds to previous findings (Lau & Passingham, 2007; van Gaal et al., 2010). The relatively global (and presumably some local) recurrent activity between frontal and posterior regions also seems to contradict theories that postulate recurrent activity in sensory regions to be sufficient for conscious experience (Lamme & Roelfsema, 2000; Lamme, 2006).

According to the HOT theory, for some information to become conscious there must be a second higher-order thought (in frontoparietal regions) about a first-order state (in sensory regions). The HOT theory does not seem to be affected much of the findings since it is neutral regarding non-conscious functions. Although, it might be hard to see how non-conscious working memory would not constitute some kind of HOT state of the maintained information. Furthermore, a non-conscious HOT about a non-conscious first-order state should cause a conscious experience (Lau & Rosenthal, 2011b), which clearly is not in the case of non-conscious working memory. However, HOT theorists would likely sidestep this concern by claiming that working memory is not the right kind of second-order state. First-order theories of consciousness are not affected as they generally posit that it is the level of activity in higher visual regions that determines if that regional content is experienced or not, and therefore allow non-conscious executive functions (Fuster, 1995, 2003a, 2015; Zeki, 2001, 2008).

Practical implications

What, if any, practical implications are there with the findings presented within this dissertation? A common question I receive when telling people about these findings is “so you have essentially shown that subliminal brainwashing is possible?” I therefore thought it relevant to shortly discuss if not too much about what practical implications there are, then at least what practical implications there are not. It seems many know about the infamous study by James Vicary from 1957. Vicary claimed to have flashed the words “Eat Popcorn” and “Drink Coke” non-consciously for 1/3 of a millisecond on the movie screen while showing the movie Picnic to an unsuspecting audience, and found an 18% and 58% increase in Coke and popcorn sales respectively. Although these claims were later revealed to be fabrications made to boost sales for Vicary’s failing marketing business, it seems to have had a lasting effect legally and in public opinion. Pratkanis (1992) critically reviewed the attempts to replicate Vicary’s claims, and concluded that subliminal persuasion is no more effective than placebo, and of no use in marketing. How is that reconciled with non-conscious memory?

Although perception and memory are necessary for persuasion, they are not sufficient for it. Showing that non-consciously perceived information can be maintained in working memory or any other memory does not imply that peoples’ opinions can be changed without them being aware of it. Non-conscious persuasion is thus a different research question with little support (Pratkanis, 1992). However, there have been some recent attempts to non-consciously influence participants' preference towards specific brands. Such subliminal advertising has found small non-conscious effects, but only if the non-conscious information is goal-relevant, the participants are motivated to pursue the goal, and the brand is not already habitually consumed (Bustin, Jones, Hansenne, & Quoidbach, 2015; Karremans, Stroebe, & Claus, 2006; Strahan, Spencer, & Zanna, 2002; Verwijmeren, Karremans, Stroebe, & Wigboldus, 2011). For example, participants can be non-consciously influenced to get a slight preference for "Lipton Ice", but only if they are thirsty (Karremans et al., 2006). It is unclear if this goal-dependent retention is in anyway related to non-conscious working memory. It could also be related to something like stimulus-reward mapping, which has been shown to occur when non-consciously presented stimuli are temporally paired with water drops as rewards to thirsty participants (Seitz, Kim, & Watanabe, 2009).

Other studies have found that non-consciously presenting affective information (Blanchfield, Hardy, & Marcora, 2014), and positive age stereotypes for older adults (Levy, Pilver, Chung, & Slade, 2014) can affect physical performance. Small non-conscious effects could potentially have a significant impact for large businesses and professional athletes, but consciously presenting information might result in larger effects. If so, presenting non-conscious information seems rather pointless. I think it is likely that consciously presenting information will have a larger (positive or negative) effect. However, Levy et al. (2014) found that the non-conscious intervention had greater effect than the conscious intervention on physical performance. Perhaps, non-conscious information might have smaller effects, but can more easily by-pass endogenous modulation of perceived information.

Synthesis

When external sensory information becomes available it is perceptually processed by reactivation of related latent (phylogenic and ontogenetic) memory. During normal conditions this perceptual reactivation is high enough to (somehow) cause a conscious experience of the sensory information (figure 6A). In addition to neural activity, this reactivation can cause short-term or long-term latent structural neural changes in the reactivated neural networks. When the external sensory information no longer is available there is nonetheless residual neural activity (i.e., sensory memory) for a few seconds (figure 6A). Any short-lived latent neural changes can last up to a minute, while long-lived latent neural changes can last for a life time, and can later be endogenously or exogenously reactivated. Initially the residual activity will be high enough to cause a residual conscious experience, but quickly drop below the threshold for conscious experience. A subset of the sensory memory can be temporarily maintained for longer periods until prospective use (i.e., in working memory).

The information is maintained in working memory by persistent recurrent activity between task-related prefrontal and posterior sensory regions. If the persistent neural activity is high enough (in the focus of attention), it will be consciously experienced, but is otherwise non-consciously maintained for prospective use as below threshold activity outside the focus of attention (figure 6B). The persistent activity during working memory can also cause short-term or long-term latent changes.

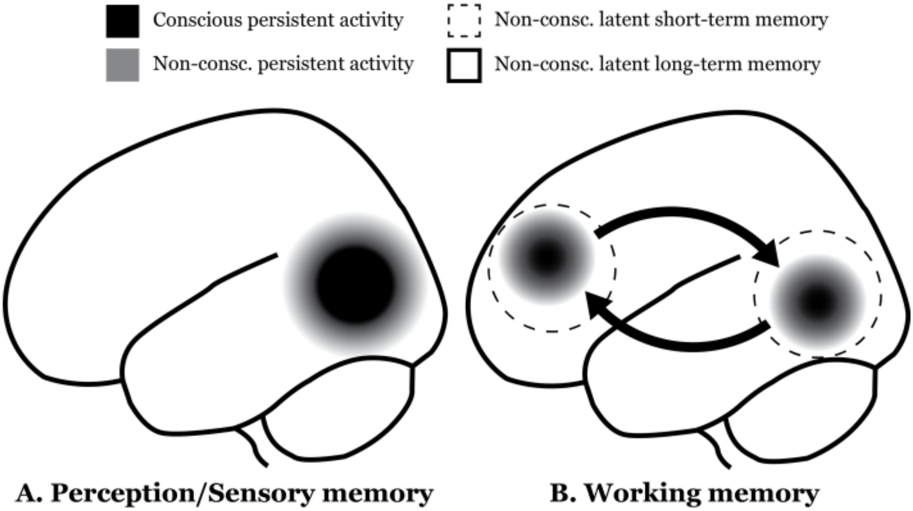

Figure 6. Illustration of conscious and non-conscious memory states.

If another subset of the actively maintained (or latently retained) sensory information becomes prioritized, then neural resources are endogenously reallocated toward the newly relevant information. The deprioritized information will either be maintained non-consciously as persistent below threshold activity or latently retained neural changes depending on how relevant it is for prospective use. Similar shifts between conscious and non-conscious states can occur in the prefrontal cortex related to the main task and potential subtasks or contingencies. For example, when mentally performing difficult multi-step arithmetic, all the pertinent information will be maintained while solving the task. However, when focusing on one of the multiple steps you might not be conscious of information concerning the other steps until focus shifts.

Perceived information does not have to be consciously experienced in the focus of attention in order to be maintained in working memory. Non-consciously perceived information can be maintained from passively viewing suppressed information (Study II - III), and while actively engaged in another working memory manipulation task (Study I). It might even be possible that focusing attention on non-consciously maintained information can boost the activity above the threshold for experience. Indeed, it has been suggested that retrospectively cuing endogenous attention to the positon of a non-consciously presented stimuli can retroactively cause conscious experience of them (Sergent et al., 2012; Thibault, van den Berg, Cavanagh, & Sergent, 2016)

For example, sometimes when thinking intensively about something, and someone asks something else, you might be conscious of that person saying something but not of what they are saying. However, when you are finished with whatever you were focusing on, you can shift attention toward what was previously said, and retroactively experience and understand it. It is possible that such retroactive experience can occur because the information was maintained non-consciously below the threshold of experience (given that you intend to respond) until attention is shifted to it.

These dynamic shifts between working memory, short-term, and long-term memory states enable an extremely flexible use of information. Whether a state is conscious or non-conscious seems to depend on the amount of allocated attentional resources, and can dynamically shift depending on task demands. Non-conscious working memory states are likely important for the normal functioning of working memory as we know it. Most studies of working memory have not specifically investigated conscious or non-conscious working memory, and have therefore essentially investigated the conjunction of conscious and non-conscious working memory.

Limitations and future directions

The empirical studies of this dissertation, as is the case with most studies, come with some caveats to consider, and to be addressed in future research. I will here briefly mention these limitations, and suggest research questions for future studies. It is not possible to know the exact functions of the prefrontal (Study I & III) and cerebellar (Study I) involvement in relation to non-conscious short-term retention based on these studies. We can only infer their functions based on previous literature about the brain, and various kinds of memory mechanisms. Future research will have to address these issues by designing experiments to specifically tease apart the different functions of these, and possibly other, brain regions during non-conscious retention. These studies can, for example, use MVPA to classify different memory content from the different brain regions similarly to what we did in Study III, but for other stimuli features and cognitive processes. Is the neural activity in the frontal cortex truly task-related and/or related to endogenous attentional control of the retained stimuli as predicted by working memory? What is the role of cerebellum in Study I? Could it relate to stimulus-response mappings similar to what is seen during conditioning? The role of cerebellum is not even clear when it comes to conscious working memory.

In our studies we do not know if the observed memory performance and BOLD signal are dependent on prospective use, which is necessary for it to be conceptualized as working memory. That is, if there was one condition with prospective action, and one without, would the memory and/or MVPA performance drop to chance level in the condition without prospective action? Although it has been previously shown that non-conscious retention can be dependent on prospective use in (Pan et al., 2013), it would be interesting to replicate such findings with neuroimaging.

It is important to keep in mind that the neuroimaging findings of Study I & III are correlations, and we can therefore not know anything about the causal mechanisms of these findings in relation to non-conscious short-term retention. Neither can we know if there is an interaction between fronto-cerebellar (Study I) and fronto-occipital (Study III) cortices, or if they only co-occur independently. Future studies can address the issue of causality by using TMS or patient groups with various lesions to determine the brain regions that are causally involved in non-conscious short-term retention (Feredoes et al., 2011). It is possible to use TMS to alternately disrupt the normal functioning of prefrontal and stimulus-specific cortical regions during non-conscious retention, similarly to what was done on primates with cortical cooling (Fuster et al., 1985). Would disrupting the neural activity cause a decrease in non-conscious memory performance, or cause the MVPA classification performance to drop to chance level in the region not disrupted by TMS? Another possibility is to investigate functional interactions by using functional connectivity analysis to see if voxels in the prefrontal cortex correlate with voxels in other regions of the brain during retention.

Another caveat is that the MVPA in Study III is not based on variance, and is therefore considered a fixed effects analysis instead of a random effects analysis. Fixed effects analyses are not as generalizable as random effects analyses, and can only infer that the observed effects are significant for the group that is tested, while random effects analyses can infer that the effects apply to others (not tested) in the same population.

There are of course many other questions, not linked to limitations in the studies of this dissertation, which should be pursued in future research. It is not known how the conscious and non-conscious mechanisms of working memory interact with each other. For example, if the same stimuli with the same features (e.g., identity and position) are non-consciously presented in two conditions, but the conscious instructions differs such that in one condition only the identity is to be remembered, and in the other only the position is to be remembered, will the instructed information be preferentially retained? In the opposite direction, to what extent can conscious working memory and decision-making make use of non-consciously presented information?

It would also be interesting to follow up on the different states of deprioritized information in working memory. For example, during which conditions is it possible to use MVPA to decode maintained information? Here one could manipulate the length of delay durations or vary the probability that the deprioritized information needs to be used later between 0% and 100% in the same experiment. Another unknown factor is the capacity of this non-conscious short-term retention. We know that conscious working memory has a very limited capacity, but it is not clear if the non-conscious retention share this limited capacity, has a little more, or perhaps a lot more capacity? Similarly, the idea of non-conscious iconic memory has not been investigated as far as I know. Is non-consciously presented information briefly retained with much higher capacity before a subset can be maintained for a more durable time? Conscious working memory is also known for its flexible use of information. To what extent can non-conscious information be manipulated for prospective use? How does non-conscious working memory relate to other kinds of implicit learning based on the basal ganglia?

In sum, the conclusions based on the studies of this dissertation about the existence of non-conscious working memory should be seen as tentative, not conclusive. As briefly mentioned in this section, these studies have presented us with fertile ground, ripe not with limitations, but with opportunity for future research.

Conclusions

Overall the findings of this dissertation show that non-consciously presented information (identity and/or position) can be retained for several seconds, sometimes with distractors, and is associated with sustained BOLD signal in frontal and posterior regions. Study I found sustained BOLD signal change in the prefrontal cortex and cerebellum, and Study III found sustained BOLD signal in the frontal cortex, and stimulus-specific BOLD signal in the occipital cortex. The sustained BOLD signal likely represents persistent neural activity of some kind, and is consistent with working memory maintenance of the non-conscious information. The non-conscious working memory effects are small in the relatively conservative experimental conditions of these studies, but do not necessarily mean that non-consciously maintained information does not have larger behavioral relevance in our everyday lives. The current results add constraints on models of working memory and theories of consciousness. Working memory models need to accommodate the possibility that information can be non-consciously maintained as persistent neural activity for prospective use, and not only as residual activity or latent neural changes. The current results also contradict the GNW theory’s assumptions that consciousness is necessary for durable retention, global recurrent activity, and working memory.

References

Abrams, J., Barbot, A., & Carrasco, M. (2010). Voluntary attention increases perceived spatial frequency. *Attention, Perception & Psychophysics*, *72*(6), 1510–1521. doi:10.3758/APP.72.6.1510

Almeida, J., Mahon, B. Z., Nakayama, K., & Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(39), 15214–8. doi:10.1073/pnas.0805867105

Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2011). Distilling the neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*, *36*(2), 737–746. doi:10.1016/j.neubiorev.2011.12.003

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory*. New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971). The Control Process of Short-Term Memory. *Scientific American*, 1–43. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.398.2237&rep=rep1&type=pdf

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *The Bell System Technical Journal*, 309–329.

Averbach, E., & Sperling, G. (1961). Short Term Storage of Information in Vision. In C. Cherry (Ed.), *Information theory* (pp. 196–211). Washington, DC: Butterworth & Co.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. doi:10.1016/j.tics.2012.06.010

Baars, B. (2005). Global workspace theory of consciousness : toward a cognitive neuroscience of human experience. *Brain*, *150*, 45–53. doi:10.1016/S0079-6123(05)50004-9

Baars, B., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, *7*(4), 166–172. doi:10.1016/S1364-6613(03)00056-1

Baars, B. J. B. (1988). *A cognitive theory of consciousness. A Cognitive*

Theory of Consciousness. Cambridge University Press. Retrieved from http://www.loc.gov/catdir/description/cam032/87020923.html

Baddeley, A. (1983). Working memory. *Philosophical Transactions of the Royal Society of London*. doi:10.1111/j.1365-2834.2011.01270.x

Baddeley, A. (1992). Consciousness and working memory. *Consciousness and Cognition*, (1), 3–6.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews. Neuroscience*, *4*(10), 829–39. doi:10.1038/nrn1201

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*, 1–29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A., & Hitch, G. J. (1974). Working Memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (pp. 47–89). New York: Academic Press.

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews. Neuroscience*, *10*(9), 659–69. doi:10.1038/nrn2667

Bahrami, B., Lavie, N., & Rees, G. (2007). Attentional load modulates responses of human primary visual cortex to invisible stimuli. *Current Biology : CB*, *17*(6), 509–13. doi:10.1016/j.cub.2007.01.070

Bar, M., & Biederman, I. (1998). Subliminal Visual Priming. *Psychological Science*, *9*(6), 464–468. doi:10.1111/1467-9280.00086

Bar, M., & Biederman, I. (1999). Localizing the cortical region mediating visual awareness of object identity. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(February), 1790–1793. doi:10.1073/pnas.96.4.1790

Barak, O., & Tsodyks, M. (2014). Working models of working memory. *Current Opinion in Neurobiology*, *25*, 20–24. doi:10.1016/j.conb.2013.10.008

Barton, J. J. S. (2011). Disorder of higher visual function. *Current Opinion in*

Neurology, *24*(1), 1–5. doi:10.1097/WCO.0b013e328341a5c2

Bayne, T., Hohwy, J., & Owen, A. M. (2016). Are There Levels of Consciousness? *Trends in Cognitive Sciences*, *20*(6), 405–413. doi:10.1016/j.tics.2016.03.009

Bergström, F. (2011). *Tracing subliminal memory traces*. Umeå University. Retrieved from https://www.diva-portal.org/smash/get/diva2:439387/FULLTEXT01.pdf

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–536. doi:10.1016/j.tics.2011.10.001

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. doi:10.1093/cercor/bhp055

Blanchfield, A., Hardy, J., & Marcora, S. (2014). Non-conscious visual cues related to affect and action alter perception of effort and endurance performance. *Frontiers in Human Neuroscience*, *8*(December), 1–16. doi:10.3389/fnhum.2014.00967

Bloch, F. (1946). Nuclear induction. *Physical Review*, *70*(7-8), 460–474. doi:10.1103/PhysRev.70.460

Block, N. (1995). On A Confusion About a Function of Consciousness. *Brain and Behavioral Sciences*, *18*(2), 227–247. doi:10.1017/S0140525X00038188

Boly, M., Seth, A. K., Wilke, M., Ingmundson, P., Baars, B., Laureys, S., … Tsuchiya, N. (2013). Consciousness in humans and non-human animals: recent advances and future directions. *Frontiers in Psychology*, *4*(October), 625. doi:10.3389/fpsyg.2013.00625

Bostan, A. C., Dum, R. P., & Strick, P. L. (2013). Cerebellar networks with the cerebral cortex and basal ganglia. *Trends in Cognitive Sciences*, *17*(5), 241–54. doi:10.1016/j.tics.2013.03.003

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, *114*(1), 38–70. doi:10.1037/0033-295X.114.1.38

Boynton, G. M., Engel, S. a, Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(13), 4207–4221.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 4–4. doi:10.1167/11.5.4

Broadbent, D. E. (1958). *Perception and communication*. Oxford: Oxford University Press.

Brunoni, A. R., & Vanderhasselt, M.-A. (2014). Working memory improvement with non-invasive brain stimulation of the dorsolateral prefrontal cortex: A systematic review and meta-analysis. *Brain and Cognition*, *86C*, 1–9. doi:10.1016/j.bandc.2014.01.008

Bustin, G. M., Jones, D. N., Hansenne, M., & Quoidbach, J. (2015). Who does Red Bull give wings to? Sensation seeking moderates sensitivity to subliminal advertisement. *Frontiers in Psychology*, *6*(June), 1–6. doi:10.3389/fpsyg.2015.00825

Buxton, R. B. (2001). The elusive initial dip. *NeuroImage*, *13*(6 Pt 1), 953–958. doi:10.1006/nimg.2001.0814

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*(1), 1–47. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10769304

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7*(3), 308–313. doi:10.1167/12.9.1387

Cave, C. B. (1997). Very long-lasting priming in picture naming. *Psychological Science*, *8*(4), 322–325. doi:10.1111/j.1467-9639.1991.tb00167.x

Chalmers, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press.

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, *40*(4), 343–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3502878

Chelazzi, L., Miller, E., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature, 363*(6427), 345–7.

doi:10.1038/363345a0

Chun, M. M., & Marois, R. (2002). The dark side of visual attention. *Current Opinion in Neurobiology*, *12*(2), 184–189. doi:10.1016/S0959-4388(02)00309-4

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychol Sci*, *22*(9), 1165–1172. doi:10.1177/0956797611419168

Cohen, M. a., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences*, *16*(8), 411–417. doi:10.1016/j.tics.2012.06.013

Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping*, *25*(4), 409–23. doi:10.1002/hbm.20118

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, *27*(3), 183–228. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6992093

Cools, R., & D'Esposito, M. (2011). Inverted-U shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry*, *29*(6), 997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–15. doi:10.1038/nrn755

Corbetta, M., & Shulman, G. L. (2011). *Spatial neglect and attention networks*. *Annual review of neuroscience* (Vol. 34). doi:10.1146/annurev-neuro-061010-113731

Corkin, S. (2002). What's new with the amnesic patient H.M.? *Nature Reviews. Neuroscience*, *3*(2), 153–160. doi:10.1038/nrn726

Courtney, S. M., Petit, L., Maisog, J., Ungerleider, L., & Haxby, J. (1998). An Area Specialized for Spatial Working Memory in Human Frontal Cortex. *Science*, *279*(5355), 1347–1351. doi:10.1126/science.279.5355.1347

Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature*, *386*(6625), 608–11. doi:10.1038/386608a0

Cowan, N. (1995). *Attention and memory: An integrated framework* (Oxford Psy.). Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87–114; discussion 114–185. doi:10.1017/S0140525X01003922

Cowan, N. (2008). *What are the differences between long-term, short-term, and working memory? Progress in brain research* (Vol. 169). Elsevier. doi:10.1016/S0079-6123(07)00020-9

Cowey, A. (2010). The blindsight saga. *Experimental Brain Research*, *200*(1), 3–24. doi:10.1007/s00221-009-1914-2

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, *375*(6527), 121–3. doi:10.1038/375121a0

Crick, F., & Koch, C. (1998). Feature Article Consciousness and Neuroscience. *Cerebral Cortex*, *8*(2), 97–107.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6*(2), 119–26. doi:10.1038/nn0203-119

D'Esposito, M., & Postle, B. R. (1999). The dependence of span and delayed-response performance on prefrontal cortex. *Neuropsychologia*, *37*(11), 1303–15.

D'Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working Memory. *Annual Review of Psychology*, 115–142. doi:10.1146/annurev-psych-010814-015031

Danker, J. F., Tompary, A., & Davachi, L. (2016). Trial-by-Trial Hippocampal Encoding Activation Predicts the Fidelity of Cortical Reinstatement During Subsequent Retrieval. *Cerebral Cortex*, 1–10. doi:10.1093/cercor/bhw146

Dannlowski, U., Ohrmann, P., Bauer, J., Kugel, H., Arolt, V., Heindel, W., & Suslow, T. (2007). Amygdala reactivity predicts automatic negative evaluations for facial emotions. *Psychiatry Research - Neuroimaging*, *154*(1), 13–20. doi:10.1016/j.pscychresns.2006.05.005

Debner, J. a, & Jacoby, L. L. (1994). Unconscious perception: attention, awareness, and control. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *20*(2), 304–17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8151275

Degonda, N., Mondadori, C. R. a, Bosshardt, S., Schmidt, C. F., Boesiger, P., Nitsch, R. M., ... Henke, K. (2005). Implicit associative learning engages the hippocampus and interacts with explicit associative learning. *Neuron*, *46*(3), 505–20. doi:10.1016/j.neuron.2005.02.030

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 200–27. doi:10.1016/j.neuron.2011.03.018

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–11. doi:10.1016/j.tics.2006.03.007

Dehaene, S., Charles, L., King, J.-R., & Marti, S. (2014). Toward a computational theory of conscious processing. *Current Opinion in Neurobiology*, *25*(1947), 76–84. doi:10.1016/j.conb.2013.12.005

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, *79*(1-2), 1–37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11164022

Dehaene, S., Naccache, L., Clec'H, G., Koechlin, E., Dehaene-Lambertz, G., Moortele, P.-F., & Le Bihan, D. (1994). Semantic Priming in a Single-Word Shadowing Task. *The American Journal of Psychology*, *107*(2), 245. doi:10.2307/1423039

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. doi:10.1146/annurev.ne.18.030195.001205

Desmond, J. E., Chen, S. H. A., & Shieh, P. B. (2005). Cerebellar transcranial magnetic stimulation impairs verbal working memory. *Annals of Neurology*, *58*(4), 553–560. doi:10.1002/ana.20604

Desmond, J. E., & Fiez, J. A. (1998). Neuroimaging studies of the cerebellum: Language, learning and memory. *Trends in Cognitive Sciences*, *2*(9), 355–362. doi:10.1016/S1364-6613(98)01211-X

Draine, S. C., & Greenwald, a G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology. General*, *127*(3), 286–303. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9742717

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology (London, England : 1953)*, *92 Part 1*, 53–78. doi:10.1348/000712601162103

Druzgal, T. J., & D'Esposito, M. (2001). Activity in fusiform face area modulated as a function of working memory load. *Cognitive Brain Research*, *10*(3), 355–364. doi:10.1016/S0926-6410(00)00056-2

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, *369*(6478), 313–315. doi:10.1038/369313a0

Duss, S. B., Oggier, S., Reber, T. P., & Henke, K. (2011). Formation of semantic associations between subliminally presented face-word pairs. *Consciousness and Cognition*, *20*(3), 928–35. doi:10.1016/j.concog.2011.03.018

Duss, S. B., Reber, T. P., Hänggi, J., Schwab, S., Wiest, R., Müri, R. M., … Henke, K. (2014). Unconscious relational encoding depends on hippocampus. *Brain : A Journal of Neurology*, *137*(12), 3355–3370. doi:10.1093/brain/awu270

Dutta, A., Shah, K., Silvanto, J., & Soto, D. (2014). Neural basis of non-conscious visual working memory. *NeuroImage*, *91*, 336–343. doi:10.1016/j.neuroimage.2014.01.016

Ebbinghaus, H. (1885). Memory: A Contribution to Experimental Psychology. *Memory: A Contribution to Experimental Psychology*, *15*, 1–7. doi:52, 281-302

Ekman, U., Eriksson, J., Forsgren, L., Mo, S. J., Riklund, K., & Nyberg, L. (2012). Functional brain activity and presynaptic dopamine uptake in patients with Parkinson's disease and mild cognitive impairment: a cross-sectional study. *Lancet Neurology*, *11*(8), 679–87. doi:10.1016/S1474-4422(12)70138-2

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. a. (2000). Remembering episodes: a selective role for the hippocampus during retrieval. *Nature Neuroscience*, *3*(11), 1149–1152. doi:10.1038/80671

Elliott, J. C., Baird, B., & Giesbrecht, B. (2016). Consciousness isn't all-or-none: Evidence for partial awareness during the attentional blink. *Consciousness and Cognition*, *40*, 79–85. doi:10.1016/j.concog.2015.12.003

Emrich, S. M., Riggall, A. C., Larocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(15), 6516–23. doi:10.1523/JNEUROSCI.5732-12.2013

Eriksen, C. W. (1960). Discrimination and Learning Without Awareness: A Methodological Survey and Evaluation. *The Psychological Review*, *67*(5), 279–300.

Eriksson, J., Larsson, A., & Nyberg, L. (2008). Item-specific training reduces prefrontal cortical involvement in perceptual awareness. *Journal of Cognitive Neuroscience*, *20*(10), 1777–87. doi:10.1162/jocn.2008.20064

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive Architecture of Working Memory. *Neuron*, *88*(1), 33–46. doi:10.1016/j.neuron.2015.09.020

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: is it really attention free? *Journal of Experimental Psychology. Human Perception and Performance*, *31*(6), 1476–1492. doi:10.1037/0096-1523.31.6.1476

Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, *8*(10), 1380–5. doi:10.1038/nn1537

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision*, *7*(1), 10. doi:10.1167/7.1.10

Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., & Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. *Proceedings of the National Academy of Sciences*, 1–6. doi:10.1073/pnas.1106439108

Ffytche, D. H., Howard, R. J., Brammer, M. J., David, a, Woodruff, P., & Williams, S. (1998). The anatomy of conscious vision: an fMRI study of visual hallucinations. *Nature Neuroscience*, *1*(8), 738–42. doi:10.1038/3738

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: Competitive content maps for attention and memory. *Trends in Cognitive Sciences*, *17*(3), 134–141. doi:10.1016/j.tics.2013.01.010

Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhauser, W. (2014). Binocular Rivalry: Frontal Activity Relates to Introspection and Action But Not to Perception. *Journal of Neuroscience*, *34*(5), 1738–1747. doi:10.1523/JNEUROSCI.4403-13.2014

Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*(2), 177–82. doi:10.1016/j.conb.2010.03.005

Fukuda, K., Vogel, E., & Mayr, U. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. doi:10.3758/17.5.673.Quantity

Fuller, S., & Carrasco, M. (2006). Exogenous attention and color perception: Performance and appearance of saturation and hue. *Vision Research*, *46*(23), 4032–4047. doi:10.1016/j.visres.2006.07.014

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*(2), 331–349.

