Low-level memory mechanisms in vision

An fMRI-study

Dag Alnæs

Main thesis - Department of psychology

UNIVERSITY of OSLO

30.04.2010

2010

Low-level memory mechanisms in vision - an fMRI-study

Dag Alnæs

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Abstract

BACKGROUND: V1 is known to contain neurons tuned to visual low-level features such as spatial frequencies and orientation. Psychophysical studies using delayed discrimination experiments have shown that these features can be retained in memory with high fidelity. According to the sensory recruitment model of working memory, visual working memory recruits the perceptual areas involved in discriminating the features that is to be retained. If this is the case, V1-neurons tuned to spatial frequencies should be recruited when information along this dimension is retained in memory. We tested this hypothesis taking advantage of the memory masking effect.

METHOD: We used fMRI to measure BOLD responses while participants performed a delayed discrimination task for spatial frequencies. While performing the discrimination task, the participants had to retain spatial frequency information about an irrelevant masker stimulus. The mask differentially interfered with discrimination accuracy on the main task, and this memory masking effect was used to probe early visual areas for differential BOLD modulation related to the masking effect. In one experiment we presented the mask and the sample stimulus to be remembered in the same retinotopic position, and in a second experiment the mask and sample were spatially separated, allowing us to investigate the spatial extent of the low-level memory representation. Early visual areas were identified using a retinotopic mapping procedure, and ROIs retinotopically coding stimuli positions were defined for each visual area in an independent localizer session.

RESULTS: When the mask and sample stimulus were presented to the same retinotopic position, the mask impaired discrimination accuracy when it differed in spatial frequency from the sample stimulus. This memory masking effect was observed as a reduced BOLD response in V1. When spatially separating the mask and sample stimulus, we found no significant decrease in BOLD activation in V1.

CONCLUSIONS: Our results indicate that neurons in V1 involved in the perceptual coding of spatial frequencies are recruited during memory of the same information, in accordance with the sensory recruitment model of visual working memory. The memory masking effect is proposed to result from cross-channel inhibition, and is a local process in the retinotopically organized visual cortex.

Contents

1 Introduction

1.1 Visual working memory

Visual working memory is a short term memory system that keeps a limited amount of visual information online, enabling manipulation of these representations as well as action guiding, and provides a bridge between perception and long term memory (Baddeley, 2003). One approach to the study of visual working memory is to assess the number of objects that participants are able to keep in memory; this approach to working memory can be described by the storehouse metaphor; how much information is retained (Magnussen, 2009). The memory of a visual object is a complex integrated representation, comprised by different basic visual features such size, shape and color, and the integration of features and manipulations of these integrated representations are processes that demands attentional resources (Wheeler & Treisman, 2002). Since the system taxes limited attentional processes, the storage capacity of visual WM is limited, typically to three or four objects (Baddeley, 2003). This limited capacity is effectively demonstrated by the phenomenon known as change blindness, which can be induced experimentally by flashing alternating images, separated by a brief blank display, of the same visual scene, but with one element changing in one of the images, e.g. the color of an object. People often fail to detect these changes which, when eventually recognized, are so evident that they become impossible to ignore (Simons & Rensink, 2005).

Another approach to the study of visual working memory is to assess the fidelity of perceptual memory representations, or how much detail is retained about a remembered stimulus along basic visual dimensions or attributes (Magnussen, 2009). The visual attributes typically studied; spatial frequency, orientation, motion and color, are those thought to be the perceptual building blocks of the integrated meaningful visual percepts of objects and scenes (De Valois & De Valois, 1990; Pasternak, Bisley, & Valkins, 2003). The memory processes studied are thus on a less abstract stage of representation in terms of the level of integration of information, and more related to the early perceptual processes in the visual processing stream (Magnussen, 2000). Fidelity of memory representations along such basic features has typically been investigated using delayed discrimination experiments. Participants are presented with a sample stimulus, followed by an inter stimulus interval (ISI), before a test

stimulus with a higher or lower value along the studied dimension is presented. The participants then decide which of the stimuli had the higher value. The difference between sample and test is manipulated along a specified dimension, and by measuring delayed discrimination thresholds as the difference is manipulated, the fidelity of the memory representation is assessed. Decay of the representation can be estimated by measuring changes in delayed discrimination threshold at different ISIs compared to simultaneous discrimination. Studies have shown that details for low-level features is extremely well retained in memory, e.g. discrimination of spatial frequencies based on memory falls in the hyperacuity range, over ISIs as long as 30 seconds (Magnussen, 2000). This capacity to retain high fidelity representations points to different underlying systems than the limited memory for visual objects, and recent theories suggest that working memory for low-level features recruits the same neural circuits that perceptually codes for the retained feature. The purpose of the current study is to investigate this hypothesis of a dual function of the perceptual system. According to the sensory recruitment model of working memory, visual areas perceptually coding for the information that is to be remembered are also recruited in retaining this information in memory.

1.2 Sensory recruitment model of working memory

Our visual percepts are created from a series of samples of the visual field, as the eyes moves and fixates at different aspects of the visual field (Irwin, 1996). To be able to construct a coherent percept across the gaps of time between saccades, some form of short term retention of visual information is needed. An important question in the field of visual neuroscience is how the brain processes and stores perceptual information, and whether neural circuits and visual areas that are involved in perceptual coding are functionally and/or anatomically distinct from the neural networks that mediate sensory working memory. The study of working memory is already a vast, and still growing field, and much of the emphasis in studies has been on the involvement of prefrontal areas in understanding how information is retained during delays when a stimulus is absent (Baddeley, 2003; D'Esposito, Postle, & Rypma, 2000). The prefrontal cortex (PFC) has been shown to be important also in visual working memory; however its role concerning the maintenance of visual information is debated. A traditional understanding has been that different sub-systems in the PFC acts as specialized memory buffers keeping information on-line (Constantinidis, Franowicz, &

Goldman-Rakic, 2001; Goldman-Rakic & Leung, 2002). Support for this view comes from imaging studies which show that sustained activation during memory intervals in PFC can be distinguished functionally based on the strength of activations and specificity for the maintenance of different types of information (Goldman-Rakic & Leung, 2002; Haxby, Petit, Ungerleider, & Courtney, 2000), and from lesion and single-cell recording studies in non-human primates showing that different neurons and regions in PFC are tuned to, and show sustained memory related activity in response to, specific types of information (Goldman-Rakic, 1995; Levy & Goldman-Rakic, 1999). However, working memory tasks are associated with activations in multiple regions of the cortex, and sustained memory related activity during working memory is a distributed process which is not restricted to the prefrontal areas (Haxby et al., 2000; Curtis & D'Esposito, 2003; D'Esposito, 2007). There is an emerging view that visual working memory also involves the recruitment of visual perceptual areas, specifically those that perceptually code the information to be remembered (Awh et al., 1999; Harrison & Tong, 2009; Pasternak & Greenlee, 2005; Postle, 2006; Super, Spekreijse, & Lamme, 2001). Supporting this idea, based on a review of data from imaging studies on working memory for visual objects, Ranganath (2006) suggests a model in which top-down signals from prefrontal areas activate object representations in the inferior temporal (IT) cortex, which is thought to be the final stage of the ventral visual processing stream (L. G. Ungerleider & Mishkin, 1982), containing neurons showing visual responses selective for object categories (Nakamura, Matsumoto, Mikami, & Kubota, 1994). Studies on working memory for faces find sustained activations in both PFC and a subregion of the IT cortex known as the fusiform face area (FFA), which contains neurons with selective visual responses to faces (Kanwisher, McDermott, & Chun, 1997), when facial information is maintained during delay periods (Druzgal & D'esposito, 2003); and in a subregion known as the parahippocampal place area (PPA) when information about places, such as buildings or houses, is retained during delays (Ranganath, Cohen, Dam, & D'Esposito, 2004).