Fuster, J., & Alexander, G. (1971). Neuron activity related to short-term memory. *Science*, *173*(3997), 652–654. doi:10.1055/s-2006-956756

Fuster, J. M. (1958). Effects of stimulation of brain stem on tachistoscopic perception. *Science*, *127*, 150.

Fuster, J. M. (1995). *Memory in the Cerebral Cortex - An Empirical Approach to Neural Networks in the Human and Nonhuman Primate.* Cambridge: MA: MIT Press.

Fuster, J. M. (1997). Network memory. *Trends in Neurosciences*, *20*(10), 451–9.

Fuster, J. M. (2003a). *Cortex and Mind: Unifying Cognition*. New York: Oxford University Press.

Fuster, J. M. (2003b). More than working memory rides on long-term memory. *Behavioral and Brain Sciences*, *26*(06), 737.

doi:10.1017/s0140525x03300160

Fuster, J. M. (2004). Upper processing stages of the perception-action cycle. *Trends in Cognitive Sciences*, *8*(4), 143–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15551481

Fuster, J. M. (2009). Cortex and memory: emergence of a new paradigm. *Journal of Cognitive Neuroscience*, *21*(11), 2047–72. doi:10.1162/jocn.2009.21280

Fuster, J. M. (2015). *The Prefrontal Cortex* (5th Editio.). London: Academic Press.

Fuster, J. M., Bauer, R. H., & Jervey, J. P. (1985). Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Research*, *330*(2), 299–307. doi:10.1016/0006-8993(85)90689-4

Fuster, J. M., & Bressler, S. L. (2015). Past makes future: role of pFC in prediction. *Journal of Cognitive Neuroscience*, *27*(4), 639–54. doi:10.1162/jocn_a_00746

Fuster, J. M., & Jervey, J. P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science (New York, N.Y.)*, *212*(4497), 952–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7233192

Gaillard, R., Cohen, L., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., ... Naccache, L. (2007). Subliminal words durably affect neuronal activity. *Neuroreport*, *18*(15), 1527–31. doi:10.1097/WNR.0b013e3282f0b6cd

Gelbard-Sagiv, H., Faivre, N., Mudrik, L., & Koch, C. (2016). Low-level awareness accompanies “unconscious” high-level processing during continuous flash suppression. *Journal of Vision*, *16*(1), 3. doi:10.1167/16.1.3

Gobell, J., & Carrasco, M. (2005). Attention alters the appearance of spatial frequency and gap size. *Psychological Science*, *16*(8), 644–651. doi:10.1111/j.1467-9280.2005.01588.x

Goense, J. B. M., & Logothetis, N. K. (2008). Neurophysiology of the BOLD fMRI Signal in Awake Monkeys. *Current Biology*, *18*(9), 631–640. doi:10.1016/j.cub.2008.03.054

Goldman-Rakic, P. . (1995). Cellular basis of working memory. *Neuron*, *14*(3), 477–485. doi:10.1016/0896-6273(95)90304-6

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *11*(3), 501–518. doi:10.1037/0278-7393.11.3.501

Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *10*(1), 164–178. doi:10.1037/0278-7393.10.1.164

Greene, A. J., Gross, W. L., Elsinger, C. L., & Rao, S. M. (2007). Hippocampal differentiation without recognition: an fMRI analysis of the contextual cueing task. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *14*(8), 548–53. doi:10.1101/lm.609807

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three Cognitive Markers of Unconscious Semantic Activation. *Science*, *273*(20 September), 1699–1702. doi:10.1007/BF00309217

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. doi:10.1016/j.tics.2005.11.006

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–77. doi:10.1146/annurev.neuro.27.070203.144220

Habib, R., Nyberg, L., & Tulving, E. (2003). Hemispheric asymmetries of memory: the HERA model revisited. *Trends in Cognitive Sciences*, *7*(6), 241–245. doi:10.1016/S1364-6613(03)00110-4

Harrison, S. a, & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632–635. doi:10.1038/nature07832

Hassin, R. R., Bargh, J. a, Engell, A. D., & McCulloch, K. C. (2009). Implicit working memory. *Consciousness and Cognition*, *18*(3), 665–78. doi:10.1016/j.concog.2009.04.003

Haxby, J. V, Connolly, A. C., & Guntupalli, J. S. (2014). Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. *Annual Review of Neuroscience*, (June), 435–456. doi:10.1146/annurev-neuro-062012-170325

Hayek, F. A. (1952). *The Sensory Order: An Inquiry Into the Foundations of Theoretical Psychology*. University of Chicago Press. Retrieved from http://books.google.ca/books/about/The_Sensory_Order.html?id=UFazm1Xy_j4C&pgis=1

Hayman, C. A. G., Macdonald, C. A., & Tulving, E. (1993). The role of repetition and associative interference in new semantic learning in amnesia: A case experiment. *Journal of Cognitive Neuroscience*, *5*(4), 375–389. doi:10.1162/jocn.1993.5.4.375</p>

Haynes, J.-D. (2015). A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron*, *87*(2), 257–270. doi:10.1016/j.neuron.2015.05.025

Haynes, J.-D., Deichmann, R., & Rees, G. (2005). Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature*, *438*(7067), 496–9. doi:10.1038/nature04169

Haynes, J.-D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, *8*(5), 686–91. doi:10.1038/nn1445

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews. Neuroscience*, *7*(7), 523–34. doi:10.1038/nrn1931

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*(6598), 334–7. doi:10.1038/383334a0

He, S., & MacLeod, D. I. (2001). Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature*, *411*(6836), 473–6. doi:10.1038/35078072

Hebb, D. H. (1949). *The Organisation of Behavior*. Wiley.

Heinzel, a, Hautzel, H., Poeppel, T. D., Boers, F., Beu, M., & Mueller, H.-W. (2008). Neural correlates of subliminal and supraliminal letter processing--an event-related fMRI study. *Consciousness and Cognition*, *17*(3), 685–99. doi:10.1016/j.concog.2008.01.008

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews. Neuroscience*, *11*(7), 523–32. doi:10.1038/nrn2850

Henke, K., Mondadori, C. R. ., Treyer, V., Nitsch, R. M., Buck, A., & Hock, C. (2003). Nonconscious formation and reactivation of semantic associations by way of the medial temporal lobe. *Neuropsychologia*, *41*(8), 863–876. doi:10.1016/S0028-3932(03)00035-6

Henke, K., Treyer, V., Nagy, E. T., Kneifel, S., Dürsteler, M., Nitsch, R. M., & Buck, A. (2003). Active hippocampus during nonconscious memories. *Consciousness and Cognition*, *12*(1), 31–48. doi:10.1016/S1053-8100(02)00006-5

Henson, R. (2000). Neuroimaging Evidence for Dissociable Forms of Repetition Priming. *Science*, *287*(5456), 1269–1272. doi:10.1126/science.287.5456.1269

Henson, R. N. . (2003). Neuroimaging studies of priming. *Progress in Neurobiology*, *70*(1), 53–81. doi:10.1016/S0301-0082(03)00086-8

Henson, R. N., Eckstein, D., Waszak, F., Frings, C., & Horner, A. J. (2014). Stimulus–response bindings in priming. *Trends in Cognitive Sciences*, 1–9. doi:10.1016/j.tics.2014.03.004

Hergenhan, B. R. (2001). *An introduction to the history of psychology* (4th ed.). Wadsworth/Thomson Learning.

Hohwy, J. (2009). The neural correlates of consciousness: new experimental approaches needed? *Consciousness and Cognition*, *18*(2), 428–38. doi:10.1016/j.concog.2009.02.006

Holender, D. (1986). Semantic Activation Without Conscious Identification in Dichotic Listening, Parafoveal Vision, and Visual Masking: A Survey and Appraisal. *Behavioral and Brain Sciences*, *9*(1), 1–23.

Horton, J. C., & Hoyt, W. F. (1991). Quadrantic visual field defects. A hallmark of lesions in extrastriate (V2/V3) cortex. *Brain : A Journal of Neurology*, *114 (Pt 4*(4), 1703–18. doi:10.1093/brain/114.4.1703

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). Functional magnetic resonance imaging. *New England Journal of Medicine*, 6–13. doi:978-0-87893-286-3

Ingvar, D. H. (1985). “Memory of the future”: an essay on the temporal organization of conscious awareness. *Human Neurobiology*, *4*(3), 127–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3905726

James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory-motor brain network. *Neuropsychologia*, *44*(14), 2937–49. doi:10.1016/j.neuropsychologia.2006.06.026

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(45), 17048–52. doi:10.1073/pnas.0605678103

Jiang, Y., Zhou, K., & He, S. (2007). Human visual cortex responds to invisible chromatic flicker. *Nature Neuroscience*, *10*(5), 657–62. doi:10.1038/nn1879

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. a, Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, *59*, 193–224. doi:10.1146/annurev.psych.59.103006.093615

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a “preattentive” feature search task. *Nature*, *387*(6635), 805–807. doi:10.1038/42940

Kanai, R., Tsuchiya, N., & Verstraten, F. a J. (2006). The scope and limits of top-down attention in unconscious visual processing. *Current Biology : CB*, *16*(23), 2332–6. doi:10.1016/j.cub.2006.10.001

Karremans, J. C., Stroebe, W., & Claus, J. (2006). Beyond Vicary’s fantasies: The impact of subliminal priming and brand choice. *Journal of Experimental Social Psychology*, *42*(6), 792–798. doi:10.1016/j.jesp.2005.12.002

Kelly, R. M., & Strick, P. L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(23), 8432–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12968006

Kentridge, R. W., Heywood, C. a, & Weiskrantz, L. (2004). Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia*, *42*(6), 831–5. doi:10.1016/j.neuropsychologia.2003.11.001

Kentridge, R. W., Nijboer, T. C. W., & Heywood, C. A. (2008). Attended but unseen: Visual attention is not sufficient for visual awareness. *Neuropsychologia*, *46*(3), 864–869. doi:10.1016/j.neuropsychologia.2007.11.036

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: rendering the visible “invisible”. *Trends in Cognitive Sciences*, *9*(8), 381–8. doi:10.1016/j.tics.2005.06.012

King, J.-R., Pescetelli, N., & Dehaene, S. (2016). Selective maintenance mechanisms of seen and unseen sensory features in the human brain. *BioRxiv.org*, 1–30. doi:http://dx.doi.org/10.1101/040030

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Roberts & Company Publishers.

Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307–321. doi:10.1038/nrn.2016.22

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. doi:10.1016/j.tics.2006.10.012

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, *11*(6), 229–35. doi:10.1016/j.tics.2007.04.005

Koenigs, M., Barbey, A. K., Postle, B. R., & Grafman, J. (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(47), 14980–6. doi:10.1523/JNEUROSCI.3706-09.2009

Koivisto, M., Mäntylä, T., & Silvanto, J. (2010). The role of early visual cortex (V1/V2) in conscious and unconscious visual perception. *NeuroImage*, *51*(2), 828–34. doi:10.1016/j.neuroimage.2010.02.042

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 857–75. doi:10.1098/rstb.2007.2093

Koutstaal, W., Wagner, A. D., Rotte, M., Maril, A., Buckner, R. L., & Schacter, D. L. (2001). Perceptual specificity in visual object priming: functional magnetic resonance imaging evidence for a laterality difference in fusiform cortex. *English*, *39*, 184 – 199.

Kravitz, D. J., Kriegeskorte, N., & Baker, C. I. (2010). High-level visual object representations are constrained by position. *Cerebral Cortex*, *20*(December), 2916–2925. doi:10.1093/cercor/bhq042

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews. Neuroscience*, *12*(4), 217–30. doi:10.1038/nrn3008

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2012). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, *17*(1), 26–49. doi:10.1016/j.tics.2012.10.011

Kravitz, D. J., Vinson, L. D., & Baker, C. I. (2008). How position dependent is visual object recognition? *Trends in Cognitive Sciences*, *12*(February), 114–122. doi:10.1016/j.tics.2007.12.006

Kreiman, G., Koch, C., & Fried, I. (2000). Imagery neurons in the human brain. *Nature*, *408*(6810), 357–61. doi:10.1038/35042575

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., ... Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(12), 5675–9. doi:10.1073/pnas.89.12.5675

Lamme, V. a F. (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Networks : The Official Journal of the International Neural Network Society*, *17*(5-6), 861–72. doi:10.1016/j.neunet.2004.02.005

Lamme, V. a F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*(11), 494–501. doi:10.1016/j.tics.2006.09.001

Lamme, V. a, & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11074267

Landau, A. N., & Bentin, S. (2008). Attentional and perceptual factors affecting the attentional blink for faces and objects. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(4), 818–830. doi:10.1037/0096-1523.34.4.818

LaRocque, J. J., Lewis-Peacock, J. a, Drysdale, A. T., Oberauer, K., & Postle, B. R. (2013). Decoding attended information in short-term memory: an

EEG study. *Journal of Cognitive Neuroscience*, *25*, 127–42. doi:10.1162/jocn_a_00305

Larocque, J. J., Lewis-Peacock, J. a, & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Frontiers in Human Neuroscience*, *8*(January), 5. doi:10.3389/fnhum.2014.00005

Lau, H. C. (2008). Are we studying consciousness yet? In L. Weiskrantz & M. Davies (Eds.), *Frontiers of Consciousness: Chichele Lectures* (pp. 245–258). Oxford: Oxford University Press.

Lau, H. C., & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(21), 5805–11. doi:10.1523/JNEUROSCI.4335-06.2007

Lau, H., & Rosenthal, D. (2011a). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, *15*(8), 365–373. doi:10.1016/j.tics.2011.05.009

Lau, H., & Rosenthal, D. (2011b). The higher-order view does not require consciously self-directed introspection: Response to Malach. *Trends in Cognitive Sciences*, *15*(11), 508–509. doi:10.1016/j.tics.2011.09.005

Laureys, S. (2005). The neural correlate of (un)awareness: Lessons from the vegetative state. *Trends in Cognitive Sciences*, *9*(12), 556–559. doi:10.1016/j.tics.2005.10.010

Lauterbur, P. C. (1973). Image formation by induced local interactions. Examples employing nuclear magnetic resonance. *Nature (London, United Kingdom)*, *242*, 190–191. doi:10.1038/242190a0

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451–468. doi:10.1037/0096-1523.21.3.451

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. doi:10.1037/0096-3445.133.3.339

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197. doi:10.3758/BF03213897

Lee, S.-H., Blake, R., & Heeger, D. J. (2005). Traveling waves of activity in primary visual cortex during binocular rivalry. *Nature Neuroscience*, *8*(1), 22–3. doi:10.1038/nn1365

Lee, S.-H., Blake, R., & Heeger, D. J. (2007). Hierarchy of cortical responses underlying binocular rivalry. *Nature Neuroscience*, *10*(8), 1048–54. doi:10.1038/nn1939

Leopold, D. a. (2012). Primary visual cortex: awareness and blindsight. *Annual Review of Neuroscience*, *35*, 91–109. doi:10.1146/annurev-neuro-062111-150356

Lewis-Peacock, J. a, Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*(1), 61–79. doi:10.1162/jocn_a_00140

Lewis-Peacock, J. a, & Postle, B. R. (2008). Temporary activation of long-term memory supports working memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(35), 8765–8771. doi:10.1523/JNEUROSCI.1953-08.2008

Levy, B. R., Pilver, C., Chung, P. H., & Slade, M. D. (2014). Subliminal Strengthening: Improving Older Individuals' Physical Function Over Time With an Implicit-Age-Stereotype Intervention. *Psychological Science*, *25*(12), 2127–2135. doi:10.1177/0956797614551970

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9596–601. doi:10.1073/pnas.092277599

Libby, L. a., Hannula, D. E., & Ranganath, C. (2014). Medial Temporal Lobe Coding of Item and Spatial Information during Relational Binding in Working Memory. *Journal of Neuroscience*, *34*(43), 14233–14242. doi:10.1523/JNEUROSCI.0655-14.2014

Lindquist, M. A. (2008). The Statistical Analysis of fMRI Data. *Statistical Science*, *23*(4), 439–464. doi:10.1214/09-STS282

Liu, T., Abrams, J., & Carrasco, M. (2009). Voluntary attention enhances contrast appearance. *Psychological Science*, *20*(3), 354–362. doi:10.1111/j.1467-9280.2009.02300.x

Logan, G. D., & Crump, M. J. C. (2010). Cognitive illusions of authorship reveal hierarchical error detection in skilled typists. *Science (New York, N.Y.)*, *330*(6004), 683–6. doi:10.1126/science.1190483

Logothetis, N. K. (1998). Single units and conscious vision. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *353*(1377), 1801–18. doi:10.1098/rstb.1998.0333

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–78. doi:10.1038/nature06976

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, a. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–7. doi:10.1038/35084005

Logothetis, N. K., & Wandell, B. a. (2004). Interpreting the BOLD signal. *Annual Review of Physiology*, *66*, 735–69. doi:10.1146/annurev.physiol.66.082602.092845

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *383*(6601), 616–8. doi:10.1038/383616a0

Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural Correlates of Perceptual Rivalry in the Human Brain. *Science*, *280*(5371), 1930–1934. doi:10.1126/science.280.5371.1930

Lundqvist, M., Herman, P., & Lansner, A. (2011). Theta and gamma power increases and alpha/beta power decreases with memory load in an attractor network model. *J Cogn Neurosci*, *23*(10), 3008–3020.

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., & Miller, E. K. (2015). Gamma and Beta Bursts Underlie Working Memory. *Neuron*, *90*(1), 152–164. doi:10.1016/j.neuron.2016.02.028

Mack, A., & Clarke, J. (2011). Gist perception requires attention. *Visual Cognition*, *22*(9), 1165–1172. doi:10.1177/0956797611419168

Mack, A., & Rock, I. (1998). *Inattentional Blindness*. MIT Press.

Makovski, T. (2012). Are multiple visual short-term memory storages necessary to explain the retro-cue effect? *Psychonomic Bulletin & Review*, *19*(3), 470–6. doi:10.3758/s13423-012-0235-9

Mansfield, P. (1977). Multi-planar image formation using. *Journal of Physics C: Solid State Physics*, *10*, 55–58.

Marois, R., Chun, M. M., & Gore, J. C. (2000). Neural correlates of the attentional blink. *Neuron*, *28*(1), 299–308. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15652305

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, *9*(6), 296–305. doi:10.1016/j.tics.2005.04.010

Marois, R., Yi, D.-J., & Chun, M. M. (2004). The neural fate of consciously perceived and missed events in the attentional blink. *Neuron*, *41*(3), 465–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14766184

Marsolek, C. J. (2008). What antipriming reveals about priming. *Trends in Cognitive Sciences*, *12*(5), 176–81. doi:10.1016/j.tics.2008.02.005

Marsolek, C. J., Deason, R. G., Ketz, N. a, Ramanathan, P., Bernat, E. M., Steele, V. R., ... Schnyer, D. M. (2010). Identifying objects impairs knowledge of other objects: a relearning explanation for the neural repetition effect. *NeuroImage*, *49*(2), 1919–32. doi:10.1016/j.neuroimage.2009.08.063

Martens, S., & Wyble, B. (2010). The attentional blink: past, present, and future of a blind spot in perceptual awareness. *Neuroscience and Biobehavioral Reviews*, *34*(6), 947–57. doi:10.1016/j.neubiorev.2009.12.005

Marzouki, Y., Grainger, J., & Theeuwes, J. (2007). Exogenous spatial cueing modulates subliminal masked priming. *Acta Psychologica*, *126*(1), 34–45. doi:10.1016/j.actpsy.2006.11.002

Mattler, U. (2005). Inhibition and decay of motor and nonmotor priming. *Perception & Psychophysics*, *67*(2), 285–300. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15971692

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 817–835.

McElree, B. (2006). Accessing Recent Events. *Psychology of Learning and Motivation - Advances in Research and Theory*, *46*(06), 155–200. doi:10.1016/S0079-7421(06)46005-9

Melcher, D., Papathomas, T. V., & Vidnyánszky, Z. (2005). Implicit attentional selection of bound visual features. *Neuron*, *46*(5), 723–729.

doi:10.1016/j.neuron.2005.04.023

Merikle, P. M., & Joordens, S. (1997). Parallels between perception without attention and perception without awareness. *Consciousness and Cognition*, *6*(2-3), 219–36. doi:10.1006/ccog.1997.0310

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: perspectives from cognitive psychology. *Cognition*, *79*(1-2), 115–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11164025

Miller, E. K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science (New York, N.Y.)*, *263*(5146), 520–2. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8290960

Mitchell, D. B. (2006). Nonconscious priming after 17 years: invulnerable implicit memory? *Psychological Science*, *17*(11), 925–9. doi:10.1111/j.1467-9280.2006.01805.x

Miyashita, Y., & Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*, *331*(7), 68–70.

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science (New York, N.Y.)*, *319*(5869), 1543–1546. doi:10.1126/science.1150769

Moran, J., & Desimone, R. (1985). Selective Attention Gates Visual Processing in the Extrastriate Cortex. *Science*, *229*, 782–784.

Moruzzi, G., & Horace, M. W. (1949). Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, *1*(1), 455–473.

Moutoussis, K., Keliris, G., Kourtzi, Z., & Logothetis, N. (2005). A binocular rivalry study of motion perception in the human brain. *Vision Research*, *45*(17), 2231–43. doi:10.1016/j.visres.2005.02.007

Moutoussis, K., & Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9527–32. doi:10.1073/pnas.142305699

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage*, *63*(1), 240–4.

doi:10.1016/j.neuroimage.2012.06.043

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*(2), 217–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9142752

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, *83*(4), 435. doi:10.2307/2183914

Naghavi, H. R., & Nyberg, L. (2005). Common fronto-parietal activity in attention, memory, and consciousness: shared demands on integration? *Consciousness and Cognition*, *14*(2), 390–425. doi:10.1016/j.concog.2004.10.003

Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2013). A meta-analysis of executive components of working memory. *Cerebral Cortex*, *23*(2), 264–82. doi:10.1093/cercor/bhs007

Nee, D. E., & Jonides, J. (2011). Dissociable contributions of prefrontal cortex and the hippocampus to short-term memory: Evidence for a 3-state model of memory. *NeuroImage*, *54*(2), 1540–1548. doi:10.1016/j.neuroimage.2010.09.002

Nee, D. E., & Jonides, J. (2013). Neural Evidence for a 3-State Model of Visual Short-Term Memory. *NeuroImage*. doi:10.1016/j.neuroimage.2013.02.019

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Crofts.

Nelissen, N., Stokes, M., Nobre, A. C., & Rushworth, M. F. S. (2013). Frontal and parietal cortical interactions with distributed visual representations during selective attention and action selection. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(42), 16443–58. doi:10.1523/JNEUROSCI.2625-13.2013

Nobre, A. C., Coull, J. T., Maquet, P., & Frith, C. D. (2004). Orienting Attention to Locations in Perceptual Versus Mental Representations. *Journal of Cognitive Neuroscience*, *16*(3), 363–373.

Norman, K. a, Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–30. doi:10.1016/j.tics.2006.07.005

Nyberg, L., Habib, R., Mcintosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences*, *97*(20), 11120 – 11124.

Nyberg, L., McIntosh, A., Houle, S., Nilsson, L.-G., & Tulving, E. (1996). Activation of medial temporal structures during episodic memory retrieval. *Nature*, *380*, 715 – 717.

Nyberg, L., McIntosh, A. R., Cabeza, R., Habib, R., Houle, S., & Tulving, E. (1996). General and specific brain regions involved in encoding and retrieval of events: what, where, and when. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(20), 11280–5. doi:10.1073/pnas.93.20.11280

Nyberg, L., Persson, J., Habib, R., Tulving, E., Mcintosh, A. R., & Cabeza, R. (1997). Large Scale Neurocognitive Networks Underlying Episodic Memory, 163–173.

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. *Journal of Cognitive Neuroscience*, *12*(6), 1013–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11177421

O'Kane, G., Kensinger, E. a, & Corkin, S. (2004). Evidence for semantic learning in profound amnesia: an investigation with patient H.M. *Hippocampus*, *14*(4), 417–25. doi:10.1002/hipo.20005

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A Biologically-Based Computational Model of Working Memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control.*

Oberauer, K. (2001). Removing irrelevant information from working memory: a cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*(4), 948–957. doi:10.1037/0278-7393.27.4.948

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *28*(3), 411–421. doi:10.1037/0278-7393.28.3.411

Oberauer, K. (2005). Control of the contents of working memory--a comparison of two paradigms and two age groups. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(4), 714–728. doi:10.1037/0278-7393.31.4.714

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(24), 9868–72. doi:10.1073/pnas.87.24.9868

Ogawa, S., Tank, D., Menon, Ra., Ellermann, J., Kim, S.-G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(July), 5951–5955.

Owen, a M., Morris, R. G., Sahakian, J. L., Polkey, C. E., Robbins, T. W., & Sahakian, B. J. (1996). Double dissociations of memory and executive\nfunctions in working memory task following frontal lobe\nexcision, temporal lobe excisions or amygdala-hippocampectomy in man. *Brain*, *119*(5), 1597–1615.

Oztekin, I., Davachi, L., & McElree, B. (2010). Are representations in working memory distinct from representations in long-term memory? Neural evidence in support of a single store. *Psychological Science : A Journal of the American Psychological Society / APS*, *21*(8), 1123–1133. doi:10.1177/0956797610376651

Oztekin, I., McElree, B., Staresina, B. P., & Davachi, L. (2009). Working memory retrieval: contributions of the left prefrontal cortex, the left posterior parietal cortex, and the hippocampus. *Journal of Cognitive Neuroscience*, *21*(3), 581–593. doi:10.1162/jocn.2008.21016

Palmer, S. (1999). *Vision Science: Photons to Phenomenology*. The MIT Press.

Pan, Y., Lin, B., Zhao, Y., & Soto, D. (2013). Working memory biasing of visual perception without awareness. *Attention, Perception & Psychophysics*, *76*(7), 2051–2061. doi:10.3758/s13414-013-0566-2

Park, J., Hebrank, A., Polk, T. a, & Park, D. C. (2012). Neural dissociation of number from letter recognition and its relationship to parietal numerical processing. *Journal of Cognitive Neuroscience*, *24*(1), 39–50. doi:10.1162/jocn_a_00085

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature*, *362*(6418), 342–5. doi:10.1038/362342a0

Pavone, E. F., Marzi, C. A., & Girelli, M. (2009). Does subliminal visual perception have an error-monitoring system? *The European Journal of*

Neuroscience, *30*(7), 1424–31. doi:10.1111/j.1460-9568.2009.06908.x

Perry, E., Walker, M., Grace, J., & Perry, R. (1999). Acetylcholine in mind: a neurotransmitter correlate of consciousness? *Trends in Neurosciences*, *22*(6), 273–80. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10354606

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, *10*(2), 257–61. doi:10.1038/nn1840

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science (New York, N.Y.)*, *316*(5826), 904–6. doi:10.1126/science.1140459

Pessoa, L., Gutierrez, E., Bandettini, P., & Ungerleider, L. (2002). Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron*, *35*(5), 975–87. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12372290

Peters, J. C., Roelfsema, P. R., & Goebel, R. (2012). Task-relevant and accessory items in working memory have opposite effects on activity in extrastriate cortex. *J Neurosci*, *32*(47), 17003–17011. doi:10.1523/JNEUROSCI.0591-12.2012

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73–89. doi:10.1146/annurev-neuro-062111-150525

Pierce, C. S., & Jastrow, J. (1884). On small differences in sensation. *Memoirs of the National Academy of Sciences*, (3), 73–83.

Pinto, Y., Sligte, I. G., Shapiro, K. L., & Lamme, V. a F. (2013). Fragile visual short-term memory is an object-based and location-specific store. *Psychonomic Bulletin & Review*, (2011). doi:10.3758/s13423-013-0393-4

Pisella, L., Berberovic, N., & Mattingley, J. B. (2004). Impaired working memory for location but not for colour or shape in visual neglect: a comparison of parietal and non-parietal lesions. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *40*(2), 379–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15156795

Pochon, J. B., Levy, R., Poline, J. B., Crozier, S., Lehéricy, S., Pillon, B., ... Dubois, B. (2001). The role of dorsolateral prefrontal cortex in the preparation of forthcoming actions: an fMRI study. *Cerebral Cortex (New York, N.Y. : 1991)*, *11*(3), 260–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11230097

Poldrack, R. A., Mumford, J. a., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. New York: Cambridge univeristy press.

Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience*, *3*(11), 1153–1159. doi:10.1038/80676

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. doi:10.1146/annurev.ne.13.030190.000325

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. doi:10.1016/j.neuroscience.2005.06.005

Postle, B. R., & D’Esposito, M. (1999). “What”-Then-Where" in visual working memory: an event-related fMRI study. *Journal of Cognitive Neuroscience*, *11*(6), 585–97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10601740

Postle, B. R., Stern, C. E., Rosen, B. R., & Corkin, S. (2000). An fMRI investigation of cortical contributions to spatial and nonspatial visual working memory. *NeuroImage*, *11*(5 Pt 1), 409–23. doi:10.1006/nimg.2000.0570

Pratkanis, A. (1992). The cargo-cult science of subliminal persuasion. *Skeptical Inquirer*, *16*(3).

Pratte, M., & Tong, F. (2014). Spatial specificity of working memory representations in the early visual cortex. *Journal of Vision*, *14*(2014), 1–12. doi:10.1167/14.3.22.doi

Purcell, E., Torrey, H., & Pound, R. (1946). Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Physical Review*, *69*(1-2), 37–38. doi:10.1103/PhysRev.69.37

Radvansky, G. A. (2014). Human memory: chapter one: Overview and history of memory research, 1–21.

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences*, *3*(1), 1–23. doi:10.1023/B:PHEN.0000041900.30172.e8

Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences*, *9*(8), 374–380. doi:10.1016/j.tics.2005.06.009

Ranganath, C., & D'Esposito, M. (2001). Medial temporal lobe activity associated with active maintenance of novel information. *Neuron*, *31*, 865–873. doi:10.1016/S0896-6273(01)00411-1

Raymond, J., Shapiro, K., & Arnell, K. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology. Human Perception and Performance*, *18*(3), 849–860.

Reber, T. P., & Henke, K. (2011). Rapid formation and flexible expression of memories of subliminal word pairs. *Frontiers in Psychology*, *2*(November), 343. doi:10.3389/fpsyg.2011.00343

Reber, T. P., Luechinger, R., Boesiger, P., & Henke, K. (2012). Unconscious Relational Inference Recruits the Hippocampus. *Journal of Neuroscience*, *32*(18), 6138–6148. doi:10.1523/JNEUROSCI.5639-11.2012

Reddy, L., Reddy, L., & Koch, C. (2006). Face identification in the near-absence of focal attention. *Vision Research*, *46*(15), 2336–43. doi:10.1016/j.visres.2006.01.020

Rees, G. (2007). Neural correlates of the contents of visual awareness in humans. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1481), 877–86. doi:10.1098/rstb.2007.2094

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science (New York, N.Y.)*, *278*(5343), 1616–1619. doi:10.1126/science.278.5343.1616

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews. Neuroscience*, *3*(4), 261–70. doi:10.1038/nrn783

Rees, G., Rees, G., Russell, C., & Frith, C. D. (2012). Inattentional Blindness

Versus Inattentional Amnesia for Fixated But Ignored Words, *2504*(1999), 2504–2508. doi:10.1126/science.286.5449.2504

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, *44*(6), 563–75.

Ress, D., & Heeger, D. J. (2003). Neuronal correlates of perception in early visual cortex. *Nature Neuroscience*, *6*(4), 414–20. doi:10.1038/nn1024

Reuss, H., Kiesel, A., & Kunde, W. (2015). Adjustments of response speed and accuracy to unconscious cues. *Cognition*, *134*, 57–62. doi:10.1016/j.cognition.2014.09.005

Reuss, H., Kiesel, A., Kunde, W., & Hommel, B. (2011). Unconscious activation of task sets. *Consciousness and Cognition*, *20*(3), 556–67. doi:10.1016/j.concog.2011.02.014

Revonsuo, A. (2010). *Consciousness: The Science of Subjectivity*. New York: Psychology Press.

Ric, F., & Muller, D. (2012). Unconscious addition: when we unconsciously initiate and follow arithmetic rules. *Journal of Experimental Psychology. General*, *141*(2), 222–6. doi:10.1037/a0024608

Riggall, a. C., & Postle, B. R. (2012). The Relationship between Working Memory Storage and Elevated Activity as Measured with Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *32*(38), 12990–12998. doi:10.1523/JNEUROSCI.1892-12.2012

Rosenbaum, R. S., Köhler, S., Schacter, D. L., Moscovitch, M., Westmacott, R., Black, S. E., … Tulving, E. (2005). The case of K.C.: Contributions of a memory-impaired person to memory theory. *Neuropsychologia*, *43*(7), 989–1021. doi:10.1016/j.neuropsychologia.2004.10.007

Rossell, S. L., Price, C. J., & Nobre, A. C. (2003). The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia*, *41*(5), 550–564. doi:10.1016/S0028-3932(02)00181-1

Rowe, J. B., Sakai, K., Lund, T. E., Ramsøy, T., Christensen, M. S., Baare, W. F. C., … Passingham, R. E. (2007). Is the prefrontal cortex necessary for establishing cognitive sets? *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(48), 13303–10. doi:10.1523/JNEUROSCI.2349-07.2007

Ruzzoli, M., Gori, S., Pavan, A., Pirulli, C., Marzi, C. a, & Miniussi, C. (2011). The neural basis of the Enigma illusion: A transcranial magnetic stimulation study. *Neuropsychologia*, *49*(13), 3648–3655. doi:10.1016/j.neuropsychologia.2011.09.020

Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nature Neuroscience*, *6*(1), 75–81. doi:10.1038/nn987

Sakai, K., Rowe, J. B., & Passingham, R. E. (2002). Active maintenance in prefrontal area 46 creates distractor-resistant memory. *Nature Neuroscience*, *5*(5), 479–84. doi:10.1038/nn846

Samaha, J. (2015). How best to study the function of consciousness? *Frontiers in Psychology*, *6*(May), 1–3. doi:10.3389/fpsyg.2015.00604

Sandberg, K., Del Pin, S. H., Bibby, B. M., & Overgaard, M. (2014). Evidence of weak conscious experiences in the exclusion task. *Frontiers in Psychology*, *5*(September), 1–8. doi:10.3389/fpsyg.2014.01080

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Consciousness and Cognition*, *19*(4), 1069–78. doi:10.1016/j.concog.2009.12.013

Schacter, D. L. (1987). Implicit Memory: History and Current Status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 501–518. doi:10.1037/0278-7393.13.3.501

Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron*, *20*(2), 185–95. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9491981

Schacter, D. L., Dobbins, I. G., & Schnyer, D. M. (2004). Specificity of priming: a cognitive neuroscience perspective. *Nature Reviews. Neuroscience*, *5*, 853–862. doi:10.1038/nrn1534

Schacter, D. L., Wig, G. S., & Stevens, W. D. (2007). Reductions in cortical activity during priming. *Current Opinion in Neurobiology*, *17*, 171–176. doi:10.1016/j.conb.2007.02.001

Scoville, & Milner. (1957). Loss of recent memory after bilateral hippocampal lesions. *Neurology, Neurosurgery, Psychiatry*, *20*(11), 11–21.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards Evoke Learning of Unconsciously Processed Visual Stimuli in Adult Humans. *Neuron*, *61*(5), 700–707. doi:10.1016/j.neuron.2009.01.016

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*(10), 1391–400. doi:10.1038/nn1549

Sergent, C., Wyart, V., Babo-Rebelo, M., Cohen, L., Naccache, L., & Tallon-Baudry, C. (2012). Cueing Attention after the Stimulus Is Gone Can Retrospectively Trigger Conscious Perception. *Current Biology*, 1–6. doi:10.1016/j.cub.2012.11.047

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*(8), 314–21. doi:10.1016/j.tics.2008.04.008

Shafi, M., Zhou, Y., Quintana, J., Chow, C., Fuster, J., & Bodner, M. (2007). Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience*, *146*(3), 1082–1108. doi:10.1016/j.neuroscience.2006.12.072

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: a neuropsychological study. *The Quarterly Journal of Experimental Psychology*, *22*(2), 261–273. doi:10.1080/00335557043000203

Sidis, B. (1898). *The psychology of suggestion*. New York: NY: Appleton.

Silvanto, J. (2014). Is primary visual cortex necessary for visual awareness? *Trends in Neurosciences*, *37*(11), 618–619. doi:10.1016/j.tins.2014.09.006

Silvanto, J., Cowey, A., Lavie, N., & Walsh, V. (2005). Striate cortex (V1) activity gates awareness of motion. *Nature Neuroscience*, *8*(2), 143–4. doi:10.1038/nn1379

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(48), 19614–19619. doi:10.1073/pnas.1211645109

Sligte, I. G., Scholte, H. S., & Lamme, V. a F. (2008). Are there multiple visual short-term memory stores? *PloS One*, *3*(2), e1699. doi:10.1371/journal.pone.0001699

Sligte, I. G., Scholte, H. S., & Lamme, V. a F. (2009). V4 activity predicts the strength of visual short-term memory representations. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(23), 7432–8. doi:10.1523/JNEUROSCI.0784-09.2009

Sligte, I. G., Wokke, M. E., Tesselaar, J. P., Scholte, H. S., & Lamme, V. a F. (2011). Magnetic stimulation of the dorsolateral prefrontal cortex dissociates fragile visual short-term memory from visual working memory. *Neuropsychologia*, *49*(6), 1578–88. doi:10.1016/j.neuropsychologia.2010.12.010

Smith, S. M. (2004). Overview of fMRI analysis. *British Journal of Radiology*, *77*(suppl_2), S167–S175. doi:10.1259/bjr/33553595

Snodgrass, M., & Shevrin, H. (2006). Unconscious inhibition and facilitation at the objective detection threshold: replicable and qualitatively different unconscious perceptual effects. *Cognition*, *101*(1), 43–79. doi:10.1016/j.cognition.2005.06.006

Soto, D., Mäntylä, T., & Silvanto, J. (2011). Working memory without consciousness. *Current Biology*, *21*(22), R912–R913. doi:10.1016/j.cub.2011.09.049

Soto, D., & Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends in Cognitive Sciences*, *18*(10), 520–525. doi:10.1016/j.tics.2014.06.005

Soto, D., & Silvanto, J. (2016). Is conscious awareness needed for all working memory processes ? *Neuroscience of Consciousness*, (February), 1–3. doi:10.1093/nc/niw009

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29. doi:10.1037/h0093759

Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring Latent Visual Working Memory Representations in Human Cortex. *Neuron*, *91*(3), 694–707. doi:10.1016/j.neuron.2016.07.006

Squire, L. R., & Dede, A. J. O. (2015). Conscious and unconscious memory systems. *Cold Spring Harbor Perspectives in Biology*, *7*(3), a021667. doi:10.1101/cshperspect.a021667

Squire, L. R., Kosslyn, S., Zola-Morgan, S., Haist, F., & Musen, G. (1992). Memory and the Hippocampus: A Synthesis From Findings With Rats, Monkeys, and Humans. *Psychological Review*, *99*(2), 195–231.

Squire, L. R., & Zola-Morgan, S. (1988). Memory: brain systems and behavior. *Trends in Neurosciences*, *11*, 170–175. doi:10.1016/0166-2236(88)90144-0

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, 1–8. doi:10.1016/j.tics.2013.12.001

Stein, T., Kaiser, D., & Hesselmann, G. (2016). Can working memory be non-conscious? *Neuroscience of Consciousness*, *2016*(1), niv011. doi:10.1093/nc/niv011

Stoodley, C. J., & Schmahmann, J. D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *NeuroImage*, *44*(2), 489–501. doi:10.1016/j.neuroimage.2008.08.039

Strahan, E. J., Spencer, S. J., & Zanna, M. P. (2002). Subliminal priming and persuasion: Striking while the iron is hot. *Journal of Experimental Social Psychology*, *38*(6), 556–568. doi:10.1016/S0022-1031(02)00502-4

Stroh, M., Shaw, M., & Washburn, M. F. (1908). A Study in Guessing. *The American Journal of Psychology*, *19*(2), 243–245. doi:doi:10.2307/1412764

Tartaglia, E. M., Mongillo, G., & Brunel, N. (2015). On the relationship between persistent delay activity, repetition enhancement and priming. *Frontiers in Psychology*, *5*(January), 1–11. doi:10.3389/fpsyg.2014.01590

Thibault, L., van den Berg, R., Cavanagh, P., & Sergent, C. (2016). Retrospective Attention Gates Discrete Conscious Access to Past Sensory Stimuli. *Plos One*, *11*(2), e0148504. doi:10.1371/journal.pone.0148504

Tong, F. (2003). Primary visual cortex and visual awareness. *Nature Reviews. Neuroscience*, *4*(3), 219–29. doi:10.1038/nrn1055

Tong, F., & Pratte, M. S. (2012). Decoding Patterns of Human Brain Activity. *Annual Review of Psychology, Vol 63*, *63*, 483–509. doi:DOI 10.1146/annurev-psych-120710-100412

Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: an update. *Annals of the New York Academy of Sciences*, *1124*, 239–61. doi:10.1196/annals.1440.004

Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*, 242–248.

Treisman, A., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive PsychologyCognitive*, *12*, 97–136.

Tse, P. U. (2005). Voluntary attention modulates the brightness of overlapping transparent surfaces. *Vision Research*, *45*(9), 1095–1098. doi:10.1016/j.visres.2004.11.001

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. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(47), 17178–83. doi:10.1073/pnas.0508010102

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8*(8), 1096–101. doi:10.1038/nn1500

Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends in Cognitive Sciences*, *19*(12), 757–770. doi:10.1016/j.tics.2015.10.002

Tulving, E. (1972). Episodic and semantic memory. *Organization of Memory*. doi:10.1017/S0140525X00047257

Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon.

Tulving, E. (2002). Episodic Memory: From Mind to Brain.

Tulving, E., Hayman, C. a, & Macdonald, C. a. (1991). Long-lasting perceptual priming and semantic learning in amnesia: a case experiment. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *17*(4), 595–617. doi:10.1037/0278-7393.17.4.595

Tulving, E., & Schacter, D. L. (1990). Priming and Human Memory Systems. *Science*, 301–306.

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26. doi:10.1016/j.cogpsych.2014.01.003

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective & Behavioral Neuroscience*, *3*(4), 255–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15040547

Wager, T. D., Spicer, J., Insler, R., & Smith, E. E. (2013). The neural bases of distracter-resistant working memory. *Cognitive, Affective & Behavioral Neuroscience*. doi:10.3758/s13415-013-0226-y

Walker, S., Stafford, P., & Davis, G. (2008). Ultra-rapid categorization requires visual attention: Scenes with multiple foreground objects. *Journal of Vision*, *8*(4), 21.1–12. doi:10.1167/8.4.21

van Boxtel, J. J. a, Tsuchiya, N., & Koch, C. (2010). Opposing effects of attention and consciousness on afterimages. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(19), 8883–8. doi:10.1073/pnas.0913292107

van Gaal, S., Lamme, V. a F., Fahrenfort, J. J., & Ridderinkhof, K. R. (2011). Dissociable brain mechanisms underlying the conscious and unconscious control of behavior. *Journal of Cognitive Neuroscience*, *23*(1), 91–105. doi:10.1162/jocn.2010.21431

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. a F. (2010). Unconscious activation of the prefrontal no-go network. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(11), 4143–50. doi:10.1523/JNEUROSCI.2992-09.2010

van Turennout, M., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience*, *3*(12), 1329–1334. doi:10.1038/81873

Warrington, E. K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia*, *12*, 419–428.

Watanabe, M., Cheng, K., Murayama, Y., Ueno, K., Asamizuya, T., Tanaka, K., & Logothetis, N. (2011). Attention But Not Awareness Modulates the BOLD Signal in the Human V1 During Binocular Suppression. *Science*, *334*(6057), 829–831. doi:10.1126/science.1203161

Verwijmeren, T., Karremans, J. C., Stroebe, W., & Wigboldus, D. H. J.

(2011). The workings and limits of subliminal advertising: The role of habits. *Journal of Consumer Psychology*, *21*(2), 206–213. doi:10.1016/j.jcps.2010.11.004

Whitwell, R. L., Milner, a D., & Goodale, M. a. (2014). The two visual systems hypothesis : new challenges and insights from visual form agnosic patient DF, *5*(December), 1–8. doi:10.3389/fneur.2014.00255

William, J. (1890). *The Principles of Psychology*. New York: Henry Holt & Company.

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society : JINS*, *17*(5), 766–80. doi:10.1017/S1355617711000683

Wunderlich, K., Schneider, K. a, & Kastner, S. (2005). Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nature Neuroscience*, *8*(11), 1595–602. doi:10.1038/nn1554

Yi, D.-J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. *Nature Neuroscience*, *7*(9), 992–996. doi:10.1038/nn1294

Zedelius, C. M., Veling, H., & Aarts, H. (2011). Boosting or choking--how conscious and unconscious reward processing modulate the active maintenance of goal-relevant information. *Consciousness and Cognition*, *20*(2), 355–62. doi:10.1016/j.concog.2010.05.001

Zeki, S. (2001). Localization and globalization in conscious vision. *Annual Review of Neuroscience*, *24*, 57–86. doi:10.1146/annurev.neuro.24.1.57

Zeki, S. (2008). The disunity of consciousness. *Progress in Brain Research*, *168*(0), 11–8. doi:10.1016/S0079-6123(07)68002-9

Zeki, S., Watson, J. D., & Frackowiak, R. S. (1993). Going beyond the information given: the relation of illusory visual motion to brain activity. *Proceedings. Biological Sciences / The Royal Society*, *252*(1335), 215–22. doi:10.1098/rspb.1993.0068

Zeman, a. (2001). Consciousness. *Brain : A Journal of Neurology*, *124*(Pt 7), 1263–89. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11408323

Zokaei, N., Ning, S., Manohar, S., Feredoes, E., & Husain, M. (2014).

Flexibility of representational states in working memory. *Frontiers in Human Neuroscience*, *8*(November), 1–12. doi:10.3389/fnhum.2014.00853

Zou, J., He, S., & Zhang, P. (2016). Binocular rivalry from invisible patterns. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1604816113

I

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 21 November 2014
doi: 10.3389/fnhum.2014.00938

Maintenance of non-consciously presented information engages the prefrontal cortex

Fredrik Bergström[1,2] * **and Johan Eriksson**[1,2]

[1] *Umeå center for Functional Brain Imaging (UFBI), Umeå, Sweden*
[2] *Department of Integrative Medical Biology, Physiology Section, Umeå University, Umeå, Sweden*

Edited by:
Richard A. P. Roche, Maynooth University, Ireland

Reviewed by:
David Soto, Imperial College London, UK
Lorraine Boran, Dublin City University, Ireland

****Correspondence:***
Fredrik Bergström, Department of Integrative Medical Biology, Physiology Section, Umeå University, 901 87 Umeå, Sweden
e-mail: fredrik.bergstrom@umu.se

Conscious processing is generally seen as required for flexible and willful actions, as well as for tasks that require durable information maintenance. Here we present research that questions the assumption that only consciously perceived information is durable (>500 ms). Using the attentional blink (AB) phenomenon, we rendered otherwise relatively clearly perceived letters non-conscious. In a first experiment we systematically manipulated the delay between stimulus presentation and response, for the purpose of estimating the durability of non-conscious perceptual representations. For items reported not seen, we found that behavioral performance was better than chance across intervals up to 15 s. In a second experiment we used fMRI to investigate the neural correlates underlying the maintenance of non-conscious perceptual representations. Critically, the relatively long delay period demonstrated in experiment 1 enabled isolation of the signal change specifically related to the maintenance period, separate from stimulus presentation and response. We found sustained BOLD signal change in the right mid-lateral prefrontal cortex, orbitofrontal cortex, and crus II of the cerebellum during maintenance of non-consciously perceived information. These findings are consistent with the controversial claim that working-memory mechanisms are involved in the short-term maintenance of non-conscious perceptual representations.

Keywords: non-conscious, durability, attention, conscious experience, perception, working memory

INTRODUCTION

The functional complexity of the human brain enables us to perceive and interact with our environment in a flexible and deliberate manner. However, despite our intuition to the contrary, we only consciously experience a fraction of the accompanying processes. We thus have the capacity to perceive more information than we can consciously experience, with the consequence that some perceived information about the external environment remains non-conscious, as demonstrated by phenomena like masking (Dehaene et al., 2001) and the attentional blink (AB; Luck et al., 1996).

Higher-level cognitive functions, associated with frontal and parietal cortical regions, have traditionally been considered the exclusive product of conscious processes, while non-conscious processes have been considered limited to automatic, lower-level functions (Koch and Crick, 2001). Correspondingly, neuroimaging studies investigating the neural correlates of conscious experiences have often found that activity in the prefrontal and parietal network (PPN) correlates with conscious perception (Rees et al., 2002; Naghavi and Nyberg, 2005). Based on the frequent involvement of the PPN in conscious perception, the Global Neuronal Workspace (GNW) model states that widespread and recurrent prefrontal and parietal activity determines if information is consciously experienced or not. According to the model, long-distance axons of the PPN, together with thalamocortical loops, form a "global workspace" that interconnects many specialized, automatic, and (otherwise) non-conscious processors. Non-conscious information is hypothesized to become conscious once it is globally broadcast via the PPN and thereby available to many brain regions, enabling depth of processing, more flexible use of information, and durable (>500 ms) representations in working memory and long-term memory (Dehaene and Changeux, 2011; Dehaene et al., 2014).The effects of non-conscious perception have been investigated and hotly debated during the past century (see Kouider and Dehaene, 2007, for review). Recent discoveries suggest that several functions previously associated uniquely with conscious processing can take place after non-conscious perception, and in some cases, activate parts of the PPN, e.g., cognitive control (Lau and Passingham, 2007), flexibility and context-specificity (Wokke et al., 2011), monetary motivation (Pessiglione et al., 2007), and error detection (Logan and Crump, 2010). Most recently, durable (up to 5 s) non-conscious perceptual representations have been demonstrated (Soto et al., 2011), challenging the common notion that non-conscious representations are extremely short-lived.

To investigate the durability of non-conscious representations we here used the AB paradigm as a way to manipulate the conscious experience of seeing a particular stimulus. As the AB effect is not consistent across trials (for a given set of parameters, participants will see a target stimulus on some trials and not see it on others), the AB phenomenon is a useful tool for creating conditions with identical experimental parameters, but with differing conscious experiences. Furthermore, the AB is known to enable relatively long non-conscious presentation durations, e.g., 100 ms (Martens and Wyble, 2010), compared with up to 50 ms for masking (Greenwald et al., 1996). The AB paradigm therefore has the potential to elicit relatively strong non-conscious brain activity (Sergent et al., 2005), and possibly more durable representations.

In a first behavioral experiment we manipulated the delay durations between unseen stimuli presentations and responses to estimate the longevity of non-conscious representations, establishing that non-conscious representations can last for up to 15 s. In a second experiment we used fMRI with a similar paradigm to investigate the neural correlates underlying the maintenance of non-conscious representations. Critically, the relatively long interval between stimulus and response enabled a within-trial separation of BOLD signal related to different trial components (stimulus presentation, delay period, and response), similar to the approach used in neuroimaging research on working memory (Curtis and D'Esposito, 2003).

According to the GNW model there should be no working-memory involvement during processing of non-conscious representations. Contrary to this prediction however, recent research has suggested that working memory operations could account for the durable retention of non-conscious representations (Soto et al., 2011; Pan et al., 2013; Soto and Silvanto, 2014). Furthermore, Dutta et al. (2014) have demonstrated BOLD signal increase in PPN during a delayed cue-target orientation discrimination task with non-conscious sample presentations. However, given the sluggishness of the BOLD signal and that the delay period used by Dutta et al. was short (1.5 s), it is unclear if the signal change was related to maintenance or to stimulus and/or response processing. If working-memory mechanisms indeed are responsible for the maintenance of non-conscious representations there should be sustained BOLD signal change in brain regions characteristically involved in working memory during the delay period, specifically, frontal and parietal cortex related to executive processes and temporal integration of previously attained perceptual knowledge and its prospective use (Cabeza and Nyberg, 2000; Wager and Smith, 2003; Fuster, 2009; Sreenivasan et al., 2014).

MATERIALS AND METHODS

PARTICIPANTS

For experiment 1 (behavioral experiment) 24 participants were recruited from the Umeå University campus area. All participants had normal or corrected-to-normal vision, gave written informed consent, and were paid for participation. Participants were excluded if they failed to comply with instructions (two participants for systematically pressing the same response instead of guessing), or if they had significantly different reported perceptual awareness ratings of target stimulus as a function of time (one participant). Twenty one participants (18–39 year age range, M = 24 years, 13 female) were thus included in the statistical analyses.

For experiment 2 (fMRI experiment) 27 participants were recruited from the Umeå University campus area. All participants were right handed and had normal or corrected-to-normal vision, gave written informed consent, and were paid for participation. The experiment was approved by the ethics committee at the University Hospital of Northern Sweden. Participants were excluded if they failed to comply with instructions (one participant for systematically pressing the same response instead of guessing), or if they had significantly different reported perceptual awareness ratings as a function of time (no participant excluded). Twenty six participants (21–29 year age range, M = 24 years, 15 female) were thus included in the statistical analyses.

STIMULI AND PROCEDURE

In experiment 1, two targets were presented in a rapid serial visual presentation (RSVP) sequence consisting of three-digit distractors (**Figure 1A**). The first target (T1) was an addition task displayed in red, which the participants were instructed to solve immediately and to retain the answer until prompted to respond. The second target (T2) was a letter (A, S, D, or F) flanked by two randomly assigned digits. By presenting T1 and T2 in a specific time sequence, visibility of T2 is severely reduced. This phenomenon is usually explained as an effect of attentional processing of T1 that hinders processing of T2, either by depleting resources or through attentional control mechanisms (Raymond et al., 1992; Martens and Wyble, 2010). A key goal of the present experimental paradigm was to allow for a relatively long T2 stimulus duration (e.g., 133 ms) by generating a strong AB effect. To this end T1 consisted of an attentionally demanding, but mathematically simple addition task, under the assumption that a more demanding T1 enhances the AB effect (Martens and Wyble, 2010). Stimulus duration was initially set to 133 ms and then adjusted online (shifted up or down between blocks in steps of one display refresh rate (60 Hz); each of six blocks consisted of 42–60 trials, depending on T1 performance (incorrect T1 response automatically added a trial until a correct T1 response or the upper trial limit was reached), totaling in 252–360 trials) to ensure an approximate 50/50 distribution between seen/unseen trials despite individual differences.