An interesting parallel to sensory recruitment in working memory is the research on mental visual imagery. These studies differs from those on working memory in the tasks employed, however parallels them in that imagery tasks often involves accessing perceptual information from memory (Kosslyn, Ganis, & Thompson, 2001). Mental imagery of faces (Ishai, Haxby, & Ungerleider, 2002) and places (O'Craven & Kanwisher, 2000) has been shown to selectively activate FFA and PPA. Mental imagery of more basic visual information also activates visual areas perceptually coding for the mentally generated information; fMRI-

studies have shown that mental imagery of retinotopic mapping stimuli can produce consistent retinotopic maps of early visual areas, including V1 (Klein et al., 2004; Slotnick, Thompson, & Kosslyn, 2005). It has also been shown that mentally generated lines can induce a perceptual phenomenon known as the tilt-after-effect (TAE) (Mohr, Linder, Linden, Kaiser, & Sireteanu, 2009). This is induced perceptually by prolonged exposure to visual patterns with a given orientation, resulting in a reduced neural response to subsequent exposure to similar orientations, which make them appear as tilted in the opposite direction. Mohr and colleagues showed that orientation-selective neurons in the visual cortex were recruited also during mentally induced TAE. Even though the neural underpinnings of imagery and perception might not be completely overlapping (some researchers have argued that deficits in mental imagery and perceptual processes can occur independently (Bartolomeo, 2002; Moro, Berlucchi, Lerch, Tomaiuolo, & Aglioti, 2008)), much evidence point to a revival of representations at a perceptual level in the visual cortex during such tasks. Interestingly, mental imagery of motion based on rules rather than remembered stimuli was found to deactivate early visual areas V1, V2 and V3 in a recent study (Kaas, Weigelt, Roebroeck, Kohler, & Muckli, 2009), while area V5/MT+ was activated during the task. At the current time the specific interactions between memory and imagery processes is yet to be understood. However, this clearly demonstrates that visual perceptual areas are recruited during a range of tasks involving visual information, even when a visual stimulus is absent.

The view emerging from the recent studies on working memory is that the distinction between functions of perception and memory, and between different types of memory seems to depend on the differential involvement of distributed systems in the brain, rather than systems contained in segregated, centralized modules (Haxby et al., 2000; Postle, 2006). From a perspective of neuroscience one can argue that the idea of the same brain circuitry involved in perceptual representation also supporting maintenance and storage of information, is a more parsimonious hypothesis compared to the postulation of separate dedicated memory buffers, in which task relevant information requires a form of transfer from the different perceptual systems to such buffers (D'Esposito, 2007). The increasing number of empirical dissociations of different types of information in visual working memory would require a division of PFC into hundreds of subsystems. Several researchers has suggested an alternative: that working memory arises from recruitment of the brain systems evolved to accomplish tasks related to perception and action guiding, and that PFC is involved in the modulation of these specialized areas, rather than being the neural substrate of specialized

storage buffers (D'Esposito, 2007; Postle, 2006; Ranganath, 2006). If this model is correct, retention of basic visual properties should recruit visual perceptual areas at the early stages in the visual processing hierarchy. In line with this model, there is evidence for sensory recruitment in working memory for even less abstract representations than faces and objects. Converging evidence from studies using different methods; psychophysics, single cell recordings and functional brain imaging, points to a sensory working memory system for retaining low-level features which recruits the same brain areas that are involved in the perceptual discrimination of the remembered features (for a review see Pasternak & Greenlee, 2005). Basic sensory dimensions appear to be stored by specialized systems, each tuned to a specific dimension, such as direction of motion in the visual system, sound frequencies in the auditory system, or tactile information about shape or vibrations in the somatosensory system. In each of the senses there seems to be certain properties that work as basic building blocks from which more complex percepts are built. So what constitutes a basic dimension in the visual processing stream?

1.3 Low-level features and the visual processing stream

The cortical visual system is divided into several different subsystems, many of which have their own retinotopically organized representation of the visual field (Wandell, Dumoulin, & Brewer, 2007). Two main principles of the visual processing system account for the large number of different visual areas; hierarchical organization and modular specialization (Grill-Spector & Malach, 2004; Zeki & Shipp, 1988). The principle of hierarchical processing in the visual system is evident in that information is processed gradually from basic and local representations, to more complex, abstract and holistic representations, while the principle of specialization points to the separate neural mechanisms and pathways involved in the processing of different aspects of the incoming visual sensory data. Functionally specialized hierarchical pathways enable paralleled extraction of different features and components in the visual scene, which then can be integrated into unified percepts (Nassi & Callaway, 2009). One example is the differentiated pathways for action and object recognition, the dorsal and the ventral visual stream, respectively (Goodale & Milner, 1992). The ventral stream, also known as the occipito-temporal pathway, is tuned to information about color and shapes and is thought to be specialized for object recognition.

The dorsal stream, or occipito-parietal pathway, is specialized for processing spatial properties and guiding action, representing properties such as location and movement. This segregation of information is already evident in the retinal projections to the lateral geniculate nucleus (LGN) of the thalamus. The parvocellular projections (known as the P-pathway) convey information of color and high spatial frequencies over small receptive fields, and transmit with slow axonal conduction speed. The magnocellular projections (M-pathway) convey achromatic information with high sensitivity to contrast, low spatial frequencies and high temporal frequencies over large receptive fields, with fast axonal conduction speeds (Nassi & Callaway, 2009). The pathways project to different layers of LGN and the segregation between the P- and M-pathway is further kept when the LGN-projections terminate in the primary visual cortex (V1). V1, the first cortical area that processes visual information, has segregated channels for processing basic visual features, with neurons tuned to low-level attributes such as orientation, size (spatial frequency), contrast, color and motion (De Valois & De Valois, 1990; Nassi & Callaway, 2009; Pasternak, et al., 2003). Neurons in V1 coding for these features are highly active in the processing of incoming sensory information.

Since these dimensions are fundamental elements in the early processing stages in visual perception, and because the neural mechanisms behind the perceptual coding of these dimensions are relatively well understood, they have been an ideal focus for psychophysical research trying to understand the mechanisms of visual working memory, and how it relates to the perceptual system (Pasternak & Greenlee, 2005).