A critical parameter for the AB is the lag between T1 and T2, such that if the lag is too short or too long, the effect is attenuated or canceled. We were here mainly interested in parameters that cause the AB to occur in approximately 50% of the trials to allow a comparison of conscious vs. non-conscious processing. However, to verify the phenomenon as such with the current experimental protocol, we used both short (strong AB effect) and long (weak AB effect) lags between T1 and T2. Each short-lag RSVP started with a central crosshair for 3000 ms after which 4–9 distractors were presented, followed by T1, one intervening distractor, and T2. Distractors were then displayed until 5, 10, or 15 s had passed from T2 presentation. The long-lag RSVPs (n = 36) were randomly interspersed with short-lag trials (n = 180), differing in the number of distractors between T1 and T2, and in delay time (always 10 s; **Figure 1B**). During trials with the longest T1-T2 lag, T2 was presented at the end

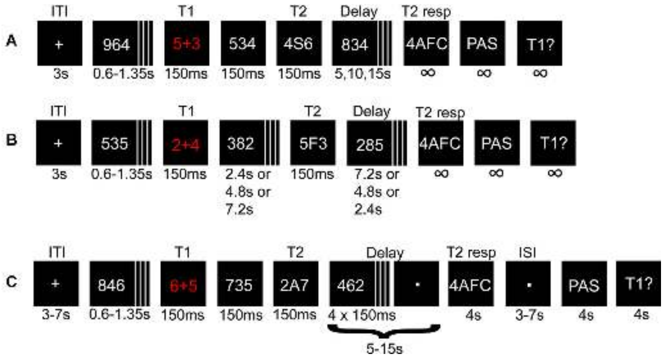

FIGURE 1 | Illustration depicting the procedure for experiment 1 and 2. A mathtask (T1) and a letter (T2) flanked by distracters were presented in a rapid serial visual presentation (RSVP). The solution to T1, and T2 identity, were held in memory during a variable delay until probed for responses. **(A)** Short- and **(B)** long-lag trials in experiment 1 and the pre-fMRI session. In **(B)** the variable delays before and after T2 presentation were adjusted such that T2 appeared early, in the middle, or at the end of the RSVP. **(C)** Illustration of a short-lag trial in the fMRI session.

of the RSVP. Thus, the participants had to attend the entire RSVP, during which they did not know whether they missed a short-lag T2 or if a long-lag T2 was to be presented at the very end.

After each RSVP participants answered three queries: (i) a four-alternative forced-choice (4AFC) task regarding T2 identity; (ii) to what degree they had a subjective experience of seeing T2; (iii) the answer to T1. The conscious experience of seeing T2 or not was judged on a four-point perceptual awareness scale (PAS; Overgaard et al., 2006; Sandberg et al., 2010). The PAS scale descriptions used were: (1) no visual experience of T2; (2) vague visual experience of T2; (3) almost clear visual experience of T2; and (4) clear visual experience of T2. All ratings above 1 were treated as indicating conscious perception. When T2 was unseen the participants had been instructed to guess when prompted regarding T2 identity by responding with the first letter that came to mind. To create a reference condition for the subjective experience of not seeing T2, there were also 36 trials without a T2 (replaced by a distractor). Therefore, trials with and without T2 that were given PAS ratings of 1 shared the same (lack of) subjective experience of T2.

Experiment 2 consisted of a pre-fMRI session (four blocks totaling 124–144 trials depending on T1 performance and/or if the upper trial limit was reached; 84 short-lag, 20 long-lag, and 20 T2-absent trials) and an fMRI session (two runs, each run consisted of two blocks totaling 116 trials; 80 short-lag and 36 T2-absent trials), and used a similar procedure as for experiment 1 with the following changes. The pre-fMRI session was modified by using three instead of four steps in the PAS scale, because combining "clear or almost clear experience" (steps 3 and 4) was easier and more intuitive to use for the participants, without losing the important distinction between "no experience" and "vague experience". The pre-fMRI session enabled us to screen for unsuitable participants, e.g., "non-blinkers" (Martens and Wyble, 2010), and to adjust the individual stimulus durations for approximate 50/50 distribution of seen/unseen trials before the fMRI session, although stimulus durations were also adjusted during the fMRI session. The fMRI session had some additional changes (**Figure 1C**): (i) added jitter to the inter-trial interval (3–7 s), delay-period (5–15 s in steps of 1 s instead of 5 s), and the inter-stimulus interval (3–7 s) between the 4AFC and PAS response, to reduce correlations between components of the statistical model; (ii) the delay period consisted of passively viewing a dot, but four distractors remained after T2 presentation to uphold the AB effect and overwrite any iconic memory representations; and (iii) response-time limits of 4 s for all responses.

fMRI ACQUISITION

The fMRI session in experiment 2 was conducted at 3T with a GE 3 Tesla Discovery MR750 scanner with a 32-channel receive-only head coil. Each subject underwent one session with two functional runs (784 volumes each) of scanning using a T2*-weighted gradient echo pulse sequence, echo planar imaging, field of view = 25 cm, matrix size = 96 × 96, slice thickness = 2.9 mm, 37 slices with no inter-slice skip and an ASSET acceleration factor of 2. The volumes covered the whole cerebrum and most of the cerebellum, the acquisition orientation was oblique axial and aligned with the anterior and posterior commissure, and was scanned in interleaved order with TE = 30 ms, TR = 2 s, flip angle = 90°. Between the two functional runs a high-resolution T1-weighted structural image was collected FSPGR with TE = 3.2 ms, TR = 8.2 ms, TI = 450 ms, and flip angle = 12°.

DATA PROCESSING AND STATISTICAL ANALYSIS

For the behavioral results of both experiments, only trials where T1 was answered correctly were used in the analyses, because if T1 was not processed there would likely not be an AB during T2.

For response times a two SD cut-off was used for each condition (PAS > 1 and PAS = 1) and participant separately.

The software used for processing and analysis of fMRI data was SPM8 (Welcome Trust Centre for Neuroimaging, London, UK), run in Matlab 7.11 (Mathworks, Inc., Sherborn, MA, USA). Before preprocessing a manual quality control was conducted using in-house software. Preprocessing was done in the following order: slice-timing correction to the first slice using Fourier phase-shift interpolation method, head-motion correction with unwarping of B0 distortions, DARTEL normalization (Ashburner, 2007) using a 12-parameter affine transformation model to MNI anatomical space, and an 8 mm FWHM Gaussian smoothing. DARTEL normalization and smoothing was applied on the contrast images after intrasubject model estimation.

For intrasubject modeling a General Linear Model (GLM) with restricted maximum likelihood estimation was used. The model consisted of the following regressors of interest: trial epochs (stimulus presentation, delay, and response)-by-trial type (short-lag or T2 absent)-by-T1 accuracy (correct or incorrect)-by-PAS rating (1, 2, or 3), and inter-trial interval. Missed 4AFC responses (because of time limit), head motion (six parameters) and physiological noise (six parameters) estimated with tCompCor (temporal variation in white matter and cerebral spinal fluid; Behzadi et al., 2007), were included as nuisance regressors. All regressors except for head motion and physiological noise were convolved with the "canonical" hemodynamic response function. The high-pass filter had a cut-off at 128 s, and the autocorrelation model was global AR (1).

For each individual and each trial epoch (stimulus presentation, delay, and response), the following conditions were compared: T2-seen > T2-absent and T2-unseen > T2-absent. Average signal change across conditions during each of the three trial epochs relative to a low-level baseline (ITI) was defined as (T2-seen + T2-unseen + T2-absent)/3 > ITI. Model estimations from each individual were taken into second-level random-effects analyses (one-sample t-tests) to account for inter-individual variability. The statistical inferences were made on the whole brain with $p \leq 0.001$ uncorrected for multiple comparisons, cluster extent ≥20.

RESULTS

BEHAVIORAL RESULTS

The average T1 performance was 82% for experiment 1, 86% for the pre-fMRI session, and 80% for the fMRI session. The average proportion of unseen T2s (PAS = 1 given correct T1) when T2 was present were 45% for experiment 1, 37% for the pre-fMRI session, and 32% for the fMRI session. The average proportion of false alarms were 12% for experiment 1, 12% for the pre-fMRI session, and 25% for the fMRI session.

There was a significant difference in T2 performance between short- and long-lag trials, (experiment 1: $F_{(1,20)} = 74.70$, $p < 0.001$; experiment 2: $F_{(1,25)} = 30.94$, $p < 0.001$), thereby replicating previous research on the AB in that the T1-T2 time interval had a high impact on T2 performance.

In experiment 1, there was a main effect of T2 visibility (T2-seen and T2-unseen trials) on T2 performance, but no main effect of delay time (**Table 1**). There was a significant visibility-by-delay time interaction, such that seen T2 performance declined over time, whereas unseen T2 performance did not. Critically, performance on unseen T2 was significantly better than chance (0.25) at all three time points (**Table 1**). A second analysis only on T2-unseen (PAS = 1) trials revealed no significant main effect of delay time (**Table 1**). Response times for T2-seen trials were significantly shorter than T2-unseen trials ($t_{(1,20)} = -3.18$, $p = 0.005$) and T2-absent trials ($t_{(1,20)} = -2.57$, $p = 0.02$). Response times for T2-unseen trials were not significantly different from T2-absent trials ($t_{(1,20)} = -0.39$, $p = 0.70$).

All behavioral results in experiment 1 were replicated in experiment 2 for the pre-fMRI and fMRI session with two exceptions: (i) there was no significant T2 visibility-by-delay time interaction in experiment 2; and (ii) T2-unseen performance was at chance-level during the third time point in the pre-fMRI session (**Table 1**).

In experiment 2, response times for T2-seen trials were significantly shorter than T2-unseen trials (pre-fMRI: $t_{(1,25)} = -3.43$, $p = 0.002$; fMRI: $t_{(1,25)} = -2.64$, $p = 0.01$) and T2-absent trials (pre-fMRI: $t_{(1,25)} = -3.76$, $p = 0.001$; fMRI: $t_{(1,25)} = -3.46$, $p = 0.002$). Response times for T2-unseen trials were not significantly different from T2-absent trials (pre-fMRI: $t_{(1,25)} = -1.28$, $p = 0.21$; fMRI: $t_{(1,25)} = -1.56$, $p = 0.13$).

fMRI RESULTS

Using the attentional blink to investigate memory

The AB phenomenon has been used extensively in previous research to investigate attention and also conscious experience. It is less commonly used to investigate aspects of memory. Here, we have used the AB to manipulate conscious perception but have designed the experiment similar to protocols investigating working memory, with a stimulus presentation, followed by a short delay, followed by a probe. Similar to previous neuroimaging research on working memory, we used multiple regression to identify BOLD signal change specifically related to different within-trial components (stimulus presentation, delay, and probe). To verify this approach, we first compared each trial epoch with a low-level baseline (the inter-trial interval), averaged across the three conditions (the participants were required to keep information online during the delay period even for T2-absent trials, as T1 was present in all trials).

Comparing stimulus presentation with the low-level baseline revealed BOLD signal change in widespread frontal, parieto-temporal, and cerebellum regions bilaterally (**Figure 2A**). The delay-period comparison against the low-level baseline revealed sustained BOLD signal change in the left inferior frontal gyrus, and bilateral occipital cortex (**Figure 2B**). Comparing the response and the low-level baseline revealed bilateral BOLD signal change in the frontal, parietal, temporal, and occipital cortex, and cerebellum (**Figure 2C**).

Comparing T2-present and T2-absent trials

Comparing T2-seen with T2-absent trials during T2 stimulus-presentation revealed wide-spread BOLD signal change, most

Table 1 | Behavioral results.

	F/t (df)	*p*-value	PAS = 1			PAS > 1		
			M (SE)	95% CI		M (SE)	95% CI	
				LL	UL		LL	UL
Experiment 1								
T2 visibility	129.98 (1, 20)	**< 0.001**	**0.34 (0.02)**	0.30	0.38	**0.75 (0.04)**	0.68	0.83
Delay	1.02 (2, 40)	0.37						
T2 visibility*delay	4.15 (2, 40)	**0.02**						
T2 acc. 5 s			**0.32 (0.02)**	0.28	0.37	**0.79 (0.03)**	0.73	0.86
T2 acc. 10 s			**0.36 (0.03)**	0.30	0.41	**0.73 (0.05)**	0.62	0.84
T2 acc. 15 s			**0.33 (0.03)**	0.28	0.39	**0.74 (0.03)**	0.66	0.81
T2 delay (PAS = 1)	1.14 (2, 40)	0.33						
T2 response times	−3.17 (1, 20)	**0.005**	1962 (88)	1780	2145	1706 (89)	1521	1890
Experiment 2: Pre-fMRI								
T2 visibility	148.25 (1, 25)	**< 0.001**	**0.33 (0.02)**	0.28	0.39	**0.78 (0.03)**	0.70	0.85
Delay	2.09 (2, 50)	0.13						
T2 visibility*delay	0.14 (2, 50)	0.87						
T2 acc. 5 s			**0.34 (0.03)**	0.28	0.41	**0.80 (0.03)**	0.74	0.87
T2 acc. 10 s			**0.34 (0.04)**	0.27	0.42	**0.78 (0.04)**	0.70	0.86
T2 acc. 15 s			0.31 (0.04)	0.23	0.39	**0.74 (0.04)**	0.66	0.83
T2 delay (PAS = 1)	0.41 (2, 50)	0.66						
T2 response times	−3.43 (1, 25)	**0.002**	1619 (143)	1326	1913	1228 (57)	1112	1345
Experiment 2: fMRI								
T2 visibility	49.73 (1, 25)	**< 0.001**	**0.41 (0.42)**	0.32	0.50	**0.75 (0.05)**	0.65	0.85
Delay	1.55 (1, 25)	0.23						
T2 visibility*delay	0.72 (2, 50)	0.49						
T2 acc. 5–8 s			**0.35 (0.05)**	0.26	0.46	**0.74 (0.05)**	0.63	0.86
T2 acc. 9–11 s			**0.47 (0.06)**	0.34	0.60	**0.76 (0.06)**	0.64	0.88
T2 acc. 12–15 s			**0.39 (0.06)**	0.27	0.51	**0.74 (0.05)**	0.64	0.84
T2 delay (PAS = 1)	1.56 (2, 50)	0.23						
T2 response times	−2.64 (1, 25)	**0.01**	1193 (70)	1049	1337	1066 (43)	977	1155

*Note: Significant values are in boldface. CI = confidence interval; LL = lower limit; UL = upper limit; T2 visibility = the main effect of visibility/PAS on T2 performance; Delay = the main effect of delay time on T2 performance; T2 visibility*delay = the interaction effect between visibility and delay time on T2 performance; T2 acc. = T2 performance after X s delay time; T2 delay (PAS = 1) = main effect of delay time on unseen T2 performance only; T2 response times = t-test comparison between seen and unseen T2 response times (ms). T-values are reported for response time comparison, and F-values for all other comparisons.*

notably in the left inferior temporal gyrus, parieto-occipital, and frontal cortex (**Figure 3A**), but also in the right parietal cortex and the left hippocampus. There was a similarly wide-spread, but much less pronounced, pattern of BOLD signal change limited to the left inferior temporal gyrus, superior parietal lobule, inferior frontal gyrus, and precentral gyrus when comparing T2-unseen with T2-absent trials.

When comparing the delay period of T2-seen with T2-absent trials no significant BOLD signal change was found. The comparison between the delay period of T2-unseen and T2-absent trials revealed sustained BOLD signal change in the right mid-lateral prefrontal cortex (mid-lateral PFC; crossing inferior and middle frontal gyrus BA 45/46), right orbitofrontal cortex (OFC), and bilateral cerebellum (crus II). Comparing the delay period of T2-unseen with T2 seen trials revealed a cluster in the mid-lateral PFC ($t = -3.76$) that overlapped with the cluster found when comparing T2-unseen with T2-absent. To investigate the relationship between BOLD signal change during the delay period and task performance we correlated beta values from the right mid-lateral PFC and OFC with unseen T2 performance across participants. There was no significant relation between regional BOLD signal change in mid-lateral PFC and task performance ($r_{(25)} = -0.12$, $p = 0.58$) or OFC and task performance ($r_{(25)} = -0.07$, $p = 0.75$).

When comparing the responses for T2-seen with T2-absent trials BOLD signal change was found in the left inferior temporal gyrus, postcentral gyrus, and superior parietal lobule (**Figure 3C**). Comparing T2-unseen and T2-absent trial responses revealed no significant BOLD signal change.

DISCUSSION

In contrast to common belief, we have shown that non-consciously presented perceptual information can be durable, here lasting with unaffected strength for at least 15 s. This result replicate and extend previous findings of durable perceptual representations by Hesselmann et al. (2011) of 4 s, and Soto et al. (2011) of 5 s, and further challenge durability as a hallmark for conscious experience. The sustained BOLD signal change in mid-lateral PFC during maintenance of the non-consciously perceived information is consistent with working memory. Our findings are in line with Dutta et al. (2014) who found that BOLD signal change in DLPFC over entire trials of masked information predicted memory performance. Critically, however, our within-trial separation of BOLD signal confirmed that sustained BOLD

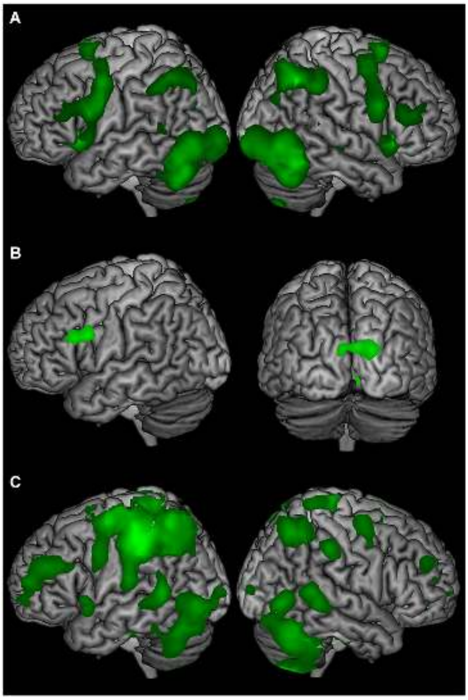

FIGURE 2 | The average BOLD signal change of T2-seen, T2-unseen, and T2-absent trials compared to a low-level baseline (ITI) for (A) stimulus presentation, (B) the delay period, and (C) the response epochs.

signal change in the PFC occurs during maintenance of non-conscious information.

Several different memory mechanisms, such as working memory, priming, or iconic memory, could in principle be responsible for durable non-conscious representations. Although the cortical aspects of iconic memory (fragile visual short-term memory) can retain information for at least 4 s, it is easily overwritten by task-irrelevant distracters (Sligte et al., 2008). Similar to Soto et al. (2011), the current AB paradigm used irrelevant distracters after stimulus-presentation. This, combined with the current delay of up to 15 s, makes iconic memory an unlikely explanation.

Furthermore, the sustained BOLD signal change found in right mid-lateral PFC, OFC, and cerebellum during the maintenance of non-consciously perceived information is inconsistent with priming. The exact mechanisms of priming are still unclear and may depend on task, material, and whether the material is masked or not (Henson, 2003). However, priming effects are likely either residual neural activity and/or latent neural changes (e.g., long-term potentiation) that facilitates or inhibits subsequent processing (Grill-Spector et al., 2006; Marsolek et al., 2010). Neither residual neural activity nor latent neural changes would elicit a sustained BOLD signal change over the entire delay period.

Thus, only working memory remains as a possible explanation for the durable non-conscious perceptual representations. Although working memory is commonly seen as intimately linked with conscious experience (Baars and Franklin, 2003; Dehaene and Changeux, 2011; Baddeley, 2012), it has been suggested on theoretical grounds that working memory indeed can operate non-consciously, as below-threshold activity (Fuster, 1995). Other recent empirical studies also support the notion of non-conscious working memory. Soto et al. (2011) have shown that 1–2 non-consciously presented items can be maintained during a distractor-filled delay of up to 5 s. Dutta et al. (2014) linked performance on a delayed task with non-consciously presented information to BOLD signal change in the PFC. Furthermore, they showed that transcranial direct current stimulation of the PFC modulated performance, demonstrating that the PFC is causally involved in such delayed performance. Pan et al. (2013) demonstrated that when a non-consciously maintained item matched an interocularly suppressed item, the latter had prior entry into conscious awareness compared to non-matching items. Critically, none of the results above can be explained by priming mechanisms, because Soto et al. (2011) and Dutta et al. (2014) used delayed cue-target orientation discrimination tasks where cue and target never matched, and Pan et al. (2013) did several control experiments to show that mere exposure to the masked stimulus was not enough for prior entry, it had to be actively maintained.

Neuroimaging studies of humans and single-unit recordings in primates have revealed that working memory maintenance is associated with sustained neural activity in lateral PFC and posterior regions. Prefrontal cortex activity has been interpreted as representing a preparatory action set, while posterior activity represents the memory content (Cabeza and Nyberg, 2000; Curtis and D'Esposito, 2003; Fuster, 2009; Sreenivasan et al., 2014). Consistent with our findings and non-conscious working memory, recent findings suggest that non-conscious information can activate task sets (Reuss et al., 2011) and lead to lateral prefrontal BOLD signal change (Lau and Passingham, 2007).

However, it is unclear why the sustained BOLD signal change in the right PFC did not correlate with memory performance, as might have been expected (Pessoa et al., 2002; Sakai et al., 2002; Wager et al., 2014). Different lateral PFC regions have been related to different executive functions during working memory (Wager and Smith, 2003; Nee et al., 2013), and it is conceivable that not all functions necessarily predicts performance. Progressively rostral regions of the PFC seem to support more abstract representations and more complex rules (Fuster, 2008; Badre and D'Esposito, 2009). The mid-lateral PFC signal change could therefore be related to task set maintenance (Sakai and Passingham, 2003) or preparation for future action (Pochon et al., 2001), while the OFC signal change might relate to maintenance of more abstract representations (Nee and Brown, 2012) or executive control functions related to coordination and

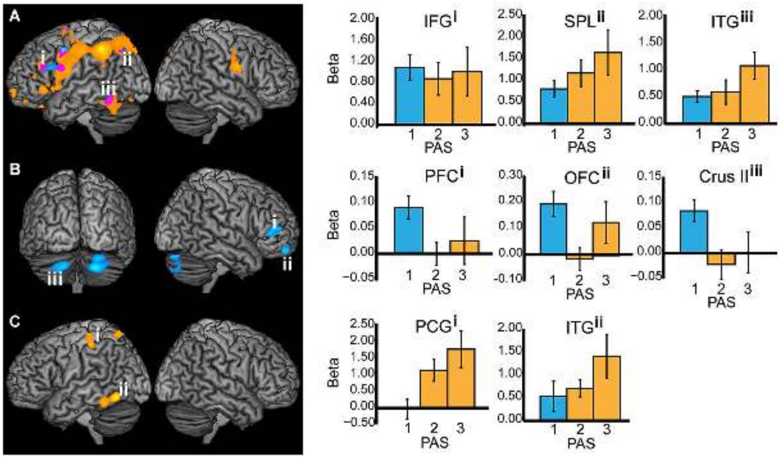

FIGURE 3 | All working memory epochs with T2-seen > T2-absent trials in orange, T2-unseen > T2-absent trials in blue, and the overlap in purple. The Y-axis: Beta values, X-axis: PAS = perceptual awareness scale, bar colors correspond to the PAS classification of T2-seen and T2-unseen trials, error bars: standard error of the mean. **(A)** Stimulus presentation (IFG = inferior frontal gyrus, $X = -46$, $Y = 28$, $Z = 24$; SPL = superior parietal lobule, $X = -26$, $Y = -68$, $Z = 44$; ITG = inferior temporal gyrus, $X = -52$, $Y = -54$, $Z = -18$). **(B)** Delay period—displayed at $p < 0.005$ for illustrative purposes (PFC = prefrontal cortex, $X = 50$, $Y = 42$, $Z = 8$; OFC = orbitofrontal cortex, $X = 20$, $Y = 58$, $Z = -22$; Crus II in the cerebellum, $X = -20$, $Y = -74$, $Z = -38$. **(C)** Response (PCG = postcentral gyrus, $X = -42$, $Y = -30$, $Z = 52$; ITG = inferior temporal gyrus, $X = -58$, $Y = -62$, $Z = -12$).

simultaneous use of several cognitive processes such as maintenance, manipulation, and monitoring (Owen et al., 2005; Barbey et al., 2011).

The lateral prefrontal cortex has contralateral input and output projections that form closed loops with crus I and II of the cerebellum (Kelly and Strick, 2003; Bostan et al., 2013). Although the exact function of the cerebellum in cognition is unclear, BOLD signal change in crus I and II together with prefrontal cortex have been associated with verbal working memory and executive functions (Stoodley and Schmahmann, 2009).

The comparison of average signal change across conditions with a low-level baseline revealed differences in the left inferior frontal gyrus, and bilaterally in the occipital cortex during maintenance. The signal change in inferior frontal gyrus is consistent with neuroimaging findings related to sub-vocalization during working memory maintenance (Paulesu et al., 1993), which suggest that sub-vocalization was used to remember the consciously perceived T1 and T2. It is unclear why there was no significant BOLD signal change during the delay period when comparing T2-seen and T2-absent trials. Possibly, the difference between consciously maintaining 1 vs. 2 items (T1 vs. T1 + T2) was not big enough to elicit detectable BOLD signal change. The fact that there was a difference during delay between T2-unseen and T2-absent trials therefore suggest that temporary maintenance of information can engage different processes depending on if the information to be maintained is conscious or not. Although consciously perceived verbal information is likely maintained by way of sub-vocalization, it seems unlikely for non-consciously perceived verbal information to be so. Instead, the non-consciously perceived verbal information might be maintained as visual representations. Consciously sub-vocalizing T1 while non-consciously maintaining T2 could therefore act as a distraction or dual-task/process that leads to increased PFC involvement (D'Esposito et al., 1995; Feredoes et al., 2011). Different maintenance processes for conscious and non-conscious representations would explain the stronger BOLD signal change in the right mid-lateral PFC during maintenance of non-consciously compared to consciously perceived information.

For T2-seen compared to T2-absent trials, BOLD signal change was evident in the left inferior temporal gyrus during stimulus presentation and response. This likely reflects representational-level processing of T2, as previous neuroimaging research has demonstrated signal change in inferior temporal gyrus during single letter perception (Flowers et al., 2004; Park et al., 2012). However, there was no significant BOLD signal change in inferior temporal gyrus during the delay period following consciously and non-consciously perceived information. Unit recordings in primates have established that stimulus-specific neurons in sensory regions are temporarily activated during working memory retention (Fuster and Jervey, 1981; Miyashita and Chang, 1988; Miller and Desimone, 1994), but human neuroimaging findings have not been as consistent. Several working-memory studies with letter-tasks did not reveal inferior temporal gyrus involvement during maintenance and manipulation of letters (Cabeza and Nyberg, 2000). Although the signal change in stimulus-specific sensory regions fail to reach the same elevated level as frontal regions during maintenance, it has recently been shown that multivariate pattern analysis can detect stimulus-specific information (Riggall and Postle, 2012). It could therefore be the case that simply maintaining one item in working memory is generally not enough to elicit elevated levels of BOLD signal in stimulus-specific sensory regions.

Consistent with the GNW model and previous findings a comparison between T2-seen and T2-absent trials revealed large and wide-spread BOLD signal change in the left inferior temporal gyrus (Flowers et al., 2004; Park et al., 2012), and PPN (Rees et al., 2002; Naghavi and Nyberg, 2005; Dehaene and Changeux, 2011) during the stimulus presentation. Corresponding BOLD signal change was also found in inferior temporal gyrus during the comparison between T2-unseen and T2-absent trials, which is consistent with previous studies on non-conscious perception (Rees et al., 2002; Marois et al., 2004; Heinzel et al., 2008; Dehaene and Changeux, 2011). However, the wide-spread PPN involvement in T2-unseen compared to T2-absent trials during stimulus presentation is inconsistent with the GNW model. Although previous masking studies have found that non-conscious processing tends to be limited (but not exclusive) to posterior sensory regions (Dehaene et al., 1994, 2001; Kouider et al., 2007), more recent studies have implicated the PPN (Kranczioch et al., 2005; Diaz and McCarthy, 2007) and prefrontal cortex (Lau and Passingham, 2007; Wokke et al., 2011) in non-conscious processing. It is conceivable that our wide-spread BOLD signal change was the result of relatively long stimulus presentation durations ($M = 129$ ms, during fMRI) compared to the 17–50 ms commonly used in masking paradigms. Furthermore, that non-consciously presented information not only can activate, but also maintain durable representations in higher-order regions, such as the mid-lateral PFC and the OFC, for 15 s is inconsistent with the GNW model's predictions that extended PFC activity and durable maintenance is unique to conscious processing (Dehaene and Changeux, 2011).

An alternative model (to GNW) of conscious experience predicts that neural processes are accompanied by conscious experience if (and only if) the neural activity reaches a certain undefined threshold. Such amplitude models (Fuster, 1995; Zeki, 2001) predict that conscious experience should correlate with higher amplitudes, and has no *a priori* reason to assume that higher-order regions, wide-spread cortical interactions, or higher-level cognitive functions such as working memory should be uniquely reserved for conscious experiences. Instead, amplitude models predict that non-conscious functions, and their underlying neural activity would be similar but weaker, which is what neuroimaging findings seem to indicate (Rees et al., 2002; Lau and Passingham, 2006; Diaz and McCarthy, 2007; Wokke et al., 2011).