1.4 Low-level working memory

Studies using the delayed discrimination paradigm have shown that basic stimulus features can be retained in memory for several seconds with minimal decay (Magnussen & Greenlee, 1999). Spatial frequency and motion have been shown to have almost perfect retention over ISIs ranging from 1 to 30 seconds (Magnussen & Greenlee, 1992; Magnussen, Greenlee, Asplund, & Dyrnes, 1990; Regan, 1985), while color, contrast and orientation show a minimal decay (Lee & Harris, 1996; Nilsson & Nelson, 1981; Vogels & Orban, 1986). This clearly separates the underlying process from iconic memory, in which the representations show decay on a timescale of milliseconds (Sperling, 1960). The fact that discriminations which are based on memories of these fundamental stimulus dimensions are as accurate as

discriminations based on on-line perceptual representations, point towards a dual function for the early visual system; both in perceptual coding and in the storage of basic stimulus dimensions, as suggested by the sensory recruitment model of visual working memory.

There exists evidence that memory for basic stimulus dimensions are represented separately by neural mechanisms specialized for a given dimension. Studies have shown that when participants are performing delayed discrimination of spatial frequencies, the thresholds are not affected by the relative orientations of sample and test gratings, and delayed discrimination of orientation is not affected by the relative differences in the spatial frequency of gratings (Bradley & Skottun, 1984; Magnussen, Idås, & Myhre, 1998). Participants are even able to simultaneously judge two stimulus dimensions in dual-task experiments - without showing elevation in thresholds compared to the single judgment task - as long as the dual task is performed on different dimensions (Greenlee & Thomas, 1993). On the other hand, if participants are required to keep track of two components of the same dimension, the thresholds increase substantially (Magnussen & Greenlee, 1997).

Interference within, but not between basic visual dimensions has also been observed for discriminations in experiments using the memory masking paradigm. If a masker stimulus is presented in the memory interval between the sample and test stimulus in a delayed discrimination task, it can interfere with the memory representation of the sample stimulus, and thereby decrease performance on the task. Memory masking was first reported by Magnussen et al. (1991), in a study of memory for spatial frequencies. They showed that as the difference between the spatial frequencies of the masker and sample/test stimuli increased, the masker stimulus increasingly interfered with the memory representation of the sample stimulus. The masker reached a maximum interference effect when the difference reached ± one octave (half or twice the frequency), with about a doubling of delayed discrimination thresholds. When the mask was similar to the sample/test, no masking effect was observed. Varying the masker stimulus along another dimension unrelated to the discrimination task (e.g. varying orientation when discriminating spatial frequencies) did not affect discrimination thresholds. The specific type of interference gives us a clue as to the nature of the representation, and it seems that memory for spatial frequency and orientation is retained by separate, specialized mechanisms. This suggests that the location of these processes is early in the visual processing hierarchy. The memory masking effect has been consistently replicated for low-level features such as spatial frequency (Lakha & Wright, 2004; Lalonde &

Chaudhuri, 2002), and speed and direction of motion (Magnussen & Greenlee, 1992; McKeefry, Burton, & Vakrou, 2007). The memory interference show the same characteristics across visual features and experiments, which indicate a similar architecture of the underlying memory networks involved.

According to the sensory recruitment model, memory for these features involves the visual areas specialized for perceptually coding the information that is retained. One such specialized area in the visual perceptual system is V5/ MT+ (Pasternak et al., 2003). It is located in the dorsal visual stream, and contains a full representation of the contralateral visual field, representing information in a retinotopic (Gardner, Merriam, Movshon, & Heeger, 2008) or spatiotopic fashion (d'Avossa et al., 2007). Most of the neurons in this area show selectivity for the perception of speed and direction of motion. There is evidence that also memory for motion recruits neurons in this area. Several studies using single cell recordings in non-human primates have shown that neurons in area MT show sustained memory related activity for motion direction (Bisley, Zaksas, Droll, & Pasternak, 2004; Zaksas & Pasternak, 2006). Similar findings has also been reported in humans using TMS (McKeefry, Burton, Vakrou, Barrett, & Morland, 2008). These researchers showed that thresholds for delayed discrimination of speed of motion increased when repetitive TMS was applied to V5 and V3a, but not when applied to V1. TMS to V5 and V3a did not affect discrimination of spatial frequencies, which indicates that this feature is processed elsewhere (sadly, TMS over V1 during discrimination of spatial frequencies was not performed). Studies have shown that spatial separation of sample and test stimuli in delayed discrimination tasks for direction of motion elevates discrimination thresholds (Ong, Hooshvar, Zhang, & Bisley, 2009; Zaksas, Bisley, & Pasternak, 2001), and that the critical spatial separation corresponds to the receptive field sizes of the V5-neurons, indicating that memory representations are confined retinotopically in the cortex. These studies support to the idea that there is a close connection also between the neural mechanisms involved in low-level perceptual analysis and those involved in keeping this information in short term memory.

As for which areas that are involved in memory for orientation and spatial frequencies, the evidence is less clear. The primary visual cortex contains neurons tuned to specific spatial frequencies and orientations, however these features are not coded completely independent of each other, instead V1-neurons are tuned to multiple dimensions (De Valois & De Valois, 1990; Pasternak et al., 2003). Mapping of the organization of V1 has shown that neurons

tuned to orientations and spatial frequencies are systematically ordered in the cortex in a pinwheel configuration, such that each part of the visual field is processed by neurons coding for different spatial frequencies across different orientations (Issa, Rosenberg, & Husson, 2008). Does visual working memory for these features recruit neurons in the primary visual cortex? The memory masking effect discussed above has been suggested to result from inhibition between multiple channels, each channel being tuned to a limited range of values along a given dimension. The selectivity of the memory masking effect is similar to the selectivity of channels shown in adaptation masking experiments, in which the repeated exposure to a high contrast stimulus of certain value along a basic dimension, such as spatial frequency, selectively raise detection thresholds for low-contrast stimuli with similar values (see section 7.1 for a closer discussion). Adaptation masking is suggested to reflect cross-channel inhibition in V1, and this similarity point to overlapping neural mechanisms in perception and memory for spatial frequencies (Magnussen, 2009). However, a study by Bennet and Cortese (1996) showed that the memory masking effect is tuned to the perceived rather than the retinal spatial frequency, meaning the masking effect obeys size constancy. As in earlier studies, Bennet and Cortese found that the masking effect increased as a function of the difference between mask and sample/test gratings, but by presenting the mask further away than the sample and test stimuli, they were able to show that the masking effect was smallest when the mask and target had identical distal frequencies (cycles pr cm, c/cm) even though the retinal frequency (cycles pr visual degree, c/deg) ratio between mask and target in this condition was 0.5 (-1 octave). This means that the memory masking involves computations at a level of the visual processing system where size constancy is processed. Magnussen and colleagues (Magnussen, Greenlee, Baumann, & Endestad, 2009) have proposed that feature specific memory mechanisms for low-level features such as spatial frequencies and orientation reside at a higher level in the visual hierarchy, but that V1-neurons tuned to the remembered feature are recruited during retrieval of the memory representation. This suggestion was based on an earlier finding by Magnussen and colleagues (1998), in which they demonstrated that even though difference in orientations of the stimuli in a discrimination task for spatial frequencies did not elevate discrimination thresholds, there was a linear increase in choice reaction times with the separation of angle between sample and test stimuli. This suggests that delayed discrimination is based on representations where both these features are coded together, and that a systematic search in a structured network of multiple tuned channels in V1 is performed to extract the spatial frequency information across

orientations, which result in the observed increases in choice reaction times. A recent fMRI-study (Baumann, Endestad, & Magnussen, 2008) investigated this finding further, using a delayed discrimination task for spatial frequencies in which the difference in orientation between the sample and test stimuli could be the zero or 90 degrees. Replicating the earlier results by Magnussen and colleagues (1998), they found that discrimination accuracy was not affected by differences in orientation, but reaction times increased when sample and test had an orthogonal orientation. In agreement with the idea of a systematic search in V1, they found increased activity in the response to the test stimulus when the orientations where orthogonal, and the change in V1-BOLD signal correlated highly with the increase of reaction times. Interestingly, they also found increased activity in extrastriate areas (BA18, or V2) when orientations differed, which may suggest this area also plays a role in the suggested extraction process. Prefrontal (BA46) and parietal (BA40) areas were also activated by the memory task, but did not show a differential effect of the stimulus orientation conditions. The retrieval of high fidelity representations thus recruits early visual cortex, at the level where the retained information is perceptually coded. The results also seems to agree with the sensory recruitment model in that the differential demands (as expressed by the differing choice reaction times) introduced by the orthogonal vs. same orientation of stimuli in the delayed discrimination task is reflected at the level of the perceptual system where these features are coded perceptually, in V1.

However, until recently, it has remained unclear whether the recruitment of V1 and extrastriate visual cortex is limited to the retrieval of low-level memory representations, or if there is a sustained memory-related activation across the delay intervals. Investigating this, Serences and colleagues (2009) used multi-voxel pattern analysis (MVPA; Norman, Polyn, Detre, & Haxby, 2006) to decode activation patterns in V1 during the delay periods in a delayed discrimination task for color and orientations. MVPA is an approach where the spatial pattern of activations are taken into account, and can thereby be used to probe what type of information gets represented by a given brain region. By analyzing the spatial pattern of activity in voxels coding retinotopically for the stimulus position in V1 during the memory interval, they were able to decode the remembered feature of the sample stimulus. That is, the pattern of activation discriminated orientation, but not color, when the task relevant feature was orientation, and vice versa. They also compared decoding accuracy when data from the last time point in the memory interval was included or excluded from the analysis. Including the last data point significantly improved decoding accuracy for the task-relevant stimulus.

The fact that information from the last data point in the delay period contributed to decoding accuracy, points to a feature-specific activation in area V1 that is sustained during the memory interval. Similar findings have been reported by Harrison and Tong (2009) for the feature orientation. However, and important for the rationale for this study, both studies failed to detect sustained BOLD-activation in V1 during the intervals using a standard univariate analysis. Offen and colleagues (2009) reported similar results. They measured BOLD-activity during delay periods in a delayed discrimination task for spatial frequencies and orientations, and found no sustained activity during the delays, i.e. the activation returned to baseline. This lack of correspondence between the MVPA findings and the results from the univariate analyses is suggested to result from suppression of neurons tuned to non-remembered values of the task relevant feature, i.e. cross-channel inhibition. The increase in BOLD-signal from neurons that are more active, and the reduction in BOLD-signal from the less active neural population would then be spatially integrated (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001; Shmuel, Augath, Oeltermann, & Logothetis, 2006), canceling each other out on the measured voxel level. In the current study we chose an approach in which we take advantage of this inhibition to explore working memory for spatial frequencies. As described above, the memory masking paradigm (Magnussen et al., 1991) introduces selective interference with memory representations for spatial frequency when presenting a masker stimulus in the delay interval which differs from the spatial frequency of the sample stimulus. The inhibitory effect induced by the mask can then be measured using univariate BOLD analysis to further investigate the role of V1 in low-level working memory.

There is also a recent finding concerning the spatial extent of sensory recruitment in low-level memory that seems to contradict earlier studies; Ester and colleagues (2009) found evidence for feature-specific activation patterns related to remembered orientations in ipsilateral V1 (relative to position of stimulus). They suggest that sensory recruitment is global rather than retinotopically confined in the visual cortex, a mechanism which could enhance the robustness of the stored information. This seems to contrast the findings by Zaksas and colleagues (2001), and by Ong and colleagues (2009), which find reduced delayed discrimination performance when stimuli are separated. We explore the spatial extent of low-level memory representations further, by conducting a second experiment where the mask is presented to a different retinotopic position than the sample and test stimuli. If the memory representations are retinotopically confined, the mask should not interfere with the delayed discrimination task when mask and sample are spatially separated.

Another divergent finding we want to explore further concerns sensory recruitment of extrastriate areas. The abovementioned study by Serences and colleagues (2009) was not able to decode rembered features in visual areas beyond V1. In contrast, Harrison and Tong (2009) were able to decode reliably the rembered features in V2, V3, V4/V3a, in addition to V1. Both these studies used delayed discrimination for orientations, so it is unclear what underlies the different results. To further investigate the involvement of extrastriate areas in low-level working memory, and since the psychophysical evidence suggest that also extrastriate areas are involved in low-level memory for spatial frequencies (Bennett & Cortese, 1996), we include these visual areas in our analysis of the memory masking phenomenon.

1.5 Current study and hypothesis

We conducted an fMRI study on memory for spatial frequency based on a memory masking paradigm. In contrast to the studies by Harrison and Tong (2009) and Offen et al. (2009), that do not find sustained activations during memory intervals using conventional univariate approaches in simple delayed discrimination tasks, we use the memory masking paradigm to modulate the strength of memory representations from trial to trial, by introducing a masker stimulus that either shares or differs in spatial frequency from the sample. Since the memory masking effect is specific to the remembered visual feature, visual areas showing activity modulation related to the masking effect can be presumed to represent this feature in memory. Since the mask and sample is separated in time, any modulation of activity caused by an interaction between these representations can be attributed to processes of mnemonic nature. This approach enables detection of even subtle changes in neural activity caused by an interaction between memory representations, thereby enabling us to test the sensory recruitment hypothesis for low-level working memory using a conventional fMRI-approach. This study therefore complements the abovementioned studies using a decoding approach, and as far as we are aware it is the first study investigating the memory masking effect using fMRI.

In one experiment, we presented all stimuli in a trial to the same position, ensuring overlap between the retinotopic representations of the mask and the sample stimulus at encoding. Motivated by the recent findings by Ester and colleagues (2009), we conducted a second experiment to explore the spatial extent of the memory masking effect. In this second experiment we presented the mask stimulus to a different position than the stimuli in the main

discrimination task, thereby producing different retinotopic activity patterns between the mask and the sample stimulus.