Comparing T2-seen with T2-absent trials during the response revealed BOLD signal change in the left inferior temporal gyrus, postcentral gyrus, and superior parietal lobule. The inferior temporal gyrus signal change is consistent with previous research, where signal change usually involves a transient peak in both prefrontal and posterior regions during stimulus presentation and response epochs, and a lower, sustained BOLD signal change during the maintenance epoch (Druzgal and D'Esposito, 2003). The contrast between T2-unseen and T2-absent trials did not reveal a significant BOLD signal difference. Interestingly, the nominal BOLD signal change in inferior temporal gyrus during response was comparable to the (significant) signal change during stimulus presentation (**Figure 3**). Thus, it seems that the inferior temporal gyrus signal change was non-significant due to higher variability rather than amplitude, which is in line with recent proposals of (low) variability as a hallmark of conscious processes (Schurger et al., 2010).

There are several valid approaches to measure and operationally define conscious experience and the lack thereof. We have here used a subjective measure of awareness. Compared with objective measures, subjective measures are more liberal and risk overestimating the extent of non-conscious processing. However, the more conservative objective measures may instead underestimate the effect of non-conscious processing by (miss) attributing them as conscious (Merikle et al., 2001). Possibly, using conservative measures of conscious experience may have biased previous findings that show non-conscious processes to be short-lived. The particular measure used here ("The Perceptual Awareness Scale") has been shown superior to other subjective measures such as confidence ratings (Cheesman and Merikle, 1986) and post-decision wagering (Persaud et al., 2007) in terms of sensitivity and exhaustiveness (Sandberg et al., 2010).

The use of a four-alternative forced-choice task with fixed response options opens up the possibility that the participants decided how to respond during the delay instead of when prompted. This is, however, not consistent with the response-time data. If participants already decided what letter to guess before prompted, then the response times for unseen T2s should be the same as for seen T2s, because in both cases the participants would know their prospective response in advance. Instead, if the participants decided what to guess when prompted, the unseen trials should have a slower response time than seen trials to account for extra deliberation time, which was the case. Indeed, there was no significant difference in response time between T2-unseen and T2-absent trials. Furthermore, the sustained BOLD signal change found when comparing the T2-unseen and T2-absent trials cannot be explained by such behavior. Given that the lack of perceptual awareness of T2 was identical for both conditions, the (deliberate) strategy must have been identical as well.

In sum, we have demonstrated that non-conscious perceptual representations can last for up to 15 s despite irrelevant distracters, and argue that this effect is best explained in terms of (non-conscious) working-memory mechanisms. Most notably, we found sustained BOLD signal change in the right mid-lateral PFC and OFC during the delay period. In addition, we found widespread frontal and parieto-temporal BOLD signal change during non-conscious perception. Although it is too early to say whether these durable non-conscious representations can truly be understood as working memory processes and not some other form of non-conscious memory, the current findings combined with recent similar research are compelling (Soto and Silvanto, 2014). Important next steps will be to convincingly show that a stimulus-specific representation is actively maintained during the delay-period (e.g., by using multivariate pattern classification algorithms), to compare potential functional differences between conscious and non-conscious working-memory operations (e.g., in terms of capacity and/or fidelity), and to determine how conscious and non-conscious working memory operations interact.

ACKNOWLEDGMENTS

The project was supported by grants from the medical faculty and the collecting foundation for medical research at Umeå University, and a grant from the Swedish Research Council, to Johan Eriksson. The project was also supported by grants from the Torsten and Ragnar Söderberg foundation and the Knut and Alice Wallenberg foundation to Lars Nyberg. We thank Lars Nyberg and Greger Orädd for comments on the manuscript, Anders Lundquist for statistical advice, and Micael Andersson for technical assistance.

REFERENCES

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage* 38, 95–113. doi: 10.1016/j.neuroimage.2007.07.007

Baars, B. J., and Franklin, S. (2003). How conscious experience and working memory interact. *Trends Cogn. Sci.* 7, 166–172. doi: 10.1016/s1364-6613(03)00056-1

Baddeley, A. (2012). Working memory: theories, models and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

Badre, D., and D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical?. *Nat. Rev. Neurosci.* 10, 659–669. doi: 10.1038/nrn2667

Barbey, A. K., Koenigs, M., and Grafman, J. (2011). Orbitofrontal contributions to human working memory. *Cereb. Cortex* 21, 789–795. doi: 10.1093/cercor/bhq153

Behzadi, Y., Restom, K., Liau, J., and Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37, 90–101. doi: 10.1016/j.neuroimage.2007.04.042

Bostan, A. C., Dum, R. P., and Strick, P. L. (2013). Cerebellar networks with the cerebral cortex and basal ganglia. *Trends Cogn. Sci.* 17, 241–254. doi: 10.1016/j.tics.2013.03.003

Cabeza, R., and Nyberg, L. (2000). Imaging cognition II: an empirical review of 275 PET and fMRI studies. *J. Cogn. Neurosci.* 12, 1–47. doi: 10.1162/08989290051137585

Cheesman, J., and Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Can. J. Psychol.* 40, 343–367. doi: 10.1037/h0080103

Curtis, C. E., and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423. doi: 10.1016/s1364-6613(03)00197-9

D'Esposito, M., Detre, J., Alsop, D. C., Shin, R. K., Atlas, S., and Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature* 378, 279–281. doi: 10.1038/378279a0

Dehaene, S., and Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227. doi: 10.1016/j.neuron.2011.03.018

Dehaene, S., Charles, L., King, J.-R., and Marti, S. (2014). Toward a computational theory of conscious processing. *Curr. Opin. Neurobiol.* 25, 76–84. doi: 10.1016/j.conb.2013.12.005

Dehaene, S., Naccache, L., Clec'H, G., Koechlin, E., Dehaene-Lambertz, G., Moortele, P.-F., et al. (1994). Semantic priming in a single-word shadowing task. *Am. J. Psychol.* 107, 245–260. doi: 10.2307/1423039

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, 752–758. doi: 10.1038/89551

Diaz, M. T., and McCarthy, G. (2007). Unconscious word processing engages a distributed network of brain regions. *J. Cogn. Neurosci.* 19, 1768–1775. doi: 10.1162/jocn.2007.19.11.1768

Druzgal, T. J., and D'Esposito, M. (2003). Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. *J. Cogn. Neurosci.* 15, 771–784. doi: 10.1162/089892903322370708

Dutta, A., Shah, K., Silvanto, J., and Soto, D. (2014). Neural basis of non-conscious visual working memory. *Neuroimage* 91C, 336–343. doi: 10.1016/j.neuroimage.2014.01.016

Feredoes, E., Heinen, K., Weiskopf, N., Ruff, C., and Driver, J. (2011). Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. *Proc. Natl. Acad. Sci. U S A* 108, 17510–17515. doi: 10.1073/pnas.1106439108

Flowers, D. L., Jones, K., Noble, K., VanMeter, J., Zeffiro, T. A., Wood, F. B., et al. (2004). Attention to single letters activates left extrastriate cortex. *Neuroimage* 21, 829–839. doi: 10.1016/j.neuroimage.2003.10.002

Fuster, J. M. (1995). *Memory in the Cerebral Cortex—An Empirical Approach to Neural Networks in the Human and Nonhuman Primate.* Cambridge, MA: MIT Press.

Fuster, J. M. (2008). *The Prefrontal Cortex.* 4th Edn. London: Academic Press.

Fuster, J. M. (2009). Cortex and memory: emergence of a new paradigm. *J. Cogn. Neurosci.* 21, 2047–2072. doi: 10.1162/jocn.2009.21280

Fuster, J. M., and Jervey, J. P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212, 952–955. doi: 10.1126/science.7233192

Greenwald, A. G., Draine, S. C., and Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science* 273, 1699–1702. doi: 10.1126/science.273.5282.1699

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006

Heinzel, A., Hautzel, H., Poeppel, T. D., Boers, F., Beu, M., and Mueller, H.-W. (2008). Neural correlates of subliminal and supraliminal letter processing–an event-related fMRI study. *Conscious. Cogn.* 17, 685–699. doi: 10.1016/j.concog.2008.01.008

Henson, R. N. (2003). Neuroimaging studies of priming. *Prog. Neurobiol.* 70, 53–81. doi: 10.1016/s0301-0082(03)00086-8

Hesselmann, G., Hebart, M., and Malach, R. (2011). Differential BOLD activity associated with subjective and objective reports during "blindsight" in normal observers. *J. Neurosci.* 31, 12936–12944. doi: 10.1523/jneurosci.1556-11.2011

Kelly, R. M., and Strick, P. L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *J. Neurosci.* 23, 8432–8444.

Koch, C., and Crick, F. (2001). The zombie within. *Nature* 411:893. doi: 10.1038/35082161

Kouider, S., and 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, 857–875. doi: 10.1098/rstb.2007.2093

Kouider, S., Dehaene, S., Jobert, A., and Le Bihan, D. (2007). Cerebral bases of subliminal and supraliminal priming during reading. *Cereb. Cortex* 17, 2019–2029. doi: 10.1093/cercor/bhl110

Kranczioch, C., Debener, S., Schwarzbach, J., Goebel, R., and Engel, A. K. (2005). Neural correlates of conscious perception in the attentional blink. *Neuroimage* 24, 704–714. doi: 10.1016/j.neuroimage.2004.09.024

Lau, H. C., and Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc. Natl. Acad. Sci. U S A* 103, 18763–18768. doi: 10.1073/pnas.0607716103

Lau, H. C., and Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *J. Neurosci.* 27, 5805–5811. doi: 10.1523/jneurosci.4335-06.2007

Logan, G. D., and Crump, M. J. C. (2010). Cognitive illusions of authorship reveal hierarchical error detection in skilled typists. *Science* 330, 683–686. doi: 10.1126/science.1190483

Luck, S., Vogel, E., and Shapiro, K. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature* 383, 616–618. doi: 10.1038/383616a0

Marois, R., Yi, D.-J., and Chun, M. M. (2004). The neural fate of consciously perceived and missed events in the attentional blink. *Neuron* 41, 465–472. doi: 10.1016/s0896-6273(04)00012-1

Marsolek, C. J., Deason, R. G., Ketz, N. A., Ramanathan, P., Bernat, E. M., Steele, V. R., et al. (2010). Identifying objects impairs knowledge of other objects: a relearning explanation for the neural repetition effect. *Neuroimage* 49, 1919–1932. doi: 10.1016/j.neuroimage.2009.08.063

Martens, S., and Wyble, B. (2010). The attentional blink: past, present and future of a blind spot in perceptual awareness. *Neurosci. Biobehav. Rev.* 34, 947–957. doi: 10.1016/j.neubiorev.2009.12.005

Merikle, P. M., Smilek, D., and Eastwood, J. D. (2001). Perception without awareness: perspectives from cognitive psychology. *Cognition* 79, 115–134. doi: 10.1016/s0010-0277(00)00126-8

Miller, E. K., and Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science* 263, 520–522. doi: 10.1126/science.8290960

Miyashita, Y., and Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331, 68–70. doi: 10.1038/331068a0

Naghavi, H. R., and Nyberg, L. (2005). Common fronto-parietal activity in attention, memory and consciousness: shared demands on integration? *Conscious. Cogn.* 14, 390–425. doi: 10.1016/j.concog.2004.10.003

Nee, D. E., and Brown, J. W. (2012). Rostral-caudal gradients of abstraction revealed by multi-variate pattern analysis of working memory. *Neuroimage* 63, 1285–1294. doi: 10.1016/j.neuroimage.2012.08.034

Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., et al. (2013). A meta-analysis of executive components of working memory. *Cereb. Cortex* 23, 264–282. doi: 10.1093/cercor/bhs007

Overgaard, M., Rote, J., Mouridsen, K., and Ramsøy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Conscious. Cogn.* 15, 700–708. doi: 10.1016/j.concog.2006.04.002

Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59. doi: 10.1002/hbm.20131

Pan, Y., Lin, B., Zhao, Y., and Soto, D. (2013). Working memory biasing of visual perception without awareness. *Atten. Percept. Psychophys.* 76, 2051–2062. doi: 10.3758/s13414-013-0566-2

Park, J., Hebrank, A., Polk, T. A., and Park, D. C. (2012). Neural dissociation of number from letter recognition and its relationship to parietal numerical processing. *J. Cogn. Neurosci.* 24, 39–50. doi: 10.1162/jocn_a_00085

Paulesu, E., Frith, C. D., and Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature* 362, 342–345. doi: 10.1038/362342a0

Persaud, N., McLeod, P., and Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nat. Neurosci.* 10, 257–261. doi: 10.1038/nn1840

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., et al. (2007). How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science* 316, 904–906. doi: 10.1126/science.1140459

Pessoa, L., Gutierrez, E., Bandettini, P., and Ungerleider, L. (2002). Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron* 35, 975–987. doi: 10.1016/S0896-6273(02)00817-6

Pochon, J. B., Levy, R., Poline, J. B., Crozier, S., Lehéricy, S., Pillon, B., et al. (2001). The role of dorsolateral prefrontal cortex in the preparation of forthcoming actions: an fMRI study. *Cereb. Cortex* 11, 260–266. doi: 10.1093/cercor/11.3.260

Raymond, J., Shapiro, K., and Arnell, K. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol. Hum. Percept. Perform.* 18, 849–860. doi: 10.1037//0096-1523.18.3.849

Rees, G., Kreiman, G., and Koch, C. (2002). Neural correlates of consciousness in humans. *Nat. Rev. Neurosci.* 3, 261–270. doi: 10.1038/nrn783

Reuss, H., Kiesel, A., Kunde, W., and Hommel, B. (2011). Unconscious activation of task sets. *Conscious. Cogn.* 20, 556–567. doi: 10.1016/j.concog.2011.02.014

Riggall, A. C., and Postle, B. R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J. Neurosci.* 32, 12990–12998. doi: 10.1523/jneurosci.1892-12.2012

Sakai, K., and Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nat. Neurosci.* 6, 75–81. doi: 10.1038/nn987

Sakai, K., Rowe, J. B., and Passingham, R. E. (2002). Active maintenance in prefrontal area 46 creates distractor-resistant memory. *Nat. Neurosci.* 5, 479–484. doi: 10.1038/nn846

Sandberg, K., Timmermans, B., Overgaard, M., and Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Conscious. Cogn.* 19, 1069–1078. doi: 10.1016/j.concog.2009.12.013

Schurger, A., Pereira, F., Treisman, A., and Cohen, J. D. (2010). Reproducibility distinguishes conscious from nonconscious neural representations. *Science* 327, 97–99. doi: 10.1126/science.1180029

Sergent, C., Baillet, S., and Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nat. Neurosci.* 8, 1391–1400. doi: 10.1038/nn1549

Sligte, I. G., Scholte, H. S., and Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLoS One* 3:e1699. doi: 10.1371/journal.pone.0001699

Soto, D., Mäntylä, T., and Silvanto, J. (2011). Working memory without consciousness. *Curr. Biol.* 21, R912–R913. doi: 10.1016/j.cub.2011.09.049

Soto, D., and Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends Cogn. Sci.* 18, 520–525. doi: 10.1016/j.tics.2014.06.005

Sreenivasan, K. K., Curtis, C. E., and D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends Cogn. Sci.* 18, 82–89. doi: 10.1016/j.tics.2013.12.001

Stoodley, C. J., and Schmahmann, J. D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *Neuroimage* 44, 489–501. doi: 10.1016/j.neuroimage.2008.08.039

Wager, T. D., and Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274. doi: 10.3758/cabn.3.4.255

Wager, T. D., Spicer, J., Insler, R., and Smith, E. E. (2014). The neural bases of distracter-resistant working memory. *Cogn. Affect. Behav. Neurosci.* 14, 90–105. doi: 10.3758/s13415-013-0226-y

Wokke, M. E., van Gaal, S., Scholte, H. S., Ridderinkhof, K. R., and Lamme, V. A. F. (2011). The flexible nature of unconscious cognition. *PLoS One* 6:e25729. doi: 10.1371/journal.pone.0025729

Zeki, S. (2001). Localization and globalization in conscious vision. *Annu. Rev. Neurosci.* 24, 57–86. doi: 10.1146/annurev.neuro.24.1.57

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 13 August 2014; paper pending published: 18 October 2014; accepted: 04 November 2014; published online: 21 November 2014.

Citation: Bergström F and Eriksson J (2014) Maintenance of non-consciously presented information engages the prefrontal cortex. Front. Hum. Neurosci. ***8****:938. doi: 10.3389/fnhum.2014.00938*

This article was submitted to the journal Frontiers in Human Neuroscience.

II

frontiers
in Psychology

ORIGINAL RESEARCH
published: 30 September 2015
doi: 10.3389/fpsyg.2015.01470

The conjunction of non-consciously perceived object identity and spatial position can be retained during a visual short-term memory task

Fredrik Bergström[1,2]* *and Johan Eriksson*[1,2]

[1] Umeå Center for Functional Brain Imaging, Umeå University, Umeå, Sweden, [2] Department of Integrative Medical Biology, Physiology Section, Umeå University, Umeå, Sweden

Edited by:
Juha Silvanto,
University of Westminster, UK

Reviewed by:
Jan Brascamp,
Michigan State University, USA
David Soto,
Imperial College London, UK

***Correspondence:**
Fredrik Bergström,
Department of Integrative Medical Biology, Physiology Section, Umeå University, The Biology Building, 901 87 Umeå, Västerbotten, Sweden
fredrik.bergstrom@umu.se

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *01 June 2015*
Accepted: *14 September 2015*
Published: *30 September 2015*

Citation:
Bergström F and Eriksson J (2015) The conjunction of non-consciously perceived object identity and spatial position can be retained during a visual short-term memory task. Front. Psychol. 6:1470. doi: 10.3389/fpsyg.2015.01470

Although non-consciously perceived information has previously been assumed to be short-lived (<500 ms), recent findings show that non-consciously perceived information can be maintained for at least 15 s. Such findings can be explained as working memory without a conscious experience of the information to be retained. However, whether or not working memory can operate on non-consciously perceived information remains controversial, and little is known about the nature of such non-conscious visual short-term memory (VSTM). Here we used continuous flash suppression to render stimuli non-conscious, to investigate the properties of non-consciously perceived representations in delayed match-to-sample (DMS) tasks. In Experiment I we used variable delays (5 or 15 s) and found that performance was significantly better than chance and was unaffected by delay duration, thereby replicating previous findings. In Experiment II the DMS task required participants to combine information of spatial position and object identity on a trial-by-trial basis to successfully solve the task. We found that the conjunction of spatial position and object identity was retained, thereby verifying that non-conscious, trial-specific information can be maintained for prospective use. We conclude that our results are consistent with a working memory interpretation, but that more research is needed to verify this interpretation.

Keywords: non-conscious, durability, priming, conscious experience, perception, working memory

Introduction

Non-consciously perceived information can be processed at all levels of the visual system (Rees et al., 2002; Kouider and Dehaene, 2007), and influence executive functions (Lau and Passingham, 2007; van Gaal et al., 2010). It is less clear for how long non-consciously perceived information can be retained and influence behavior. It was previously assumed that non-consciously perceived information is extremely fleeting, and would cease to be detectable within 500 ms after stimulus offset (Greenwald et al., 1996; Mattler, 2005; Dehaene and Changeux, 2011). However, recent studies have found that non-consciously perceived information is more durable than previously assumed. For example, Reber et al. (2012) found that repeatedly presented masked word-pairs could influence decision-making 1 min later, and that hippocampal BOLD signal change at encoding predicted the outcomes. Bar and Biederman (1998, 1999) found behavioral evidence of non-consciously encoded visual repetition priming effects on naming tasks after 15–20 min, and

Gaillard et al. (2007) found electrophysiological repetition effects (but no behavioral effects) 47 min after single presentations of masked words. These studies suggest that non-consciously perceived information can form long-lasting latent neural representations reminiscent of consciously encoded hippocampus-based memory (Cabeza and Nyberg, 2000) and visual repetition priming (Henson, 2003) mechanisms.

However, whether or not non-consciously perceived information can be actively maintained for prospective use after stimulus offset (i.e., working memory) is still unclear. Initial findings suggest that working memory can operate on non-consciously perceived information. Soto et al. (2011) used delayed cue-target orientation discrimination tasks to demonstrate that 1–2 non-consciously presented items can be maintained during a distractor-filled delay of up to 5 s. Dutta et al. (2014) later used fMRI to link performance on the delayed cue-target discrimination task to BOLD signal change in the dorsolateral prefrontal cortex (DLPFC), and transcranial direct current stimulation to causally link DLPFC to performance. Pan et al. (2014) found that when a non-consciously presented face matched an interocularly suppressed face, the latter had prior entry into conscious awareness compared to non-matching items, but only when the face was needed for prospective use. We have previously used an attentional-blink paradigm to demonstrate that non-consciously presented information can be maintained during a distracter-filled delay for up to 15 s, which was associated with BOLD signal change in the prefrontal cortex (Bergström and Eriksson, 2014). Accordingly, it has been suggested that working memory can operate on non-consciously perceived information (Soto and Silvanto, 2014).

Little is known about the properties of this non-conscious visual short-term memory (VSTM) and we here aim to further clarify the nature of non-conscious memory representations. To this end we used continuous flash suppression (CFS; Tsuchiya and Koch, 2005) to render stimuli non-conscious during a delayed match-to-sample (DMS) task. CFS has become popular in studies of non-conscious processes, because CFS can in an efficient and easily controlled manner suppress the conscious experience of stimuli for long periods of time (e.g., up to 3 min; Tsuchiya and Koch, 2005). Such long presentation durations can potentially enable more reliable and durable non-conscious representations, compared to for example masked stimuli. Here we wanted to replicate our previous findings that non-consciously perceived information can be retained for at least 15 s (Bergström and Eriksson, 2014) with CFS, and determine which properties of the non-conscious representations that can be retained and influence DMS performance.

Experiment I

In Experiment I we used CFS to suppress faces that expressed angry or neutral emotions presented within a spatial quadrant of the visual field, with variable delay durations (5 or 15 s), to investigate if non-consciously perceived visual information can be retained for similar delay durations as with attentional blink (Bergström and Eriksson, 2014) and masking (Soto et al., 2011) manipulations.

Materials and Methods

Participants

Nineteen healthy participants were recruited from the Umeå University campus area. All participants had normal or corrected-to-normal vision, right-eye dominance, gave written informed consent, and were paid for participation. Two participants were excluded for systematically giving the same response instead of guessing when stimulus was not experienced, one for having an extremely high DMS accuracy of stimuli that were reportedly not experienced ($d' = 2.16$, > 3 SD above the group mean), and one where CFS did not consistently suppress the appearance of the target stimuli. Thus, 15 participants ($M = 23$ years, 11 females) were included in the statistical analyses.

Stimuli and Procedure

The experiment consisted of 270 DMS trials dispersed on three presentation conditions (50 conscious, 170 non-conscious, and 50 "baseline" trials, see below for description), randomly distributed between two delay durations (5 or 15 s) between the stimulus presentation and the DMS response (**Figure 1**). The stimuli consisted of gray-scaled and Gaussian blurred (1 pixel radius) images of faces (height: 1.6–1.8°, width: 1.3–1.5°, average luminance: 4.3 cd/m^2) expressing angry or neutral emotions (four faces of each emotion) at 75% opaqueness level (to lower contrast relative to a gray background), and positioned in one of four spatial quadrants. The stimuli were presented on a computer monitor in front of a mirror stereoscope that isolated the visual input from the left side of the monitor to the participants left eye, and vice versa for the right side. The monitor was placed at a length that enabled all visual input to be presented within 6° horizontally and 9.6° vertically. The stimulus to be held in VSTM was presented for 3 s, either to both eyes simultaneously (consciously experienced), or only to the non-dominant (left) eye while colored squares of random composition (mondrians; height: 4.2°, width: 4.2° luminance: ~30 cd/m^2) where flashed with a frequency of 10 Hz to the dominant eye to suppress the stimulus from conscious experience (Tsuchiya and Koch, 2005). During the baseline trials mondrians were presented to the dominant eye while an empty gray background (height: 4.2°, width: 4.2°, luminance: 8.6 cd/m^2) was presented to the non-dominant eye. Critically, the visual experience of baseline and non-conscious trials is the same (experiencing only mondrians).

After the delay period a DMS response was prompted by a probe with an identical face, emotion, and spatial position as previously presented (match), or with a different face, emotion, and spatial position (non-match). The participants were instructed to decide if the previously presented stimulus was a spatial match or non-match to the probe (thus, the participants were not instructed to remember face identity or emotional expression). If they had not experienced the target stimulus (i.e., only experienced mondrians) they were instructed to guess on the first alternative that came to mind (match/no match). After the DMS response, they were prompted to estimate the conscious experience of the stimulus on a three-point perceptual awareness scale (PAS; Sandberg et al., 2010). The participants were instructed and trained to use the PAS scale as follows: 1 = no perceptual experience, 2 = vague perceptual experience, and 3 =

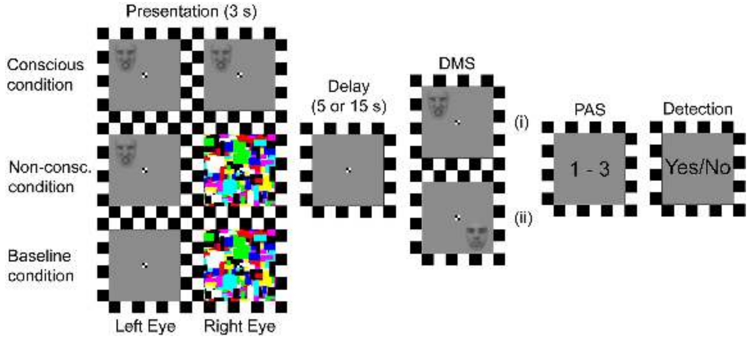

FIGURE 1 | Trial procedure. Depending on the presentation condition, two identical stimuli (faces), stimulus and mondrians, or empty background and mondrians were presented to the left and right eye respectively. The spatial position of the stimulus was then to be retained for a 5 or 15 s delay period, until a probe prompted the participants to respond whether or not the probe's position matched the previously presented position. Next, the participants gave an estimate of their perceptual experience of the stimulus. Finally, they responded whether or not a stimulus had been present. DMS, delayed match-to-sample task; PAS, perceptual awareness scale; (i), probe matches presentation; (ii), probe does not match presentation.

clear or almost clear perceptual experience of the target stimulus. Lastly, the participants were prompted to make a detection response to determine if a target stimulus had been presented at all (yes or no). If they had not perceptually experienced a stimulus they were to guess per the same instructions as for the DMS task. After participants had received instructions, they performed a practice run of the experiment with the instructor until their behavior was consistent with the instructions, after which the actual experiment started. After the experiment the participants were debriefed and asked about their behavior in relation to the instructions.

Statistical Analyses

Trials with a DMS response time (RT) of <250 ms or >M + 3 SD were excluded as outliers prior to any statistical analyses (Ratcliff, 1993). Only trials in the baseline and non-conscious presentation conditions with PAS = 1, and trials in the conscious condition with PAS = 3 were used in the statistical analyses, and will for simplicity hereby be referred to as baseline, non-conscious, and conscious trials. Signal detection theory (d′) was used to calculate performance on the discrimination (DMS) and detection tasks (Macmillan and Creelman, 1991). For DMS d′ the signal was defined as the spatial quadrant where a face appeared. Hits were therefore defined as a match between presentation target and probe together with a "match" response, and false alarms (FAs) as a non-match between presentation target and probe together with a "match" response. For the detection task, hits were defined as the presence of a target stimulus together with a "yes" response, and FA were defined as the absence of a target stimulus (i.e., baseline trials) together with a "yes" response.