Based on the abovementioned empirical findings, and using the memory masking paradigm, we wanted to test the following hypotheses:

- Spatial frequency coding neurons in V1 are recruited in the retention of spatial frequency information. We predict that the introduction of a masker stimulus will produce a weaker BOLD-response in visual areas involved in the storage of spatial frequency information.
- Low-level representations of spatial frequency information are localized in areas with retinotopic organization. The process behind memory masking is suggested to be lateral inhibition between narrowly tuned feature specific channels. If memory representations are retinotopically organized, memory masking should be impaired when mask and sample are presented to different visual hemifields. On a behavioral level, we expect to find smaller effects of memory masking when mask and sample stimuli are spatially separate (Experiment 2), than when they are presented to the same location (Experiment 1). Correspondingly, on a physiological level (fMRI), we expect to find weaker or no modulations of the BOLD response in early visual areas due to memory masking when mask and sample stimuli are spatially separate (Experiment 2).
- In addition we want to investigate the involvement of extrastriate areas in low-level memory, by exploring the effects of memory masking throughout the early visual system.

Since our area of interest in the current study is retinotopic areas coding for stimulus position in the early visual cortex, we have chosen a region of interest (ROI) approach to increase detection power. Our hypotheses are directed at processes in specific visual areas, therefore we conduct a retinotopic mapping session for all our participants. This enables us to map each participant's visual cortex and to define regions of interest within visual areas on an individual basis. The retinotopic mapping approach gives a much higher precision compared to definitions of visual areas based on stereotaxic coordinate systems, such as MNI and Talairach, since both anatomical landmarks and functional organization are known to show individual differences (Wandell et al., 2007). This ensures that we look at functionally equivalent regions across the participants. In addition, we perform a separate localizer session to identify regions within the different visual areas responding to the positions occupied by

the stimuli in the main experiments. Thus, this study will investigate activity modulations across conditions for regions of interests in visual areas V1-V4/V3a.

2 Materials and Methods

2.1 Overview of experiments

This study has two parts. One psychophysical, conducted in a standard laboratory setting, where behavioral data was acquired. This part comprised two experiments and an initial threshold estimation part. The other part uses functional magnetic resonance imaging, where blood-oxygen-level dependent (BOLD) measurements as well as behavioral data were acquired as participants performed tasks while positioned inside the MR-unit. This part consists of two experiments, plus a retinotopic mapping session and a localizer session.

2.2 Participants

Three male participants took part in the experiments. The mean age of the participants was 28.8 years (age range, 27-30 years). All participants were right handed, had normal or corrected to normal vision, and no history of ocular disease. The participants were experienced psychophysical observers and were thoroughly trained on the experimental task before the reported data were measured. None of the participants were paid for the participation.

2.3 Stimuli and stimulus presentation

In all experiments and for estimating discrimination thresholds, stimuli were Gabor gratings created using custom made scripts written in MatLab (version 7.1; The MathWorks, Inc., USA). The sinusoid constituting the spatial frequency content of the grating subtended 10° of visual angle, and was tapered with a Gaussian kernel with a standard deviation of 1.25°. The background, to which the sinusoid faded, had RGB-values of 127,127 and 127 respectively. The sinusoid could have an orientation of 0° (vertical) or 90° (horizontal). The phase of the sinusoid was kept constant within an experimental trial, but varied randomly between trials. The Gabor gratings had a maximum Michelson's contrast of 0.9, so the area of the patch with a Michelson's contrast over 0.1 had a diameter of approximately 5.2° of visual angle. Gabor grating stimuli were presented at four different positions in the experiments,

located in each of the four visual field quadrants. The distance between center of the screen (fixation), and the center of the Gabor grating was 6° of visual angle for all four positions.

For the fMRI localizer session, stimuli were created by masking a black and white dartboard of maximum contrast with a black mask, producing windows with diameters of 5.2° of visual angle centered at the four positions of interest. The patches of the dartboard were approximately scaled following the linear cortical magnification factor, $M(\theta) = 13.48\theta^{-1}$ (Qiu et al., 2006).

A fixation cross (width/height: 0.5° of visual angle) indicated the center of the display in all sessions, except for in the retinotopic mapping session, where a central fixation point was used.

In the psychophysical sessions, stimuli were presented on a calibrated 19-inch Eizo FlexScan L768 monitor (Eizo Nanao Corporation, Japan). Screen resolution was set at 1024x768, with a vertical refresh rate of 60 Hz. The participants indicated their responses using a PST serial response box (Psychology Software Tools, Inc., USA). The subject's head was stabilized and viewing distance kept constant at 57 cm using an adjustable chin-and-forehead rest.

In all fMRI sessions, except the retinotopic mapping procedure, stimuli were back-projected on a screen inside the scanner with a modified F20 sx+ SXGA+ DLP digital projector (Projectiondesign, Norway). Screen resolution was set at 1024x768 with a vertical refresh rate of 60 Hz. We used a SyncBox (Nordic NeuroLab, Bergen) to synchronize stimulus presentations with trigger signals from the MR-scanner. The participants indicated their responses using two response grips (NordicNeuroLab, Norway). Viewing distance in the scanner was 64 cm, and the size of the stimuli on the screen was adjusted to account for the increase in distance compared with the psychophysical sessions.

In the retinotopic mapping session stimuli were presented using MR-compatible video-goggles (NordicNeuroLab, Bergen). The goggles have poorer spatial resolution compared to the projector (800x600), but has the advantage of a continuous visual field (not interrupted by the head coil) that spans 30 degrees in the horizontal direction and 23 degrees in the vertical direction. Vertical refresh rate was set to 85 Hz. The stimuli used are presented in the retinotopic mapping section.

For the psychophysical threshold estimation sessions and the retinotopic mapping procedure, the stimuli were delivered using the Psychophysics Toolbox (v.3) extension for MatLab (Brainard, 1997; Pelli, 1997). For the other experiments, including the fMRI-localizer session, E-Prime 1.2 was used (Psychology Software Tools, Inc., USA).

2.4 Experimental design

2.4.1 Psychophysical experiments

The two experiments were modified versions of the memory masking paradigm used by Lalonde & Chaudhuri (2002). This version of the paradigm differs from the original (e.g. Magnussen et al., 1991) in that the interfering stimulus (the mask) is presented before the sample stimulus to be remembered. Additionally, the mask is involved in a second, much easier, discrimination task to ensure that it is actively encoded. All stimuli in our study were presented in the visual periphery. Other than the positions of the stimuli, all parameters were identical for the two experiments. Both experiments were tested psychophysically before fMRI measurements were conducted.

Memory masking experiments have traditionally used the method of constant stimuli to calculate psychometric functions of discrimination at different mask/sample ratios. This is a very time consuming method as each psychometric function (i.e. each condition) requires a large amount of observations to be fitted correctly, and is not suitable for an fMRI investigation. We therefore chose another approach: first we calculated the level at which each participant could discriminate the difference between the sample and the test stimulus correctly 75% of the trials using an adaptive procedure. This level corresponds to the steepest point of the slope of the psychometric function, thus changes in the subject's discrimination capacity (due to experimental manipulations) produce the largest effects on performance around this level. Second, we used the calculated level as the difference between the sample and the test stimulus in 3 masking conditions: the mask could be identical to the sample stimulus, or it could differ with ±1 octave (since masking effects usually are largest at this difference). The second comparison, which was included to ensure participants actively processed the mask, was a much easier task as the difference between the two stimuli was set to be at an estimated 85% correct discrimination level.