Results

The presence of CFS efficiently suppressed the visual input to the non-dominant eye from conscious experience in the non-conscious condition (**Table 1**). In the following results, all trials with PAS > 1 were removed to ensure no visibility at all of the target stimulus in non-conscious and baseline conditions. A 2 × 2 (visual experience × delay time; **Figure 2**) repeated measures ANOVA was used to test if DMS performance (d′) changed as a function of (delay) time and/or visibility. There was a main effect of visibility [$F_{(1,\ 14)} = 8.3$, $p < 0.001$], but not time [$F_{(1,\ 14)} = 0.02$, $p = 0.89$] and no interaction effect between visibility and time [$F_{(1,\ 14)} = 0.94$, $p = 0.35$]. Since there was neither a main effect of time nor an interaction effect between time and visibility we proceeded to treat the two time points together. T-tests were used to determine if memory performance (**Figure 2A**) was above chance (i.e., $d' > 0$), and revealed that non-conscious [$t_{(14)} = 2.24$, $p = 0.02$, one-tailed, $M = 0.22$, $SE = 0.10$, $P(\text{hits}) = 0.47$, $P(\text{FA}) = 0.39$] and conscious [$t_{(14)} = 127$, $p < 0.001$, one-tailed, $M = 3.35$, $SE = 0.03$, $P(\text{hits}) = 0.95$, $P(\text{FA}) = 0.09$] DMS d′ was greater than zero. Thus, the DMS results replicated previous research (Bergström and Eriksson, 2014).

TABLE 1 | CFS efficiency.

Conditions	PAS		
	1	2	3
Baseline	48	2	0
Non-consc.	128	33	9
Conscious	1	1	48

Average trial frequency of each presentation condition distributed on reported perceptual awareness scale (PAS) responses per participant.

For the detection response, a 2 × 2 repeated measures ANOVA showed a main effect of visibility [$F_{(1,\ 14)} = 13.4$, $p < 0.001$], but not time [$F_{(1,\ 14)} = 1.75$, $p = 0.21$] and no interaction effect between visibility and time [$F_{(1,\ 14)} = 1.27$, $p = 0.28$]. Treating both time points together, t-tests showed that detection

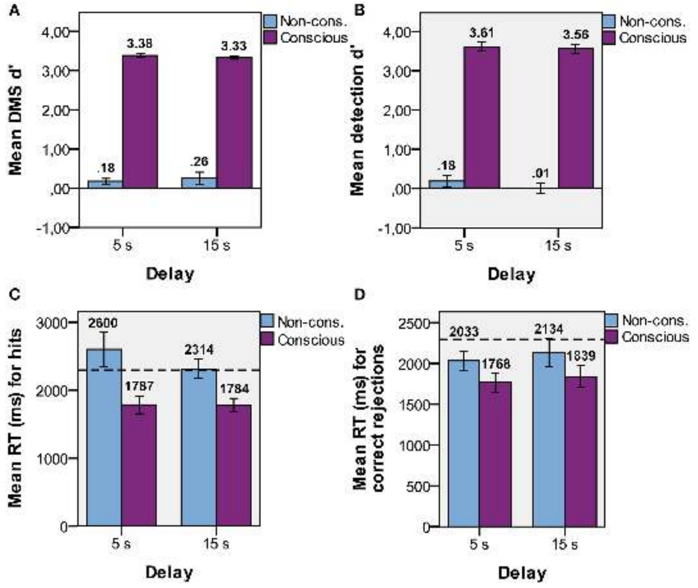

FIGURE 2 | Task performance. Mean d′ performance for non-conscious and conscious **(A)** delayed match-to-sample (DMS; discrimination), and **(B)** detection tasks for each delay duration with standard error. Mean DMS response time (RT) for **(C)** hits and **(D)** correct rejections for each delay duration with standard error. Dotted lines represent baseline RT.

d′ (**Figure 2B**) was greater than zero for conscious trials [$t_{(14)} = 33$, $p < 0.001$, one-tailed, $M = 3.58$, $SE = 0.11$, $P(\text{hits}) = 0.97$, $P(\text{FA}) = 0.11$], but not for non-conscious trials [$t_{(14)} = 0.83$, $p = 0.21$, one-tailed, $M = 0.10$, $SE = 0.12$, $P(\text{hits}) = 0.51$, $P(\text{FA}) = 0.48$]. Thus, participants could not tell whether a target stimulus had been presented or not for the non-conscious trials. The non-conscious DMS d′ and detection d′ were not significantly different from each other [$t_{(14)} = 1.07$, $p = 0.30$].

Repeated measures ANOVAs were used to determine if response time of the DMS task (RT) differed as a function of time and/or visibility for hits and correct rejections (CRs) separately. We hypothesized that RT for both hits and CRs would be faster than baseline if the representation was held in working memory, while only hits would be faster if the representation depended on repetition priming (since there is no stimulus repetitions in CRs). The ANOVA on hits revealed a main effect of visibility [$F_{(1,\ 14)} = 21$, $p < 0.001$], but not time [$F_{(1,\ 14)} = 0.96$, $p = 0.34$] and no interaction effect between visibility and time [$F_{(1,\ 14)} = 1.02$, $p = 0.33$]. Paired t-tests on RTs combined over time showed that hits (**Figure 2C**) were faster than baseline ($M = 2295$, $SE = 116$) for conscious trials [$t_{(14)} = -4.32$, $p = 0.001$, $M = 1782$, $SE = 89$] but not for non-conscious trials [$t_{(14)} = 1.31$, $p = 0.21$, $M = 2428$, $SE = 147$]. The ANOVA on CRs revealed that the main effect of visibility was at trend [$F_{(1,\ 14)} = 3.38$, $p = 0.09$], but no effect of time [$F_{(1,\ 14)} = 1.07$, $p = 0.32$] or an interaction effect [$F_{(1,\ 14)} = 0.03$, $p = 0.86$]. RT on CRs (**Figure 2D**) was faster than baseline for both conscious [paired t-tests; $t_{(14)} = -3.72$, $p = 0.002$, $M = 1797$, $SE = 117$] and non-conscious trials [$t_{(14)} = -3.69$, $p = 0.002$, $M = 2080$, $SE = 131$].

If participants generated a guess of the target directly after the presentation and consciously held that guess in WM during the delay, then the RTs for non-conscious and baseline trials should be equal to the RT for conscious trials. To investigate this issue, we calculated paired t-tests for RTs averaged across delay-time, hits, misses, FAs, and CRs. RT for conscious trials ($M = 1788$, $SE = 93$) were faster than RTs for non-conscious [$t_{(14)} = -3.46$, $p = 0.004$, $M = 2241$, $SE = 131$] and baseline [$t_{(14)} = -4.38$, $p = 0.001$] trials, while there was no difference between non-conscious and baseline trials [$t_{(14)} = -0.97$, $p = 0.35$]. These results show that there is a difference between conscious and non-conscious and also baseline trials, presumably because of the extra deliberation time before the guessing response, and indicate that participants had not already guessed the target before the probe appeared. Guessing performance above chance level is thus consistent with non-consciously retained information.

Experiment II

Experiment I showed that non-consciously perceived information was retained for up to 15 s in a DMS paradigm using CFS to present stimuli non-consciously. However, since the samples were either fully matched or not matched at all in relation to the three information components (spatial position, face identity, and emotion), it is not possible to know which component(s) that were retained in memory. Indeed, spatial attentional effects rather than item-specific memory representations may have driven performance on non-conscious trials. In Experiment II we therefore investigated the content of the information retained by presenting tools at specific spatial positions, and by using probe stimuli that matched in terms of both object and position information (similar to Experiment I), only object, only position, or neither of the information components. It is possible that non-conscious memory mechanisms only retain one of the information components (e.g., the spatial position) despite a conscious task set to retain both. However, if item-specific representations, defined as an arbitrary combination of spatial position and object identity that change from trial to trial, are retained across the delay, the conjunction of spatial- and object information would be preferentially retained over one of the components.

Furthermore, given the uniform effect across delay duration (5 vs. 15 s) in Experiment I, all trials in Experiment II had a 5 s delay period. We chose to use tools as stimuli instead of faces for two reasons: (i) we reasoned that within-category discrimination might be easier for tools because of their distinctly different shapes/features, and (ii) previous research have indicated that tools may be more reliably processed during interocular suppression with CFS (Fang and He, 2005; Almeida et al., 2008).

Materials and Methods

Participants

Nineteen healthy participants were recruited from the Umeå University campus area. All participants had normal or corrected-to-normal vision, right-eye dominance, gave written informed consent, and were paid for participation. Two participants were excluded for systematically giving the same response instead of guessing when the stimulus was not experienced, and one where CFS did not consistently suppress the appearance of the target stimuli. Thus, 16 participants ($M = 25$ years, 9 females) were included in the statistical analyses.

Stimuli and Procedure

The experiment consisted of 396 trials dispersed on three presentation conditions (117 conscious, 222 non-conscious, and 57 baseline trials). The procedure for Experiment II (**Figure 3**) was identical to that of Experiment I in all aspects except for the following alterations. Firstly, the delay duration between stimuli presentation and response was set to 5 s. Secondly, the detection response was excluded. Thirdly, the stimuli material used was changed from faces to six different gray silhouettes of tools (height: 1.7°, width: 1.7°, Gaussian blur: 1 pixel radius, luminance: 8.4 cd/m^2). The participants were instructed to remember both the tool and its spatial position. For it to be a "match," the probe stimulus had to be the same tool and be in the same spatial position (full match). If the probe contained the same tool at a different spatial position (object match), different tool at the same spatial position (spatial match), or different tool

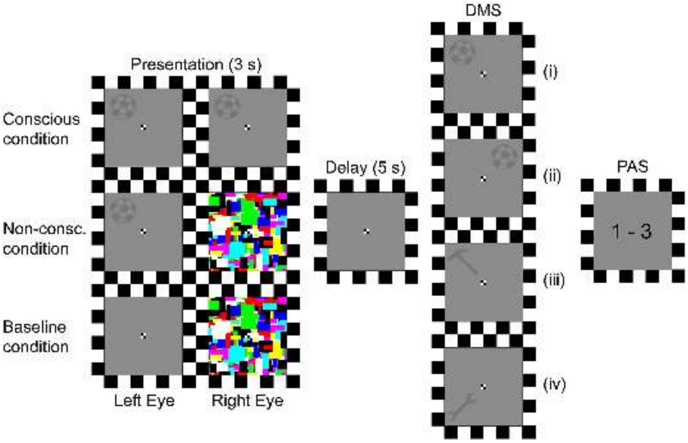

FIGURE 3 | Trial procedure. Depending on the presentation condition, two identical stimuli (tools), stimulus and mondrians, or empty background and mondrians were presented to the left and right eye respectively. The stimulus' position and identity was then to be retained for a 5 s delay period, until a probe prompted the participants to respond whether or not the probe's position and identity matched the previously presented position and identity. Finally, they gave an estimate of their perceptual experience of the stimulus. DMS, delayed match-to-sample task; PAS, perceptual awareness scale; (i), Probe identity and position matches presentation; (ii), Probe identity matches presentation; (iii), Probe position matches presentation; (iv), Probe does not match presentation.

at a different spatial position (non-match), it should be answered with a "no match" response.

Statistical Analyses

Outliers and inclusion criteria for statistical analyses were identical to Experiment I. For DMS d′ the signal was defined as the object identity and its spatial position. Hits were therefore defined as a (position and identity) match between sample and probe together with a "match" response, and FAs as a non-match (which includes cases where only position, only identity, or neither was a match) between sample and probe together with a "match" response.

Results

The presence of CFS efficiently suppressed the visual input to the non-dominant eye from conscious experience in the non-conscious condition (**Table 2**). In the following results, all trials with PAS > 1 were removed to ensure no visibility at all of the target stimulus in non-conscious and baseline conditions. *T*-tests were used to determine if memory performance was above chance. DMS d′ (**Figure 4A**) was greater than zero for non-conscious trials [$t_{(15)} = 3.17, p = 0.003$, one-tailed, $M = 0.22$, $SE = 0.07$, $P(\text{hits}) = 0.56$, $P(\text{FA}) = 0.48$] and conscious trials [$t_{(15)} = 38, p < 0.001$, one-tailed, $M = 4.12, SE = 0.11, P(\text{hits}) = 0.97, P(\text{FA}) = 0.03$].

To examine the relative contributions of spatial position, object identity, and the conjunction of both in driving the non-conscious DMS d′ effect, we looked at the proportion of trials within each of the categories full match, object match, spatial match, and non-match, where participants responded "match." In short, we compared hit rate (full match) and the false alarm rates when only object identity matched (object FA), only spatial position matched (spatial FA), and when neither matched (baseline FA; **Figure 5**). A repeated-measures ANOVA showed that there was a difference [$F_{(3,\ 15)} = 3.57, p = 0.02$] among hits/FAs. The planned paired *t*-tests (corrected for three comparisons with the Holm-Bonferroni procedure) revealed that the hit rate for non-conscious trials were greater than baseline FA [$t_{(15)} = 2.29, p = 0.018$, one-tailed], object FA [$t_{(15)} = 2.67, p = 0.009$, one-tailed], and spatial FA [$t_{(15)} = 3.12, p = 0.004$, one-tailed]. There was no difference among FA rates [repeated-measures ANOVA, $F_{(2,\ 15)} = 0.19, p = 0.83$]. These results confirm that the conjunction of spatial position and object identity was retained throughout the delay.

Paired *t*-test comparisons of conscious DMS reaction time (RT; ms; **Figure 4B**) demonstrated that hits [$t_{(15)} = 4.22, p = 0.001, M = 1194, SE = 87$] and CRs [$t_{(15)} = 2.79, p = 0.014, M = 1448, SE = 110$] were faster than baseline ($M = 1809, SE = 180$). However, the paired *t*-test comparisons on non-conscious DMS RT for hits [$t_{(15)} = -1.09, p = 0.29, M = 1721, SE = 145$]

TABLE 2 | CFS efficiency.

Conditions	PAS		
	1	2	3
Baseline	53	3	0
Non-consc.	194	26	2
Conscious	3	2	112

Average trial frequency of each presentation condition distributed on reported perceptual awareness scale (PAS) responses per participant.

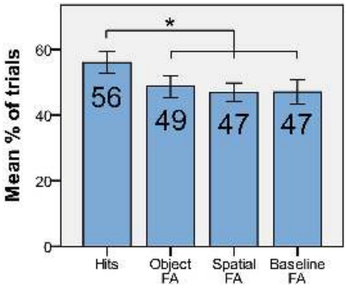

FIGURE 5 | The average proportion of hits, object-FA, spatial-FA, and baseline-FA with standard error for non-conscious trials. *$p < 0.05$, one-tailed.

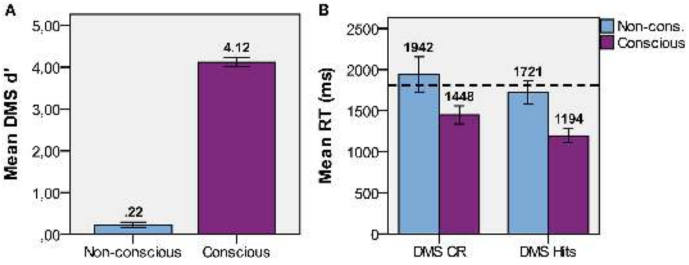

FIGURE 4 | Task performance. (A) Mean d′ performance for non-conscious and conscious delayed match-to-sample (DMS; discrimination) tasks. **(B)** Mean DMS response time (RT) for hits and correct rejections with standard error, and dotted lines that represents baseline RT.

and CRs [$t_{(15)} = 1.83$, $p = 0.09$, $M = 1942$, $SE = 213$] were not faster than baseline. Indeed, the CRs RT was at trend with regard to being slower than baseline RT.

Paired t-tests of RTs (aggregating hits, misses, FAs, and CRs) showed that RTs of conscious trials ($M = 1342$, $SE = 99$) were faster than RTs of non-conscious [$t_{(15)} = -3.33$, $p = 0.005$, $M = 1841$, $SE = 187$] and baseline [$t_{(15)} = -3.44$, $p = 0.004$] trials, while there was no difference between non-conscious and baseline trials [$t_{(15)} = 0.70$, $p = 0.50$]. These results are consistent with Experiment I in showing that participants had not already decided what to guess before the probe appeared.

Discussion

In Experiment I, we found that non-consciously perceived visuospatial information can be maintained and influence behavior 5–15 s after stimulus offset, thereby replicating recent findings (Bergström and Eriksson, 2014). Although the participants were instructed to only remember the spatial position, it is possible that face identity and/or emotion information also contributed to performance. For CFS, some have found more reliable processing in the dorsal visual pathway compared to the ventral pathway (Fang and He, 2005; Almeida et al., 2008), but emotional faces (Tsuchiya et al., 2009; Faivre et al., 2012) have also shown to be processed to some extent (see Yang et al., 2014, for a review of CFS research).

In Experiment II, we specifically investigated the content of the non-conscious VSTM representations. We found that spatial position and object identity could be arbitrarily bound on a trial-by-trial basis, and retained for prospective use after 5 s. The specificity of the retained information (i.e., the conjunction of spatial position and object identity rather than one of the components) was further substantiated by the fact that the average proportion of hits was significantly greater than baseline FA, whereas spatial FA and object FA were not. However, in ongoing experiments in our lab we have noted that spatial information may be driving behavioral performance in similar tasks, and we therefore do not exclude the possibility that specific information components may dominate regulation of behavior in different experiments. For example, it is possible that the conscious task set of the participants can bias the non-conscious short-term retention toward a specific content.

Although our current finding that the conjunction of spatial position and object identity can be retained is consistent with previous research on working memory (Jiang et al., 2000; Olson, 2005; Wood, 2011), it does not rule out the influence of other memory mechanisms. However, the use of only six tools and four spatial positions that were reused over the course of the experiment provided proactive interference from previous trials, which makes short-lived unconsolidated long-term memory mechanisms an unlikely explanation for the non-conscious DMS d′ performance (Endress and Potter, 2014). For the same reason, non-consciously encoded hippocampus-based long-term memory (Degonda et al., 2005; Reber et al., 2012; Duss et al., 2014), where each trial is encoded as a specific episodic representation (i.e., a specific temporal context paired with a specific stimulus content), also seems unlikely. Furthermore, the non-conscious retention of information cannot be explained by residual activity in rod and cone receptors (i.e., iconic memory) since such activity tend to subside within 1 s of stimulus offset (Coltheart, 1980; Sligte et al., 2008).

It has recently been suggested that, contrary to common belief, visual object recognition may be position dependent (Kravitz et al., 2008). Kravitz et al. (2010) found decreased object priming performance with changes in spatial position, and a weaker ability to differentiate between object identity (based on BOLD signal change in high-level object-selective cortex) across positions compared to within positions. They therefore argued that high-level object representations are position dependent. If object representations indeed are position dependent and thereby automatically processed together during visual object recogniton, then it is reasonable to assume that this also is the case in lower-level visual memories (e.g., repetition priming), and not unique to higher-level visual memories (e.g., working memory). If true, the arbitrary, trial-specific matching of spatial position and object identity does not by itself exclude an influence of repetition priming.

The DMS task together with the arbitrary binding of spatial position and a limited set of objects would effectively minimize any automatic stimuli-response mapping, and thereby minimize potential priming effects. However, masked priming effects can remain without stimulus-response mapping (Van den Bussche et al., 2009). Although non-consciously encoded priming has been assumed to be short-lived (<500 ms; Greenwald et al., 1996; Mattler, 2005; Dehaene and Changeux, 2011), there are cases of long-lasting effects (Bar and Biederman, 1998, 1999). The discrepancy between short- and long-lasting non-consciously encoded priming effects might partly be explained by a focus on semantic rather than repetition priming when drawing conclusions about non-conscious priming overall. For example, it could be that non-consciously encoded semantic priming (≤100 ms; Greenwald et al., 1996; Draine and Greenwald, 1998) is less durable than non-consciously encoded repetition priming (15 and 20 min; Bar and Biederman, 1998, 1999). Indeed, there is a similar difference in longevity between consciously encoded semantic and repetition priming (Henson, 2003). It would therefore be prudent to assume, a priori, that non-consciously encoded visual repetition priming effects might last for a few seconds and possibly affect performance in a DMS task.

We hypothesized that RTs would be faster than baseline for hits and CRs if the information was held in working memory, but only for hits if the facilitation was caused by repetition priming, since there were no stimuli repetitions to be facilitated during CRs. The RTs for conscious trials in both experiments confirmed our hypothesis regarding working memory, but the RTs for non-conscious trials were variable. Consistent with working memory, the first experiment showed faster RTs for CRs compared to baseline, but this was not true for the second experiment that instead showed a trend to the opposite. Furthermore, non-conscious hits were not faster than baseline in any of the experiments, which is inconsistent with repetition priming effects. The absence of repetition priming effects despite significant discrimination performance on the DMS task is in line with our previous findings (Bergström and Eriksson, 2014), the assumption that repetition priming is less sensitive than recognition memory (Berry et al., 2006), and studies showing that priming has a negligible effect on recognition tasks (Poldrack and Logan, 1997; Conroy et al., 2005). Taken together, the non-conscious RTs did not show convincing support for working memory *per se*, but did not support repetition priming as a likely explanation either.

A likely strategy during trials with a consciously seen target is to verbalize the information, which is consistent with debriefing statements from the participants. Relatedly, a possible objection to the interpretation that DMS performance on trials with non-conscious targets reflects non-conscious working memory is that participants might have verbalized a conscious representation (e.g., "hammer in upper right quadrant") by guesswork based on non-conscious perception of the target, and then consciously maintained the guess until probed. Such verbalization should generate similar RTs for conscious and non-conscious trials. However, control analyses showed that the RT for non-conscious trials were slower than for conscious trials, suggesting that participants did not use verbalization during non-conscious trials. This is also what the participants reported during post-experiment debriefing. Nevertheless, in principle, such response-time differences could be caused by increased uncertainty rather than strategy differences. Future research may clarify this issue.

There are several accepted approaches for defining the presence/absence of conscious experience. We have here used a subjective measure (the PAS). Subjective measures risk not being completely exhaustive (Reingold and Merikle, 1988) of conscious experiences, and might therefore overestimate non-conscious effects. On the other hand, objective thresholds have been criticized for not exclusively (Reingold and Merikle, 1988) measuring conscious experience, and thereby underestimating non-conscious effects. Indeed, it has been argued that task performance can be an unreliable measure of conscious experience (Lau, 2008). There is currently no consensus on how to most exhaustively and exclusively measure the absence/presence of conscious experience during perception (Boly et al., 2013). However, Sandberg and colleagues (Sandberg et al., 2010, 2014) have shown the PAS to be more exhaustive (and thus more conservative) than other subjective measures such as confidence ratings (Cheesman and Merikle, 1986), as well as objective measures like post-decision wagering (Persaud et al., 2007) and exclusion tasks (Debner and Jacoby, 1994). Nevertheless, caveats regarding subjective measures should be considered in relation to the current results.

In conclusion, we found that non-consciously perceived visuospatial information could be retained for prospective use at least 15 s after stimuli offset, and that object identity and spatial position could be arbitrarily bound and retained for prospective use with a fidelity high enough to enable within-category discrimination after 5 s. Our findings are consistent with the notion of non-conscious working memory, although we cannot, based on the current experiments, completely rule out other memory mechanisms.

Acknowledgments

The project was supported by grants from the Swedish Research Council, the medical faculty and the collecting foundation for medical research at Umeå University, to JE. The research leading to these results has received funding from the European Union Seventh Framework Program (FP7/2007-2013) under grant agreement no. 604102 (Human Brain Project), to Lars Nyberg. We thank Lars Nyberg for input on the experimental design, Göran Westling for building the stereoscope, Anders Lundqvist for statistical advice, and Johan Sanders for assisting with data collection during Experiment II.

References

Almeida, J., Mahon, B. Z., Nakayama, K., and Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15214–15218. doi: 10.1073/pnas.0805867105

Bar, M., and Biederman, I. (1998). Subliminal visual priming. *Psychol. Sci.* 9, 464–468. doi: 10.1111/1467-9280.00086

Bar, M., and Biederman, I. (1999). Localizing the cortical region mediating visual awareness of object identity. *Proc. Natl. Acad. Sci. U.S.A.* 96, 1790–1793. doi: 10.1073/pnas.96.4.1790

Bergström, F., and Eriksson, J. (2014). Maintenance of non-consciously presented information engages the prefrontal cortex. *Front. Hum. Neurosci.* 8:938. doi: 10.3389/fnhum.2014.00938

Berry, C. J., Henson, R. N. A., and Shanks, D. R. (2006). On the relationship between repetition priming and recognition memory: insights from a computational model. *J. Mem. Lang.* 55, 515–533. doi: 10.1016/j.jml.2006.08.008

Boly, M., Seth, A. K., Wilke, M., Ingmundson, P., Baars, B., Laureys, S., et al. (2013). Consciousness in humans and non-human animals: recent advances and future directions. *Front. Psychol.* 4:625. doi: 10.3389/fpsyg.2013.00625

Cabeza, R., and Nyberg, L. (2000). Imaging cognition II: an empirical review of 275 PET and fMRI studies. *J. Cogn. Neurosci.* 12, 1–47. doi: 10.1162/08989290051137585

Cheesman, J., and Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Can. J. Psychol.* 40, 343–67. doi: 10.1037/h0080103

Coltheart, M. (1980). Iconic memory and visible persistence. *Percept. Psychophys.* 27, 183–228. doi: 10.3758/BF03204258

Conroy, M. A., Hopkins, R. O., and Squire, L. R. (2005). On the contribution of perceptual fluency and priming to recognition memory. *Cogn. Affect. Behav. Neurosci.* 5, 14–203. doi: 10.3758/CABN.5.1.14

Debner, J. A., and Jacoby, L. L. (1994). Unconscious perception: attention, awareness, and control. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 304–17. doi: 10.1037/0278-7393.20.2.304

Degonda, N., Mondadori, C. R. A., Bosshardt, S., Schmidt, C. F., Boesiger, P., Nitsch, R. M., et al. (2005). Implicit associative learning engages the hippocampus and interacts with explicit associative learning. *Neuron* 46, 505–520. doi: 10.1016/j.neuron.2005.02.030

Dehaene, S., and Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227. doi: 10.1016/j.neuron.2011.03.018

Draine, S. C., and Greenwald, A. G. (1998). Replicable unconscious semantic priming. *J. Exp. Psychol. Gen.* 127, 286–303. doi: 10.1037/0096-3445.127.3.286

Duss, S. B., Reber, T. P., Hänggi, J., Schwab, S., Wiest, R., Müri, R. M., et al. (2014). Unconscious relational encoding depends on hippocampus. *Brain* 137, 3355–3370. doi: 10.1093/brain/awu270

Dutta, A., Shah, K., Silvanto, J., and Soto, D. (2014). Neural basis of non-conscious visual working memory. *Neuroimage* 91C, 336–343. doi: 10.1016/j.neuroimage.2014.01.016

Endress, A. D., and Potter, M. C. (2014). Large capacity temporary visual memory. *J. Exp. Psychol. Gen.* 143, 548–565. doi: 10.1037/a0033934

Faivre, N., Berthet, V., and Kouider, S. (2012). Nonconscious influences from emotional faces: a comparison of visual crowding, masking, and continuous flash suppression. *Front. Psychol.* 3:129. doi: 10.3389/fpsyg.2012.00129

Fang, F., and He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat. Neurosci.* 8, 1380–1385. doi: 10.1038/nn1537

Gaillard, R., Cohen, L., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., et al. (2007). Subliminal words durably affect neuronal activity. *Neuroreport* 18, 1527–1531. doi: 10.1097/WNR.0b013e3282f0b6cd

Greenwald, A. G., Draine, S. C., and Abrams, R. L. (1996). Three Cognitive markers of unconscious semantic activation. *Science* 273, 1699–1702. doi: 10.1126/science.273.5282.1699

Henson, R. N. (2003). Neuroimaging studies of priming. *Prog. Neurobiol.* 70, 53–81. doi: 10.1016/S0301-0082(03)00086-8

Jiang, Y., Olson, I. R., and Chun, M. M. (2000). Organization of visual short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 683–702. doi: 10.1037/0278-7393.26.3.683

Kouider, S., and 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, 857–875. doi: 10.1098/rstb.2007.2093

Kravitz, D. J., Kriegeskorte, N., and Baker, C. I. (2010). High-level visual object representations are constrained by position. *Cereb. Cortex* 20, 2916–2925. doi: 10.1093/cercor/bhq042

Kravitz, D. J., Vinson, L. D., and Baker, C. I. (2008). How position dependent is visual object recognition? *Trends Cogn. Sci.* 12, 114–122. doi: 10.1016/j.tics.2007.12.006

Lau, H. C. (2008). "Are we studying consciousness yet?" in *Frontiers of Consciousness: Chichele Lectures,* eds L. Weiskrantz and M. Davies (Oxford: Oxford University Press), 245–258. doi: 10.1093/acprof:oso/9780199233151.003.0008

Lau, H. C., and Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *J. Neurosci.* 27, 5805–5811. doi: 10.1523/JNEUROSCI.4335-06.2007

Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide.* New York, NY: Cambridge University press.