All participants went through 2x3 sessions for estimating discrimination thresholds, using an adaptive, maximum likelihood procedure, QUEST (Watson & Pelli, 1983),before being tested in the experiments. Each session consisted of 40 trials of a two-interval forced-choice (2-IFC) delayed discrimination task. A sample stimulus appeared in one of the four stimulus positions for 250 ms, and was followed by presentation of a test stimulus for 250 ms, in the same position. The participants had to decide which of the two stimuli had the higher spatial frequency. The spatial frequency of the sample stimulus varied randomly between 2 and 5 cycles per visual degree (c/deg), and the test stimulus had a higher or a lower value, the amount of difference decided by the QUEST-algorithm for each trial. The participants were informed that the target could appear in either stimulus interval with equal probability. Stimulus position varied randomly between trials. For the 3 first sessions, an interstimulus-interval (ISI) of 3,000 ms was used, and the desired threshold estimate was set to a hit rate of 75%. For the last 3 sessions the ISI was set to 9,000 ms, and the desired threshold estimate to 85% hit rate. The resulting estimates were used to set appropriate differences for each participant between the different stimuli in the experimental trials.

The main structure was similar for both experiments and is visualized in Figure 1. A central fixation cross appeared and stayed on the screen until the last response was produced. The participants were instructed to keep fixation throughout the trial. 1000 ms after the fixation cross the mask stimulus (S1) appeared for 250 ms. After an ISI (ISI-1) of 3,000 ms, the sample stimulus (F1) appeared for 250 ms. Then, after an ISI (ISI-2) of 3,000 ms, the first test stimulus (F2) appeared for 250 ms. The participant indicated which interval had the higher spatial frequency (F1 or F2) with a response (RESP-1). 2,500 ms after the offset of the test stimulus, the second test stimulus (S2) appeared for 250 ms. The participant again indicated which interval had the higher spatial frequency (S1 or S2) with a response (RESP-2) within 2500 ms after the offset of S2. The fixation cross then disappeared, indicating that the trial was over. Thus, one trial lasted 13,000 ms. To avoid confusion of the stimuli in a trial and to control for the possibility of priming effects confounding the encoding of the sample stimulus (F1), the orientations of the Gabor gratings in the S1-S2 task and the F1-F2 task were always orthogonal to each other. The specific orientation (0°or 90°) of the two stimulus pairs was randomized across trials.

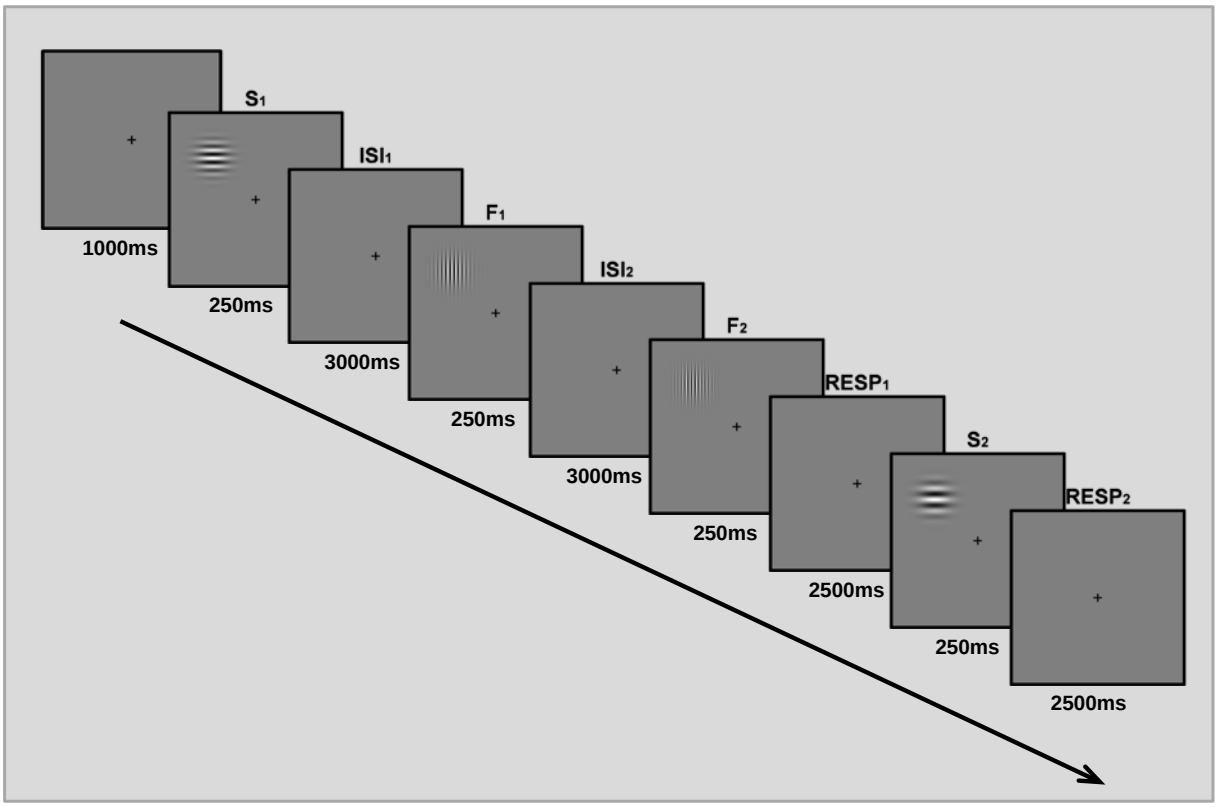

Figure 1 - Overview of experiment

The main structure of an experimental trial (similar for psychophysical and fMRI trials). In the illustration, all stimuli are presented in the same position. This was the case in Experiment 1 (psychophysical and fMRI). In Experiment 2 (psychophysical and fMRI), stimulus F1 and F2 were presented at the far opposite position of stimulus S1 and S2 - in the example trial above, F1 and F2 would have been presented in the lower right quadrant. The F1-F2 comparison is the main task of interest in the current experiments. The difference between stimulus F1 and F2 was adjusted in a separate threshold estimation sequence to produce a baseline hit rate of approximately 75%. Stimulus S2 was included mainly to ensure the encoding of stimulus S1. The difference between stimulus S1 and S2 was adjusted to produce a hit rate of approximately 85%, thus this comparison was perceived as an easier task than the F1-F2 task. In the example above, stimulus S1 has a spatial frequency one octave lower then stimulus F1 (Mask/Sample Grating Ratio = 0.5). Note also that S1 and S2 are oriented horizontally, while F1 and F2 are oriented vertically. This orthogonality was always present between S1-S2, and F1-F2. However, the specific orientations varied between trials.