Mattler, U. (2005). Inhibition and decay of motor and nonmotor priming. *Percept. Psychophys.* 67, 285–300. doi: 10.3758/BF03206492

Olson, I. R. (2005). Remembering " what " brings along " where " in visual working memory. *Percept. Psychophys.* 67, 185–194. doi: 10.3758/BF03206483

Pan, Y., Lin, B., Zhao, Y., and Soto, D. (2014). Working memory biasing of visual perception without awareness. *Atten. Percept. Psychophys.* 76, 2051–2062. doi: 10.3758/s13414-013-0566-2

Persaud, N., McLeod, P., and Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nat. Neurosci.* 10, 257–261. doi: 10.1038/nn1840

Poldrack, R. A., and Logan, G. D. (1997). Fluency and response speed in recognition judgments. *Mem. Cognit.* 25, 1–10. doi: 10.3758/BF03197280

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychol. Bull.* 114, 510–32. doi: 10.1037/0033-2909.114.3.510

Reber, T. P., Luechinger, R., Boesiger, P., and Henke, K. (2012). Unconscious Relational inference recruits the hippocampus. *J. Neurosci.* 32, 6138–6148. doi: 10.1523/JNEUROSCI.5639-11.2012

Rees, G., Kreiman, G., and Koch, C. (2002). Neural correlates of consciousness in humans. *Nat. Rev. Neurosci.* 3, 261–270. doi: 10.1038/nrn783

Reingold, E. M., and Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Percept. Psychophys.* 44, 563–575. doi: 10.3758/BF03207490

Sandberg, K., Del Pin, S. H., Bibby, B. M., and Overgaard, M. (2014). Evidence of weak conscious experiences in the exclusion task. *Front. Psychol.* 5:1080. doi: 10.3389/fpsyg.2014.01080

Sandberg, K., Timmermans, B., Overgaard, M., and Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Conscious. Cogn.* 19, 1069–1078. doi: 10.1016/j.concog.2009.12.013

Sligte, I. G., Scholte, H. S., and Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLoS ONE* 3:e1699. doi: 10.1371/journal.pone.0001699

Soto, D., Mäntylä, T., and Silvanto, J. (2011). Working memory without consciousness. *Curr. Biol.* 21, R912–R913. doi: 10.1016/j.cub.2011.09.049

Soto, D., and Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends Cogn. Sci.* 18, 520–525. doi: 10.1016/j.tics.2014.06.005

Tsuchiya, N., and Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nat. Neurosci.* 8, 1096–1101. doi: 10.1038/nn1500

Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., and Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nat. Neurosci.* 12, 1224–1225. doi: 10.1038/nn.2380

Van den Bussche, E., Van den Noortgate, W., and Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychol. Bull.* 135, 452–477. doi: 10.1037/a0015329

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., and Lamme, V. A. (2010). Unconscious activation of the prefrontal no-go network. *J. Neurosci.* 30, 4143–4150. doi: 10.1523/JNEUROSCI.2992-09.2010

Wood, J. N. (2011). When do spatial and visual working memory interact? *Atten. Percept. Psychophys.* 73, 420–439. doi: 10.3758/s13414-010-00

Yang, E., Brascamp, J., Kang, M.-S., and Blake, R. (2014). On the use of continuous flash suppression for the study of visual processing outside of awareness. *Front. Psychol.* 5:724. doi: 10.3389/fpsyg.2014.00724

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

III

Neural evidence for non-conscious short-term memory

Fredrik Bergström[1, 2] & Johan Eriksson[1, 2]

1. Umeå center for Functional Brain Imaging (UFBI), Umeå University, Sweden
2. Department of Integrative Medical Biology, physiology section, Umeå University, Sweden

Corresponding author:

Johan Eriksson

Department of Integrative Medical Biology, physiology section, Umeå University

901 87 Umeå, Sweden

e-mail: johan.eriksson@umu.se; telephone: +46(0)90-786 51 37; fax: +46(0)90-786 66 96

Running title: Non-conscious short-term memory

Abstract

Recent studies have found that non-consciously perceived information can be retained for several seconds, a feat that has been attributed to non-conscious working memory processes. However, these studies have mainly relied on subjective measures of awareness and the neural processes responsible for such short-term retention remains unclear. Here we used continuous flash suppression to render stimuli non-conscious in a delayed match-to-sample task, and fMRI to investigate the neural correlates of non-conscious short-term (5-15 s) retention. The participants' behavioral performance was at chance level when they reported no experience of the sample stimulus. Critically, multivariate pattern analysis of BOLD signal during the delay period could classify presence vs. absence of the sample stimulus based on signal change in frontal cortex, and its spatial position based on signal change in occipital cortex. In addition, univariate analyses revealed increased BOLD signal change in prefrontal regions during memory recognition. Thus, our findings demonstrate short-term maintenance of information presented non-consciously, as confirmed by chance performance behaviorally. This non-consciously retained information seems to rely on persistent neural activity in frontal and occipital cortex, and may engage further cognitive control processes during memory recognition.

Keywords: consciousness, continuous flash suppression, fMRI, subjective measure, working memory

Intuitively, we seem to have rich conscious experiences of our external and internal environment as we navigate our way through the world. However, it is commonly assumed that our conscious experience reflects but a small fraction of neural processes that occurs mainly non-consciously. It was previously assumed that non-conscious processing was simple and automatic, while conscious processing was flexible and strategic (Koch & Crick, 2001; Kouider & Dehaene, 2007). In recent years, however, there has been a shift in our understanding of non-conscious processing. We now know that non-conscious processing can occur at higher perceptual levels (for reviews see, Koch, Massimini, Boly, & Tononi, 2016; Kouider & Dehaene, 2007; Rees, Kreiman, & Koch, 2002), and can influence cognitive control functions in the frontal cortex (Lau and Passingham 2007; van Gaal et al. 2010).

A similar paradigm shift has begun regarding non-conscious memory. It was previously believed that non-consciously perceived information quickly faded, and was undetectable after 500 ms (Dehaene & Changeux, 2011; Draine & Greenwald, 1998; Greenwald, Draine, & Abrams, 1996; Mattler, 2005). However, studies have shown that non-conscious repetition priming can have effects lasting 15-20 min (Bar and Biederman 1998, 1999) and even up to 47 min (Gaillard et al., 2007). It was furthermore assumed that non-conscious (non-procedural) memory only existed in the form of priming since working memory, as well as long-term retention involving the hippocampus, were strongly associated with conscious experience (Graf and Schacter 1985; Squire et al. 1992; Baddeley and Andrade 2000; Dehaene and Naccache 2001; Tulving 2002; Baars and Franklin 2003; Baddeley 2003; Baars 2005; Dehaene and Changeux 2011; Squire and Dede 2015). There are now several studies showing hippocampi-based retention of non-consciously encoded information (Henke et al. 2003; Degonda et al. 2005; Reber et al. 2012; Duss et al. 2014).

Lately, the dominating view that working memory only pertains to conscious information has also been challenged (Soto and Silvanto 2014). Working memory is the temporary retention of information for prospective use (Baddeley & Hitch, 1974; Baddeley, 1983; Fuster, 1995, 2015), and have typically been associated with persistent neural activity in the prefrontal cortex related to the task at hand, and in posterior regions related to the memorandum (for reviews, see Eriksson, Vogel, Lansner, Bergström, & Nyberg, 2015; Fuster, 2009; Sreenivasan, Curtis, & D'Esposito, 2014). Several recent studies report that non-conscious information can be retained for durations up to 15 s, even with distractors occurring between the sample and memory probe (Bergström & Eriksson, 2014, 2015; Soto, Mäntylä, & Silvanto, 2011). This short-term retention of non-consciously perceived information has been associated with sustained BOLD signal change in the prefrontal cortex during retention (Bergström and Eriksson 2014), and activity in prefrontal cortex has been causally linked to task performance using transcranial direct current stimulation (Dutta et al. 2014). Moreover, Pan and colleagues demonstrated that non-conscious retention depended on whether or not the information was needed for prospective action, a key feature of working memory (Pan et al. 2014). However, there have been some critique against these findings (Samaha 2015; Stein et al. 2016), in that subjective ratings of conscious experience have been used in most previous studies, which can be biased towards underreporting. That is, participants may have reported "no experience" when actually having a "vague experience" on some of the trials. Objective measures of performance therefore provide more conservative evidence for stimuli being non-conscious.

To further verify the phenomenon of short-term memory of non-conscious stimuli and to investigate how the brain accomplishes such retention, we here used continuous flash suppression (CFS) to render stimuli non-conscious while participants performed a delayed match-to-sample task during fMRI scanning. Based on previous research on both conscious and non-conscious short-term memory, we expected to find sustained fMRI BOLD signal change during the delay period (Bergström and Eriksson 2014; Sreenivasan et al. 2014). Also, multivariate pattern analysis was used as a more sensitive analysis technique to complement the standard univariate approach, and to provide further information on the type of representations maintained during task performance (Lewis-Peacock and Postle 2008; Lewis-Peacock et al. 2012).

Materials and Methods

Participants

Thirty-four participants were recruited from the Umeå University campus. All participants had normal or corrected to normal vision, right eye-dominance, were right-handed, gave written informed consent, and were paid for participation. The experiment consisted of two sessions: one pre-fMRI and one fMRI session. The pre-fMRI session was used to screen for participants on whom CFS worked well (i.e., reporting target stimuli as unseen on > 80% of the suppressed trials). Four participants were excluded prior to the fMRI session for experiencing the target stimulus on too many trials. Two participants were excluded from analyses of behavioral data related to the pre-fMRI session for systematically pressing “no match” (instead of guessing) when not experiencing target stimulus, but were instructed to change behavior before the fMRI session. Four participants were excluded from the fMRI session; two for failing to follow task instructions, one for excessive head motion, and one for being a statistical outlier in both sessions (pre-fMRI d’ = -.59, > 2 SD; fMRI d’ = -.66, > 2 SD below group mean) when not experiencing target stimuli. Thus, 25 participants (M = 25 years, 17 females) were included in the pre-fMRI session, and 26 participants (M = 25 years, 17 females) in the fMRI session.

Stimuli and procedure

The pre-fMRI session consisted of 360 delayed match-to-sample trials dispersed on three presentation conditions: 120 conscious, 180 non-conscious, and 60 trials with no target, hereafter referred to as baseline trials (figure 1). The fMRI session consisted of 192 delayed match-to-sample trials (44 conscious, 108 non-conscious, and 40 baseline trials). Each trial was drawn randomly from the three conditions and began with an inter-trial-interval (ITI; 3-7 s for the pre-fMRI session, 3-9 s for the fMRI session) before the sample presentation. The sample consisted of a grey silhouette of a tool (pre-fMRI: 1.7° x 1.7°, fMRI: 1.5° x 1.5°, Gaussian blur: 1 pixel radius) that was presented in one quadrant of a computer monitor. The tool was semi-randomly selected from a set of six tools, such that the tool and quadrant was not directly repeated from one trial to the next. A mirror stereoscope was used to isolate visual input from one side of the monitor to the participants corresponding eye. The monitor was placed such that all visual input could be presented within 6° horizontally and 9.6° vertically for the pre-fMRI, and 5.4° horizontally and 8.7° vertically for the fMRI session. The sample was presented for 3 s, either to both eyes simultaneously (consciously experienced), or only to the non-dominant (left) eye while colored squares of random composition (mondrians; pre-fMRI: 4.2° x 4.2°, fMRI: 3.8° x 3.8°) where flashed (10 Hz) to the dominant eye to suppress the stimulus from conscious experience (Tsuchiya and Koch 2005). During the suppressed presentations the sample was presented at gradually stronger contrast within the initial 400 ms of the 3 s to facilitate suppression. During the baseline trials mondrians were presented to the dominant eye while an empty gray background (pre-fMRI: 4.2° x 4.2°, fMRI: 3.8° x 3.8°) was presented to the non-dominant eye. Critically, the visual experience of baseline and non-conscious trials was the same (experiencing only mondrians).

After a delay period (5 s during the pre-fMRI session, 5-15 s for the fMRI session), a memory probe was presented until the participant responded (maximum 5 s). The probe could match the sample in terms of object identity and spatial position, only object identity, only spatial position, or neither. The participants were instructed that in this memory recognition task a "match" consisted of the probe being the same object in the same spatial position (full match). If the probe contained the same object at a different spatial position (identity match), different object at the same spatial position (position match), or different object identity at a different spatial position (non-match), it should be answered with a "no match" response. If participants did not experience a target stimulus (i.e., only experienced mondrians) they were instructed to guess on the first alternative that came to mind/gut feeling (match or no match) when the probe appeared. During the pre-fMRI session there were equal proportions of "match" and "no match" trials. However, during the fMRI session there was a larger proportion of match than no match trials because we intended to focus our analyses on comparisons between hits > baseline and hits > misses. Out of the conscious trials there were 20 full match, 8 identity match, 8 position match, and 8 non-match trials (i.e., 24 "no match" trials). Out of the non-conscious trials there were 78 full match, 10 identity match, 10 position match, and 10 non-match trials.

Next the participants were prompted to make a detection response to determine if a sample stimulus had been presented at all (yes or no). If they had not perceptually experienced a sample they were to guess per the same instructions as for the memory recognition task. Lastly, they estimated their conscious experience of the stimulus on a three-point perceptual awareness scale (PAS; Sandberg, Timmermans, Overgaard, & Cleeremans, 2010). The participants were instructed and trained to use the PAS scale as follows: 1 = no perceptual experience, 2 = vague perceptual experience, and 3 = clear or almost clear perceptual experience, of the target stimulus. During the pre-fMRI session there was no time limit when responding, but during the fMRI session responses had an upper time limit of 5 s, after which the experiment automatically continued with the next response prompt or trial.

After participants had received instructions during the pre-fMRI session, they performed a practice run of the experiment with the instructor until their behavior was consistent with the instructions, after which the actual pre-fMRI session began. Participants were debriefed and asked about their behavior in relation to the instructions after both sessions.

fMRI acquisition

The fMRI data was collected with a GE 3 Tesla Discovery MR750 scanner (32-channel receive-only head coil). Each participant underwent one fMRI session with two functional runs (1230 volumes each) of scanning using a T2*-weighted gradient echo pulse sequence, echo planar imaging, field of view = 25 cm, matrix size = 96 × 96, slice thickness = 3.4 mm, 37 slices with no inter-slice skip and an ASSET acceleration factor of 2. The volumes covered the whole cerebrum and most of the cerebellum, the acquisition orientation was oblique axial and aligned with the anterior and posterior commissures, and was scanned in interleaved order with TE = 30 ms, TR = 2 s, flip angle = 78°. Between the two functional runs a high-resolution T1-weighted structural image was collected FSPGR with TE = 3.2 ms, TR = 8.2 ms, TI = 450 ms, and flip angle = 12°.

Data processing and statistical analysis

Trials with a response time (RT) of < 250 ms or > 5 s were excluded prior to statistical analyses (Ratcliff 1993). Only trials in the baseline and non-conscious presentation conditions with PAS = 1, and trials in the conscious condition with PAS = 3 were used in the statistical analyses, and will for simplicity hereby be referred to as baseline, non-conscious, and conscious trials. Signal detection theory (d′) was used to calculate performance on the delayed match-to-sample recognition task and the detection task (Macmillan and Creelman 1991). For recognition d′ the signal was defined as the object identity and its spatial position. A hit was therefore defined as a (identity and position) match between sample and probe together with a "match" response, and a false alarm (FA) as a non- or partial-match between sample and probe together with a "match" response. For the detection task, a hit was defined as the presence of a sample stimulus together with a "yes" response, and a FA as the absence of a sample stimulus (i.e., baseline trials) together with a "yes" response.

Preprocessing and univariate analyses of fMRI data. The software used for processing and analysis of fMRI data was SPM8 (Welcome Trust Centre for Neuroimaging, London, UK), run in Matlab 7.11 (Mathworks, Inc., Sherborn, MA, USA). Before preprocessing, a manual quality inspection using in-house software was done. Preprocessing was done in the following order; slice-timing correction to the first slice using Fourier phase-shift interpolation method, head-motion correction with unwarping of B0 distortions, DARTEL normalization (Ashburner 2007) using a 12-parameter affine transformation model to MNI anatomical space, and an 8 mm FWHM Gaussian smoothing. DARTEL normalization and smoothing was applied on the contrast images after intrasubject model estimation. For intrasubject modeling a General Linear Model (GLM) with restricted maximum likelihood estimation was used.

The model consisted of the following regressors of interest: Presentation conditions (conscious and non-conscious) by trial phases (sample presentation, delay, and DMS response) by PAS rating (1, 2, or 3) by signal detection category (hits, misses, false alarms, and correct rejections), and baseline by trial phases by PAS rating, and lastly the ITI. The model also included the following nuisance regressors: missed responses (because of time limit), head motion (six parameters) and physiological noise (six parameters) estimated with temporal variation in white matter and cerebral spinal fluid (Behzadi, Restom, Liau, & Liu, 2007). All regressors except for head motion and physiological noise were convolved with the "canonical" hemodynamic response function as defined in SPM8. The high-pass filter had a cut-off at 128 s, and the autocorrelation model was global AR(1). Model estimations from each individual were taken into second-level random-effects analyses (one-sample t-tests) to account for inter-individual variability. Statistical inferences were made on the whole brain with $p \leq 0.001$, uncorrected for multiple comparisons, $k \geq 20$, unless otherwise specified.

Multivariate pattern analyses of fMRI data. The fMRI data was preprocessed by correcting for slice timing and head motion as described under univariate analysis prior to being analyzed with the Princeton MVPA Toolbox. For individual feature selection we created binary masks from univariate F contrasts of conscious (hits and correct rejections) compared to baseline trials for the phase of interest (presentation, delay, or recognition response). The feature selection masks were threshold at $p \leq .0001$, uncorrected, $k = 0$ for the whole brain classifications of all phases except recognition ($p \leq .001$), which otherwise excluded almost all voxels for some individuals. The feature selection masks were selected to capture the voxels most relevant for the processing during the specific phase of interest. We thus assumed that voxels related to conscious processing are relevant for non-conscious processing. For the presentation and delay phases we also constructed spatially limited feature selection masks, defined as the conjunction between the previously described univariate F contrasts ($p \leq .001$) and occipital, temporal, parietal, or frontal lobe masks based on the WFU Pickatlas (Maldjian et al. 2003). Classification during the ITI was used as a sanity-check, and done on BOLD signal from the scan prior to sample presentation onset.

To pinpoint the trial phases of interest we used the regressors from the univariate GLM, and set different cut-off values to reflect the difference in durations (> .4 for presentation, and > .14 for recognition). Thus, we only included fMRI data from when the regressors were at their respective peaks. However, given the sluggishness of the BOLD signal, we handled the delay periods differently. We shifted the delay onset times forward four scans to avoid residual BOLD signal from the stimulus presentation, while coming closer to the peak of the delay-related BOLD signal. We also cut the delay duration short so that it always ended with the scan prior to recognition probe onset, to ensure that no BOLD signal from the probe could influence classification performance of the delay. As a result, trials with a delay duration < 10 seconds were not used in the classification of delay-related BOLD signal.

The included voxel values were then passed through a high-pass filter (128 s cut-off), replaced by z-score normalized values, and averaged over time. The analyses used a leave-k-out cross-validation procedure where k is the number of categories to be classified (i.e., two; either sample presence vs. absence or left vs. right visual field). When there were more trials of one category we randomly excluded trials from that category until there was an equal amount of trials in each category. Because the MVPA needed an equal amount of trials in each category and there were fewer baseline than non-conscious trials, we could only use a subset of all non-conscious trials. We therefore opted to use non-conscious hits and correct rejections vs. baseline trials, in case correct trials carried some additional information despite behavioral performance being at chance.

To preclude that finger-related BOLD signal confounded the classification performance during analyses of the recognition phase, we controlled for which finger/button was used to make the response (index finger for “match”, and middle finger for “non-match” responses). For example, if classifying hits and correct rejections vs. baseline trials, an equal number of hits and baseline trials with “match” responses, and the same amount of correct rejections and baseline trials with “non-match” responses, were selected.

Following Polyn, Natu, Cohen, & Norman (2005), we used a backpropagation-based neural network algorithm to train and test the BOLD signal patterns in the data, an OnOff-value was calculated as a measure of classification performance, and statistical significance was tested using a non-parametric permutation test. Each participant’s OnOff value was computed by correlating the two classifier estimates (how well the current test pattern matches each category’s characteristic/trained pattern) with the actual conditions (answer key) for all test iterations. From the resulting two-by-two correlation matrix an OnOff value was derived by subtracting the average of the off-diagonal from the average of the on-diagonal elements. An overall OnOff value was estimated by averaging across participants' OnOff values. For the non-parametric permutation (group level) test we scrambled the individual OnOff matrices. The actual overall OnOff value was then compared to the null distribution of 10 000 scrambled OnOff values to generate a (one-tailed) p-value.

Results

In the following results, all trials with PAS > 1 were removed to ensure no visual experience of the target stimulus in non-conscious (3.9% for the pre-fMRI session and < 1% for the fMRI session) and baseline conditions (pre-fMRI: 3.3%; fMRI: 0), and all trials with PAS < 3 were removed from the conscious condition (pre-fMRI: 2.5%, fMRI: 2.3%).

Behavioral performance

One-tailed t-tests were used to test whether memory performance (d') was above chance (d' > 0) for the recognition and detection tasks. For conscious trials, memory performance was above chance during both sessions for the recognition task [pre-fMRI: t(24) = 86, p < .001, M = 4.56, SE = .06; fMRI: t(25) = 89, p < .001, M = 3.85, SE = .04] and the detection task [pre-fMRI: t(23) = 33, p < .001, M = 4.32, SE = .13; fMRI: t(25) = 113, p < .001, M = 4.05, SE = .04]. For non-conscious trials, recognition [t(24) = 3.44, p = .001, M = .16, SE = .05] and detection [t(23) = 3.49, p = .001, M = .15, SE = .04] was significantly better than chance during the pre-fMRI session. Importantly, neither recognition [t(25) = .41, p = .69, M = .02, SE = .04] nor detection [t(25) = .31, p = .76, M = .02, SE = .05] was better than chance during the fMRI session. Thus, in these trials the sample to be remembered was non-conscious according to objective criteria.

Similar results were evident for response times. Paired t-tests of conscious recognition response times (ms) during the pre-fMRI session showed that hits [t(24) = -7.40, p < 0.001, M = 1236, SE = 41] and correct rejections [t(24) = -4.84, p < 0.001, M = 1505, SE = 47] were faster than baseline response times (M = 2027, SE = 113), and hits were faster than correct rejections [t(24) = -8.82, p < .001]. This was true also for the fMRI session (hits [1287+/-41] < correct rejections [1517+/-64] < baseline [2074+/-105], all *p* < .001). However, for non-conscious recognition during the pre-fMRI session, hit (2014+/-119), correct rejection (2030+/-108), and baseline (2027+/-113) response times did not differ from each other (all *p* > .49). Because non-conscious recognition performance was at chance during the fMRI session, comparing reaction times across "hits", "misses", etc., was considered less meaningful. Previous research has demonstrated different neurocognitive processes during working-memory retrieval depending on whether the memory probe matches or does not match the sample (Bledowski et al. 2012; Rahm et al. 2013; Schon et al. 2015). We therefore compared the response times of match and non-match non-conscious trials. Paired t-tests revealed that match response times [t(25) = .24, p = .81, M = 2081, SE = 90] and non-match response times [t(25) = 1.42, p = .17, M = 2119, SE = 105] were no different from baseline, or each other [t(25) = -.95, p = .35]. Thus, there were no response time differences among non-conscious trial types during the pre- or fMRI sessions.

fMRI Results

We examined the BOLD signal change related to conscious processes by contrasting conscious (hits and correct rejections) to the baseline trials separately at the three trial phases: sample presentation, delay, and recognition response. When contrasting conscious > baseline sample presentations we found widespread BOLD signal change bilaterally in the occipital, temporal, parietal, and prefrontal cortex. For the delay period there was sustained BOLD signal change bilaterally in the occipital cortex, and at memory recognition there was widespread BOLD signal change in occipital, temporal, parietal, and prefrontal cortex bilaterally. The comparison of sample-probe non-match > match probes was associated with extensive BOLD signal change lateralized in the left temporal, parietal, frontal cortex, and the right putamen/pallidum (figure 2A), while comparing match > non-match probes was associated with a higher BOLD signal change in the left occipital cortex.

Since non-conscious recognition and detection performance was at chance, we treated hits, misses, false alarms, and correct rejections together when comparing non-conscious to baseline trials at the different trial phases. There was no significant BOLD signal change related to non-conscious > baseline sample presentations, nor during the delay. However, during recognition BOLD signal increased in the right anterior insula and inferior frontal cortex (table 1). Furthermore, when comparing non-match > match trials there was significant BOLD signal change in occipital, temporal, and parietal cortex, and in the left pre-SMA, partly overlapping with the corresponding results for conscious non-match > match (figure 2A, table 1). For non-conscious sample-probe match > non-match, there was a significant signal change in the right cerebellum (table 1).

The results related to non-match trials are important because they may demonstrate memory effects that cannot be ascribed to simple repetition between sample and probe. However, the results from the non-match > match comparison may be driven by reduced BOLD signal for match trials (e.g., repetition suppression) rather than BOLD signal change specifically related to non-match trials. To verify that there were significant signal change related specifically to non-match trials we compared non-match > baseline trials, which revealed increased BOLD signal in the left inferior and middle frontal gyrus and supramarginal gyrus, partly overlapping with corresponding results based on conscious trials (figure 2B, table 1). This activity pattern did not overlap with the results from the non-match > match comparison, suggesting that the results from the latter comparison was at least in part driven by signal change related to match trials.

Multivariate pattern analyses. The recognition-related results for non-conscious trials demonstrate that some of the non-consciously presented information indeed was retained despite behavioral performance being at chance. It is therefore somewhat surprising that there was no significant signal change at least related to the non-conscious sample presentation. Multivariate pattern analysis (MVPA) is more sensitive than univariate analyses, and previous studies have found significant classification performance without significant univariate BOLD signal change during working-memory tasks (Riggall and Postle 2012; Emrich et al. 2013). We therefore used MVPA to further investigate BOLD signal change during the different trial periods. Specifically, we trained the classifier to differentiate between non-conscious (PAS = 1) and baseline (PAS = 1) trials separately for each trial phase (sample presentation, delay, recognition).

Classification accuracy was significantly better than chance (which corresponds to an OnOff value of zero) for the sample presentation period ($p < .0001$, $M = .26$, $SE = .05$) using a whole-brain feature mask that was based on conscious vs. baseline trials. To address whether there was any sample-specific information present in the brain activity pattern we tried to classify spatial position of the sample object (left or right visual field) in regions that are associated with spatial processing, by combining the whole-brain feature mask with cerebral lobes as defined in the WFU Pickatlas (i.e., the conjunction between the whole-brain feature mask and the different lobes). Classification of spatial position was significant in the parietal ($p = .01$, $M = .13$, $SE = .05$) and occipital ($p = .045$, $M = .10$, $SE = .06$), but not frontal ($p = .25$, $M = .04$, $SE = .05$) cortex.

Next we used a delay-based feature mask (created from conscious vs. baseline trials) to classify non-conscious sample presence/absence during the delay period. Classification performance was not significant (p = .11, M = .08, SE = .06) when using the whole-brain feature mask. However, subsequent analyses showed that classification was successful in frontal (p = .008, M = .20, SE = .07), but not parietal (p = .45, M = .01, SE = .09), temporal (p = .07, M = .08, SE = .06) or occipital (p = .25, M = .04, SE = .06) cortex. Given the sluggishness of the BOLD signal, it is possible that the algorithm was classifying residual signal from the sample presentation rather than signal change related to the delay period. To control for such "spillover" we proceeded to use the delay-based frontal mask (i.e., the same voxels that was successfully classified during the delay) to classify sample presence vs. absence during the presentation phase. Critically, classification was at chance (p = .32, M = .03, SE = .07), which excludes the possibility that classification performance for the delay period was due to residual BOLD signal from the sample presentation.

To address whether sample-specific information was retained during the delay, we tried to decode spatial position. Classification was successful in the occipital (p = .004, M = .16, SE = .06), but failed in parietal (p = .64, M = -.02, SE = .07), and frontal (p = .15, M = .08, SE = .07) cortex. To control for residual BOLD signal we used the delay-based occipital mask to decode spatial position during the sample presentation. Classification performance was at trend (p = .07, M = .08, SE = .05), which suggests that the relatively high classification performance for the delay phase cannot purely be attributed to residual BOLD signal from the sample presentation.

Using the recognition-based feature mask (created from conscious vs. baseline trials) classification performance was better than chance for non-conscious sample presence vs. absence during recognition (p < .0001, M = .42, SE = .07), which corroborate our univariate findings and demonstrate that the non-conscious information is retained until recognition despite chance-level performance. Importantly, classification performance was also better than chance when using only non-match trials ($p < .0001$, M = .50, SE = .07).

Control Analyses. It is conceivable that participants accidentally pressed PAS = 1 despite consciously experiencing the target stimulus on some trials, and that such mislabeling could explain the effects attributed to non-conscious processing. We therefore did control analyses to address the potential effect of mislabeling trials in the univariate and multivariate analyses. We assumed that the amount of mislabeling during suppressed trials could be approximated by the number of mislabeled trials in the non-suppressed trials (i.e., trials where the sample stimulus was clearly visible but participants pressed "PAS = 1"; 2.3%, see above). To control for mislabeling effects in the fMRI data we divided the baseline trials in two bins. We then contaminated one bin of baseline trials with the corresponding number of potentially mislabeled trials (rounding up, corresponding to 1 trial, or 5%), using conscious correct rejection trial (because correct rejections, i.e., non-match trials, had the most pronounced BOLD signal change) for each participant. The univariate comparison of contaminated > pure baseline trials did not reveal any significant BOLD signal change at the previously set threshold, and lowering the threshold to $p \leq .01$ uncorrected, $k = 0$, reveled only a few smaller clusters, mostly in white matter, and none that overlapped with previous results. To control for effects of mislabeled trials in the MVPAs we trained the algorithms to differentiate between the contaminated and pure baseline bins. Classification performance was at chance during the sample presentation ($p = .37$, M = .03, SE = .08), delay (frontal cortex, $p = .20$, M = .06, SE = .07; occipital cortex, $p = .68$, M = -.04, SE = .08), and recognition phase ($p = .46$, M = .01, SE = .09).

Discussion

Consistent with recent research, the current fMRI results demonstrate short-term memory effects of information presented non-consciously. Here, the effects were most pronounced during the memory recognition (test) phase, but using MVPA there was also evidence for BOLD signal change during the sample presentation and during the delay phase. Critically, behavioral performance during scanning was at chance level, providing strong support for the non-conscious nature of the sample presentation.

Sustained activity during the delay phase

Sustained activity during short memory delays is often considered a characteristic feature of working memory, as it represents a likely neural mechanism for short-term retention of task-relevant information. During the delay phase the MVPA successfully classified presence vs. absence of the non-consciously presented sample based on BOLD signal change in the frontal cortex, and its spatial position (left vs. right) in the occipital cortex. We cannot infer the exact function of the frontal BOLD signal in relation to retention, except that it is likely not related to spatial information. Previous working-memory literature commonly link the prefrontal cortex to task-specific information, while item-specific information (i.e., the memorandum) is suggested to be retained in posterior regions (Eriksson, Vogel, Lansner, Bergström, & Nyberg, 2015; Fuster, 2009; Sreenivasan, Curtis, & D’Esposito, 2014). Surprisingly, it was only possible to decode spatial position and not presence vs. absence in the occipital cortex during the delay. Possibly, the BOLD signal during the delay phase was heterogeneous relative present/absent categorization when a sample had been presented, because the signal would toggle between representing left and right samples within the “sample present” category (i.e., different signals within the same category), to a degree that was not apparent during the sample presentation itself. It is also noteworthy that it was not possible to decode spatial position based on BOLD signal in the parietal cortex, even though classification was significant during the sample-presentation phase.

Speculatively, only lower-level visuospatial information was actively maintained during the delay.

Considering how weak the current BOLD signal was during the delay phase of non-conscious trials (only detectable using MVPA), it is unclear how “sustained” the corresponding neural activity in fact was. Specifically, it may reflect intermittent rather than persistent neural activity (Lundqvist et al. 2016), and/or metabolically demanding synaptic events that may not be reflected in increased neural spiking (Goense and Logothetis 2008), but may still reflect short-term retention of mnemonic information (Shafi et al. 2007; Mongillo et al. 2008). Moreover, we were here able to demonstrate only crude evidence that the sustained information was specifically related to the sample (left/right visual field). These results extend previous findings regarding the type of information that is maintained during the delay phase following non-conscious presentation of a memory sample (King et al. 2016). Specifically, King et al. demonstrated successful decoding of target presence/absence based on MEG signals during the delay phase, but failed to decode more specific information. Based on behavioral measures we have previously demonstrated that the conjunction of spatial position and object identity can be maintained during a delayed match-to-sample task almost identical to the current task (Bergström and Eriksson 2015). We here failed to replicate these findings and speculate that the extent of encoding during continuous flash suppression may have differed across participants (previous vs. current experiment) and across experimental setups (pre-fMRI vs. fMRI session), such that stronger suppression may lead to a weaker sample representation. Further research is needed to more clearly determine the kind of information that can be maintained in memory after a non-conscious stimulus presentation.

Memory recognition

During the memory recognition phase, there was significant BOLD signal change in the right anterior insula and the right inferior frontal cortex when comparing non-conscious to baseline trials. In line with recent research that specifically address neurocognitive processes during working-memory retrieval (Bledowski et al. 2006, 2012; Nee and Jonides 2008; Rahm et al. 2013), there was also significant activity differences when comparing memory probes that did vs. did not match the sample, both for conscious and non-conscious trials. Critically, there was significant BOLD signal change also when comparing only non-matching probes with baseline probes both for the univariate analysis and the MVPA – results that cannot be explained in terms of simple repetition suppression from sample to probe, but rather indicates more complex mnemonic processing that is potentially similar to working-memory retrieval operations during more traditional (i.e., conscious) working-memory tasks. Specifically, previous research has demonstrated increased BOLD signal in left lateral prefrontal cortex during the test phase when demands on sample-probe comparison processes are higher, for example when the memory probe does not match the item currently in the focus of attention (Nee and Jonides 2008; Bledowski et al. 2012; Rahm et al. 2013). Here, BOLD signal change in left middle frontal and supramarginal gyrus increased during non-match trials both when the sample was conscious and non-conscious (figure 2B).

Several frontal and parietal regions have previously been associated with different cognitive control processes relevant for memory probe recognition, including attentional deployment (Nee and Jonides 2008), sample-probe similarity assessment (Bledowski et al. 2012), and decision making (Rahm et al. 2013). Based on the current activity pattern during the memory recognition phase, we speculate that cognitive control processes are also engaged during non-conscious memory recognition. In line with this proposal, several previous studies have reported activation of cognitive control processes related to non-conscious stimuli (Lau and Passingham

2007; van Gaal et al. 2010; Charles et al. 2013; Reuss et al. 2015). The specific cognitive roles played by the currently activated regions remain unclear, not least because behavioral performance was at chance level.

Consciousness and working memory

Historically, working memory and conscious experience have been tightly linked (Soto and Silvanto 2014). More recently, neuroimaging studies have found overlapping neural correlates of working memory and conscious perception in the prefrontal and parietal cortex (Rees et al. 2002; Naghavi and Nyberg 2005) (but see no-report paradigms; Frassle et al. 2014; Tsuchiya et al. 2015; Koch et al. 2016). Based on such overlapping findings, prominent models of conscious experience have asserted that prefrontal and parietal activity, and by extension working memory, plays an important role in conscious experience. According to the Global Neuronal Workspace (GNW) model, non-conscious information is processed locally in specialized modules and is relatively short-lived (< 500 ms). For information to become consciously experienced and maintained in working memory, it needs to be globally broadcast via the frontoparietal network, and is thereby available to many brain regions (Dehaene & Naccache, 2001; Dehaene & Changeux, 2011). However, our findings seem to contradict some of the GNW model's assumptions. Firstly, non-conscious information can be retained for several seconds. Secondly, the information is retained by persistent fronto-occipital activity. Thirdly, conscious experience does not seem to be necessary for working memory. The current findings are thus consistent with the commonly held notion that most of our cerebral processing are parallel and non-conscious, and that we only consciously experience a small fraction of it all, which seems to suggest that non-conscious processing might be very global.

It has recently been demonstrated that unattended information in a short-term memory task can fail to be decoded with MVPA during delay periods, but can nevertheless be used for solving the task and can be "revived" by re-directing attention to the relevant representation (Lewis-Peacock et al. 2012; LaRocque et al. 2013). Such findings are in line with state-based models of working memory, where memory representations are in different "states of access" depending on attentional deployment (Fuster 1995; Oberauer 2002; McElree 2006; Cowan 2008; Jonides et al. 2008). Our current findings suggest that sample representations can be in a heightened state of access, reflected here as significant BOLD signal change during the delay period, even though it has never been "inside" the focus of attention.

Critique of non-conscious short-term memory

Recent critique against findings related to non-conscious working memory have pointed out that subjective measures of awareness, which has been used in most previous research on this topic, might be biased towards underreporting (Samaha 2015; Stein et al. 2016). Results may therefore be driven by information that was in fact conscious, despite subjective reports indicating no conscious experience. Objective measures are more conservative but increase the risk of false negative findings, and the two approaches may be seen as complementary (Seth et al. 2008). Stein et al. (2016) argued that even if the subjective measure indeed were to be bias-free, results could still be explained by participants indeed having a non-conscious perception, but this is then transformed into a conscious "guess" that can be maintained in (conscious) working memory. By contrast, several aspects of the current findings support the notion of non-conscious short-term memory. Firstly, the participants' recognition and detection performance was at chance during the fMRI session, meaning that the sample presentation was non-conscious according to objective criteria. This finding is consistent with results reported by Pan et al. (2014), where memory performance related to non-consciously presented faces was at chance level while indirect measures (time to breaking suppression) was significantly altered by the sample presentation. Secondly, we performed control analyses with regard to accidental mislabeling of trials. These control analyses (univariate and multivariate fMRI) showed that accidental mislabeling at frequencies similar to mislabeling of conscious trials could not by itself drive the effects seen for non-conscious trials. Thirdly, the participants were instructed to wait until the probe appeared before making their guesses, and when debriefed about any particular strategies during the non-conscious trials, they report none. If, as Stein et al. (2016) suggests, participants guess the object's identity and position and hold that conscious guess in working memory, conscious and non-conscious reaction times should not differ. However, reaction times for conscious trials were significantly faster than non-conscious trials, which in turn did not differ from baseline trials. This suggests that participants had not already made their guesses before the probe appeared. Taken together, the current findings provide strong support for the notion of non-conscious short-term memory.

Conclusions

In conclusion, we found neural evidence for maintenance of non-consciously presented information during several seconds and engagement of brain regions associated with cognitive control during memory recognition. The maintenance of sample-unspecific information in the frontal cortex and sample-specific information in the occipital cortex is consistent with current conceptions of how information is maintained in working memory. These findings imply that working-memory models need to accommodate a representational state where information can be maintained without ever being inside the focus of attention. Furthermore, the findings contradict some of the assumptions of the global neuronal workspace model, namely that non-conscious processing cannot be global, maintained in short-term memory, or engage cognitive control processes.

Funding

This work was supported by grants from the Swedish Research Council, the medical faculty and the collecting foundation for medical research at Umeå University, to JE. The research leading to these results has received funding from the European Union Seventh Framework Program (FP7/2007-2013) under grant agreement no. 604102 (Human Brain Project), to Lars Nyberg.

Acknowledgements

We thank Lars Nyberg for input on the experimental design, Göran Westling for building the stereoscope, and Anders Lundqvist for statistical advice.

References

Ashburner J. 2007. A fast diffeomorphic image registration algorithm. Neuroimage. 38:95–113.

Baars BJ. 2005. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. Prog Brain Res. 150:45–53.

Baars BJ, Franklin S. 2003. How conscious experience and working memory interact. Trends Cogn Sci. 7:166–172.

Baddeley a D, Andrade J. 2000. Working memory and the vividness of imagery. J Exp Psychol Gen. 129:126–145.

Baddeley A. 2003. Working memory: looking back and looking forward. Nat Rev Neurosci. 4:829–839.

Baddeley AD. 1983. Working memory. Philos Trans R Soc London.

Baddeley AD, Hitch GJ. 1974. Working Memory. In: Bower GA, editor. Recent advances in learning and motivation. New York: Academic Press. p. 47–89.

Bar M, Biederman I. 1998. Subliminal Visual Priming. Psychol Sci. 9:464–468.

Bar M, Biederman I. 1999. Localizing the cortical region mediating visual awareness of object identity. Proc Natl Acad Sci U S A. 96:1790–1793.

Behzadi Y, Restom K, Liau J, Liu TT. 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage. 37:90–101.

Bergström F, Eriksson J. 2014. Maintenance of non-consciously presented information engages the prefrontal cortex. Front Hum Neurosci. 8:938: 1–10.

Bergström F, Eriksson J. 2015. The conjunction of non-consciously perceived object identity and spatial position can be retained during a visual short-term memory task. Front Psychol. 6:1470:1–9.

Bledowski C, Cohen Kadosh K, Wibral M, Rahm B, Bittner R a, Hoechstetter K, Scherg M, Maurer K, Goebel R, Linden DEJ. 2006. Mental chronometry of working memory retrieval: a combined functional magnetic resonance imaging and event-related potentials approach. J Neurosci. 26:821–829.

Bledowski C, Kaiser J, Wibral M, Yildiz-Erzberger K, Rahm B. 2012. Separable neural bases for subprocesses of recognition in working memory. Cereb Cortex. 22:1950–1958.

Charles L, Van Opstal F, Marti S, Dehaene S. 2013. Distinct brain mechanisms for conscious versus subliminal error detection. Neuroimage. 73:80–94.

Cowan N. 2008. What are the differences between long-term, short-term, and working memory? Prog Brain Res. 169:323–338.

Degonda N, Mondadori CR a, Bosshardt S, Schmidt CF, Boesiger P, Nitsch RM, Hock C, Henke K. 2005. Implicit associative learning engages the hippocampus and interacts with explicit associative learning. Neuron. 46:505–520.

Dehaene S, Changeux J-P. 2011. Experimental and theoretical approaches to conscious processing. Neuron. 70:200–227.

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

Draine SC, Greenwald a G. 1998. Replicable unconscious semantic priming. J Exp Psychol Gen. 127:286–303.

Duss SB, Reber TP, Hänggi J, Schwab S, Wiest R, Müri RM, Brugger P, Gutbrod K, Henke K. 2014. Unconscious relational encoding depends on hippocampus. Brain. 137:3355–3370.

Dutta A, Shah K, Silvanto J, Soto D. 2014. Neural basis of non-conscious visual working memory. Neuroimage.

Emrich SM, Riggall a. C, LaRocque JJ, Postle BR. 2013. Distributed Patterns of Activity in Sensory Cortex Reflect the Precision of Multiple Items Maintained in Visual Short-Term Memory. J Neurosci. 33:6516–6523.

Eriksson J, Vogel EK, Lansner A, Bergström F, Nyberg L. 2015. Neurocognitive Architecture of Working Memory. Neuron. 88:33–46.

Frässle S, Sommer J, Jansen A, Naber M, Einhauser W. 2014. Binocular Rivalry: Frontal Activity Relates to Introspection and Action But Not to Perception. J Neurosci. 34:1738–1747.

Fuster JM. 1995. Memory in the Cerebral Cortex - An Empirical Approach to Neural Networks in the Human and Nonhuman Primate. Cambridge: MA: MIT Press.

Fuster JM. 2009. Cortex and memory: emergence of a new paradigm. J Cogn Neurosci. 21:2047–2072.

Fuster JM. 2015. The Prefrontal Cortex. 5th Editio. ed. London: Academic Press.

Gaillard R, Cohen L, Adam C, Clemenceau S, Hasboun D, Baulac M, Willer J-C, Dehaene S, Naccache L. 2007. Subliminal words durably affect neuronal activity. Neuroreport. 18:1527–1531.

Goense JBM, Logothetis NK. 2008. Neurophysiology of the BOLD fMRI signal in awake monkeys. Curr Biol. 18:631–640.

Graf P, Schacter DL. 1985. Implicit and explicit memory for new associations in normal and amnesic subjects. J Exp Psychol Learn Mem Cogn. 11:501–518.

Greenwald AG, Draine SC, Abrams RL. 1996. Three cognitive markers of unconscious semantic activation. Science (80-). 273:1699.

Henke K, Treyer V, Nagy ET, Kneifel S, Dürsteler M, Nitsch RM, Buck A. 2003. Active hippocampus during nonconscious memories. Conscious Cogn. 12:31–48.

Jonides J, Lewis RL, Nee DE, Lustig C a, Berman MG, Moore KS. 2008. The mind and brain of short-term memory. Annu Rev Psychol. 59:193–224.

King J-R, Pescetelli N, Dehaene S. 2016. Selective maintenance mechanisms of seen and unseen sensory features in the human brain. bioRxiv.

Koch C, Crick F. 2001. The zombie within. Nature. 411:893.

Koch C, Massimini M, Boly M, Tononi G. 2016. The neural correlates of consciousness: progress and problems. Nat Rev Neurosci. in press:307–321.

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. 362:857–875.

LaRocque JJ, Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. 2013. Decoding attended information in short-term memory: an EEG study. J Cogn Neurosci. 25:127–142.

Lau HC, Passingham RE. 2007. Unconscious activation of the cognitive control system in the human prefrontal cortex. J Neurosci. 27:5805–5811.

Lewis-Peacock JA, Drysdale AT, Oberauer K, Postle BR. 2012. Neural evidence for a distinction between short-term memory and the focus of attention. J Cogn Neurosci. 24:61–79.

Lewis-Peacock JA, Postle BR. 2008. Temporary activation of long-term memory supports working memory. J Neurosci. 28:8765–8771.

Lundqvist M, Rose J, Herman P, Brincat SL, Buschman TJ, Miller EK. 2016. Gamma and Beta

Bursts Underlie Working Memory. Neuron. 90:152–164.

Macmillan NA, Creelman CD. 1991. Detection Theory: A User's Guide. New York: Cambridge University Press.

Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage. 19:1233–1239.

Mattler U. 2005. Inhibition and decay of motor and nonmotor priming. Percept Psychophys. 67:285–300.

McElree B. 2006. Accessing Recent Events. Psychol Learn Motiv - Adv Res Theory. 46:155–200.

Mongillo G, Barak O, Tsodyks M. 2008. Synaptic theory of working memory. Science. 319:1543–1546.

Naghavi HR, Nyberg L. 2005. Common fronto-parietal activity in attention, memory, and consciousness: shared demands on integration? Conscious Cogn. 14:390–425.

Nee DE, Jonides J. 2008. Neural correlates of access to short-term memory. Proc Natl Acad Sci U S A. 105:14228–14233.

Oberauer K. 2002. Access to information in working memory: exploring the focus of attention. J Exp Psychol Learn Mem Cogn. 28:411–421.

Pan Y, Lin B, Zhao Y, Soto D. 2014. Working memory biasing of visual perception without awareness. Atten Percept Psychophys. 76:2051–2062.

Polyn SM, Natu VS, Cohen JD, Norman KA. 2005. Category-specific cortical activity precedes retrieval during memory search. Science (80-). 310:1963–1966.

Rahm B, Kaiser J, Unterrainer JM, Simon J, Bledowski C. 2013. fMRI characterization of visual working memory recognition. Neuroimage. 90:413–422.

Ratcliff R. 1993. Methods for dealing with reaction time outliers. Psychol Bull. 114:510–532.

Reber TP, Luechinger R, Boesiger P, Henke K. 2012. Unconscious Relational Inference Recruits the Hippocampus. J Neurosci. 32:6138–6148.

Rees G, Kreiman G, Koch C. 2002. Neural correlates of consciousness in humans. Nat Rev Neurosci. 3:261–270.

Reuss H, Kiesel A, Kunde W. 2015. Adjustments of response speed and accuracy to unconscious cues. Cognition. 134:57–62.

Riggall a. C, Postle BR. 2012. The Relationship between Working Memory Storage and Elevated Activity as Measured with Functional Magnetic Resonance Imaging. J Neurosci. 32:12990–12998.

Samaha J. 2015. How best to study the function of consciousness ? 6:1–3.

Sandberg K, Timmermans B, Overgaard M, Cleeremans A. 2010. Measuring consciousness: is one measure better than the other? Conscious Cogn. 19:1069–1078.

Schon K, Newmark RE, Ross RS, Stern CE. 2015. A Working Memory Buffer in Parahippocampal Regions: Evidence from a Load Effect during the Delay Period. Cereb Cortex.

Seth AK, Dienes Z, Cleeremans A, Overgaard M, Pessoa L. 2008. Measuring consciousness: relating behavioural and neurophysiological approaches. Trends Cogn Sci. 12:314–321.

Shafi M, Zhou Y, Quintana J, Chow C, Fuster J, Bodner M. 2007. Variability in neuronal activity in primate cortex during working memory tasks. Neuroscience. 146:1082–1108.

Soto D, Mäntylä T, Silvanto J. 2011. Working memory without consciousness. Curr Biol. 21:R912–R913.

Soto D, Silvanto J. 2014. Reappraising the relationship between working memory and conscious awareness. Trends Cogn Sci. 18:520–525.

Squire LR, Dede AJO. 2015. Conscious and Unconscious Memory Systems. CSH Perspect Biol.

Squire LR, Kosslyn S, Zola-Morgan S, Haist F, Musen G. 1992. Memory and the Hippocampus: A Synthesis From Findings With Rats, Monkeys, and Humans. Psychol Rev. 99:195–231.

Sreenivasan KK, Curtis CE, D’Esposito M. 2014. Revisiting the role of persistent neural activity during working memory. Trends Cogn Sci. 1–8.

Stein T, Kaiser D, Hesselmann G. 2016. Can working memory be non-conscious? Neurosci Conscious. 2016:niv011.

Tsuchiya N, Koch C. 2005. Continuous flash suppression reduces negative afterimages. Nat Neurosci. 8:1096–1101.

Tsuchiya N, Wilke M, Frässle S, Lamme VAF. 2015. No-Report Paradigms : Extracting the True Neural Correlates of Consciousness. Trends Cogn Sci. 19:757–770.

Tulving E. 2002. Episodic memory: from mind to brain. Annu Rev Psychol. 53:1–25.

van Gaal S, Ridderinkhof KR, Scholte HS, Lamme V a F. 2010. Unconscious activation of the prefrontal no-go network. J Neurosci. 30:4143–4150.

Table 1. BOLD signal change during non-conscious memory recognition.

Region	**Left/Right**	**Peak t-value**	**XYZ**	**Cluster size**
		NON-CONSCIOUS PROBE > BASELINE PROBE		
Inferior frontal gyrus	R	4.36	[44 14 12]	31
Anterior insula	R	4.09	[32 24 6]	32
		PROBE MATCH > PROBE NON-MATCH		
Cerebellum	R	4.44	[46 -50 -50]	31
		PROBE NON-MATCH > PROBE MATCH		
Superior frontal gyrus/pre-SMA **(i)**	L	5.36	[-10 22 54]	189
Superior frontal gyrus	L	4.64	[-8 54 30]	325
Insula	L	4.65	[-40 -6 10]	74
Precentral gyrus	R	4.74	[28 -14 62]	257
	L	4.53	[-32 -14 64]	229
Postcentral gyrus	L	4.39	[-50 -4 40]	85
Superior parietal lobule	L	3.79	[-38 -44 64]	29
Inferior parietal lobule	L	4.90	[-60 -26 38]	458
Middle temporal gyrus **(ii)**	L	5.24	[-54 -28 -4]	240
	L	5.16	[-48 -10 -42]	248
	L	3.84	[-56 -6 -22]	36
	R	4.63	[42 -38 -2]	78
Temporal pole	L	4.11	[-44 20 -18]	44
Calcarine	L/R	4.04	[12 -82 16]	252
Lingual gyrus	L/R	4.96	[12 -70 -2]	690
		PROBE NON-MATCH > BASELINE PROBE		
Middle frontal gyrus **(iii)**	L	3.73	[-36 38 40]	22
Inferior frontal gyrus	L	4.40	[-56 8 4]	45
SMA	L	4.15	[-10 2 68]	30
Supramarginal gyrus **(iv)**	L	4.03	[-60 -28 40]	54
Cuneus	R	3.83	[8 -84 18]	53

Note: (i) and (ii) partly overlapped with conscious probe non-match > match, while (iii) and (iv) partly overlapped with conscious probe non-match > baseline (see also figure 2).

Figure captions

Figure 1. Trial procedures.

Depending on the presentation condition, two identical target samples (tools), a sample and mondrians, or an empty background and mondrians, were presented to the left and right eye respectively. The object identity and spatial position of the sample was then to be retained for a 5 s (pre-fMRI session) or variable 5-15 s (fMRI session) delay period, until a probe prompted the participants to respond whether or not the probe's identity and position matched the previously presented sample. Next, participants responded whether or not a sample had been present. Finally, the participants gave an estimate of their perceptual experience of the sample. PAS = perceptual awareness scale, (i) probe identity and position matches sample, (ii) probe identity matches sample, (iii) probe position matches sample, (iv) probe does not match sample.

Figure 2. BOLD signal change during memory probe recognition.

A. BOLD signal change during conscious (hot colors) and non-conscious (cool colors) trials for memory probe recognition when the probe did not match the sample compared with sample-probe match (non-match > match). Overlap (purple) is evident in medial frontal (upper) and middle temporal (lower) regions. B. BOLD signal change during conscious (hot) and non-conscious (cool) trials for non-matching probes compared with baseline trials. Overlap (purple) is evident in middle frontal and supramarginal gyrus.

Figure 1.

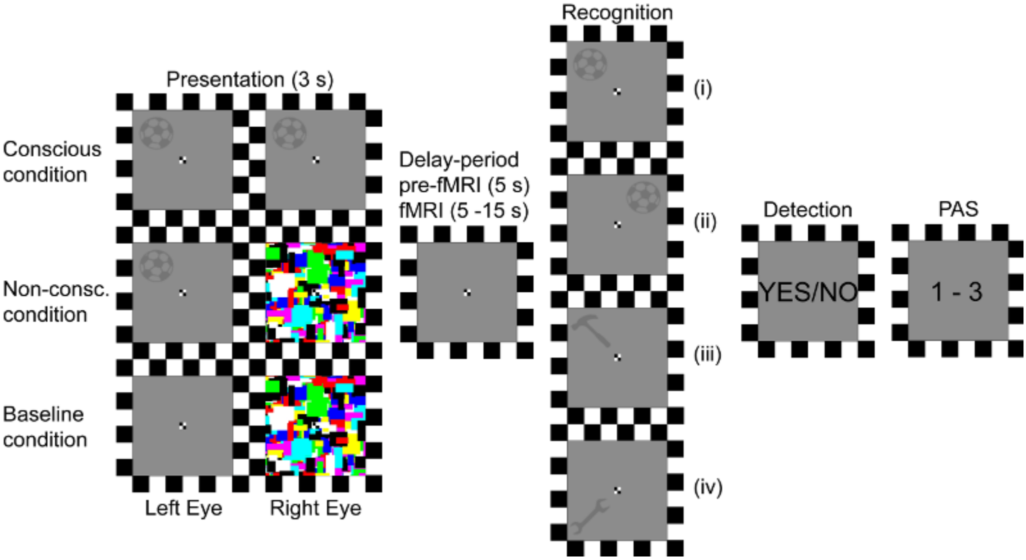

Figure 2.