The intervals between F1 and F2, and S2 and S2 (3,000 ms and 9,000 ms, respectively), were similar to the ISIs used in the threshold estimation procedure above. The estimated individual thresholds, 75% hit rate at an ISI of 3,000 ms, 85% hit rate at an ISI of 9,000 ms, were used as the percentage difference between the stimuli in the two tasks (F1-F2 and S1-S2) in a trial. The test stimuli (F2 and S2) could increase or decrease with this

percentage, and the participants knew that both directions of change occurred with equal probability.

The spatial frequency relationship between the mask stimulus (S1) and the sample stimulus (F1) varied from trial to trial in 3 established ratios: the spatial frequency of S1 could be the same as for F1 (Mask/Sample Grating Ratio = 1), one octave above F1 (Mask/Sample Grating ratio = 2), or one octave below F1 (Mask/Sample Grating Ratio = 0.5). The stimuli used in the experiments varied across a spatial frequency range of 1.2 - 6 c/deg, with an average frequency of 3 c/deg. One round of the experiment consisted of 216 trials, divided into 3 testing blocks separated by breaks. Each mask/sample-ratio was tested 72 times per experiment. These 3 main conditions, as well as the other task conditions (F2 higher/lower than F1, S2 higher/lower than S1) were counterbalanced and randomly selected within a round. All participants were tested 3 rounds on each of the two experiments, producing 648 responses each on the task of main interest (F1-F2): 216 observations per mask/sample-ratio per experiment.

The two experiments varied only in the relative positions of the S1-S2 task and the F1-F2 task within a trial. In Experiment 1, all stimuli in a trial were presented in the same stimulus position. The positions varied randomly between trials, with the criterion that the same position could not be sampled two trials in a row. In Experiment 2, S1 and S2 were presented in one position, while F1 and F2 were presented in the far opposite position (e.g. S1 and S2: upper left quadrant; F1 and F2: lower right quadrant). Each starting position was tested 54 times per round of the experiments. The sessions constituting the two experiments were run interleaved, with the sampling order counterbalanced between participants.

2.4.2 fMRI-experiments

The stimuli and the structure of the two fMRI experiments were practically identical to psychophysical Experiment 1 and psychophysical Experiment 2. However, to ensure that the hemodynamic response went back to an approximate baseline within areas retinotopically coding a stimulus position between trials, no position could be tested two trials in a row. For Experiment 2, where S1/S2 and F1/F2 were presented on opposite sides within a trial, the same diagonal could not be sampled in adjacent trials.

The average intertrial interval (ITI) was 1 TR, but was randomly jittered with ±1/2 TR (700 ms) in 2/3 of the trials. Thus, there was at least an interval of 16,800 ms between offset of the last stimulus in a trial (S2) and onset of the next stimulus (S1) in the same position (same diagonal for Experiment 2). As for the psychophysical experiments, the orientations of the Gabor gratings in the S1-S2 task and the F1-F2 task were always orthogonal to each other. This difference between the two tasks constituting a trial should not affect the hemodynamic response at the level of investigation, since it has been shown that horizontally and vertically oriented gratings elicit equally strong BOLD responses in V1 (Furmanski & Engel, 2000).

Based on the findings from the psychophysical experiments, the two masking conditions (Mask/Sample Grating Ratio = 2, Mask/Sample Grating Ratio = 0.5) were collapsed into one condition: Mask/Sample Grating Ratio $\neq 1$. Thus, for the fMRI experiments there were two main conditions. An experimental run in the scanner contained 88 trials, 22 trials at each stimulus position (S1 position for Experiment 2). Each condition was sampled the same number of times at each position. The duration of an experimental run in the scanner was approximately 20.5 minutes. Each participant finished 4 runs of each experiment, spread over 4 testing sessions, producing 352 responses each on the task of main interest (F1-F2): 176 observations per mask/sample grating ratio per experiment. The two experiments were run interleaved, with the sampling order counterbalanced between participants.

2.5 MR data acquisition

2.5.1 fMRI-data

We used magnetic resonance imaging at 3T (Achieva, Philips Medical Systems, Best, The Netherlands) at the Interventional Centre at Oslo University Hospital to measure blood-oxygen level dependent (BOLD). An 8-channel Philips SENSE head coil was used. The functional images were acquired using a T2*-weighted echo-planar imaging (EPI) sequence. Before each functional scanning session, a survey sequence with 7 sagital slices was used to precisely locate the calcarine sulcus.

For the experiments and the ROI localizers, 24 transversally oriented slices (interleaved, no gap) were oriented along the calcarine sulcus to cover the visual cortex.

Repetition time (TR) was 1,400 ms, echo time (TE) 30 ms, and flip angle 70°. Field of view measured 192x192x48 mm, and voxel size 2x2x2 mm. A scanning session consisted of two experimental runs, in each run 890 functional volumes were acquired. For the ROI-localizer, 620 functional volumes were acquired.

For the retinotopic mapping procedure 31 transversally oriented slices (interleaved, no gap) covering the whole brain were measured. TR was 2,000 ms, TE 30 ms and flip angle 80°, field of view 192x192x62 and voxel size 2x2x2 mm. The retinotopic mapping session consisted of three runs for the polar angle mapping, and 1 run for eccentricity mapping. In each run 260 functional volumes were acquired.

2.5.2 Anatomical data

Anatomical images were acquired using a T1-weighted turbo field echo (TFE) pulse sequence with TR 9,64 ms, TE 4,59 ms and a flip angle of 8°. The whole-brain anatomical volume consisted of 192 sagitally oriented slices with an isometric voxel size of 1x1x1 mm. The field of view measured 256x256 mm. Slices were oriented along the AC-PC line. Two anatomical volumes were acquired for each participant.

2.6 Image preprocessing

2.6.1 Functional images

All imaging data were exported from the scanner in Philips PAR-REC-format, and were imported into BrainVoyager QX (v2.2, Brain Innovation, Maastricht, The Netherlands; Goebel, Esposito, & Formisano, 2006) and converted to FMR-format for preprocessing and analysis. First the images were corrected for slice time differences using information about TR and slice scanning order. Then motion correction was performed to determine and correct for small head movements. Each volume of a run was aligned to the first volume of that run (within-session alignment) using a 6 parameters rigid body transformation. Inspection of the estimated translation and rotation parameters showed only sub-millimeter movement for all participants and all runs. Low frequency drifts were removed using a temporal high pass-filter of 0.01Hz. To ensure the spatial precision of voxels in our ROI-analysis, no spatial smoothing was performed on the data, but we expect some spatial blurring to be introduced from the

motion correction and normalization procedures. FMRs were then co-registered to each participant's preprocessed high resolution anatomical by a 3 step procedure; first an automatic orientation coarsely aligned the FMRs with the structural image according to the spatial information from image headers. The images were then manually co-registered by visual inspection, before an automatic rigid body fine tuning of the co-registration was performed. The FMRs were then converted into a Talairach normalized volume time course format (VTC) using the estimated co-registration parameters and the Talairach transformation data from the structural images. For the two experiments the VTCs were preprocessed following similar routines as Offen et al. (2009) and Sligte, Scholte and Lamme (2009), also conducting ROI-based univariate analysis on memory related activity in early visual areas; the time series were, in addition to the high pass filtering, temporally smoothed with a low-pass filter of 2.8s, and normalized using z-transformation.

2.6.2 Anatomical images

The high resolution anatomicals were converted from PAR-REC to the internal Brain Voyager VMR-format. First the two separate images for each participant were corrected for spatial inhomogenities in intensity, by analyzing changes in white matter intensities across image space (Vaughan et al., 2001). The skull was stripped from the brain, and the cerebellum removed using manual procedures, to ensure optimal FMR-VMR and VMR-VMR co-registration. The two anatomical images for each participant were co-registered and then averaged together to produce a single high-resolution anatomical volume for each participant. These anatomical volumes were then transformed into Talairach space using a sinc interpolation algorithm (Figure 2A). The white-gray matter boundary was estimated and segmented by analysis of voxel intensity histograms specifically for the occipital cortex, to obtain as high spatial precision as possible in this area. White matter bridges were removed using automated algorithms in BrainVoyager QX. The resulting white-matter segments (Figure 2B) were then used to create 3D-meshes of the cortical surfaces (2C). The meshes were inflated (2D), and manually inspected and corrected for topological errors. The inflated meshes were then cut along the calcarine sulcus, and flattened to get a 2D representation of each participant’s visual cortex (shown in the retinotopic mapping session).

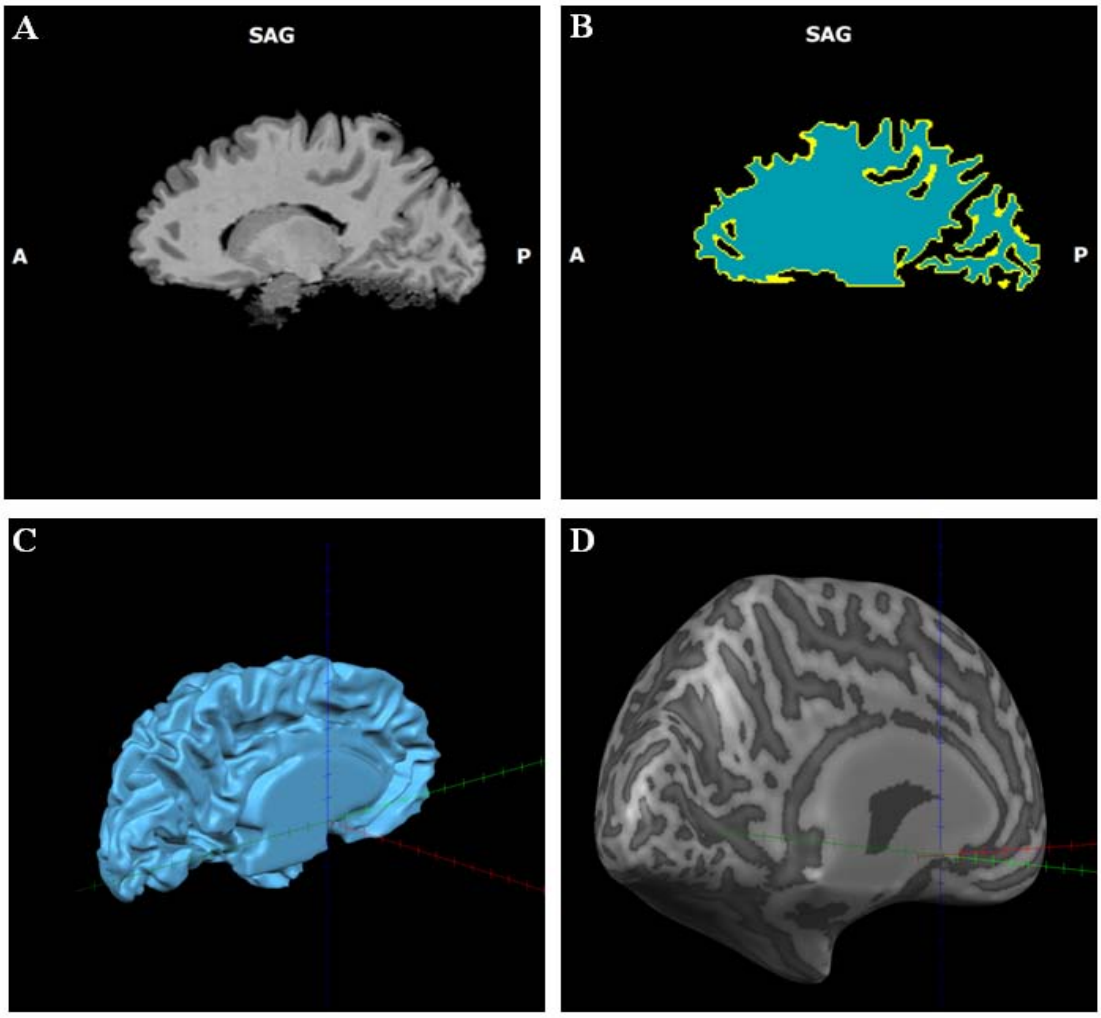

Figure 2 - Preprocessing of structural images

A) The two anatomical scans for each subject were corrected for intensity inhomogeneities and the skull and cerebellum were removed manually, before the two images were averaged together to get one high quality anatomical volume for each subject. The images were then transformed in to Talairach coordinates. **(B)** An automatic algorithm estimated the white-grey matter boundary. The estimated threshold value was adjusted on the basis of intensity histograms of the occipital cortex, to ensure the best possible spatial precision in this area. **(C)** A 3d-mesh of the cortical surface was generated on the basis of the white-matter segment, and **(D)** inflated. A cut was made along the calcarine fissure, before the inflated brain was flattened to get 2D-representations of the visual cortex (shown in retinotopic mapping session)

3 Retinotopic mapping and ROI-localizers

3.1 Visual field maps

Projections from the retina of the eye, to the layers in lateral geniculate nucleus (LGN), and further on to the visual cortex, preserve the topographic organization of the retinal image. Neighboring regions on the retina thus project to neighboring regions in the visual cortex (Grill-Spector & Malach, 2004). The visual field is divided into two hemifields along the vertical meridian. The retinal projections from the temporal part of the retina project ipsilaterally, while nasal fibers cross over at the optic chiasm and project contralaterally. Each visual hemifield is therefore represented in the contralateral visual cortex of the brain (e.g. left visual hemifield is represented by right visual cortex).

The primary visual area is located along the calcarine sulcus, with foveal visual representations in the posterior part of the calcarine cortex. As one moves from the posterior calcarine in the anterior direction, visual field representations change from center to the periphery of visual field. This dimension of the visual field map is referred to as eccentricity. The visual hemifield is further divided along the vertical and horizontal meridian, with the horizontal meridian represented along the fundus of the calcarine sulcus, and the upper and lower vertical meridians represented in the lower and upper lips of the calcarine, respectively. This dimension is referred to as polar angle. Even though the visual field changes with movements of the eyes, this organization is fixed with respect to the retinal image. Polar angle and eccentricity is therefore defined in relation to the participant's fixation point, and the visual field maps are therefore known as retinotopic maps. The visual system is divided into several different retinotopically organized areas in the occipital cortex. They are suggested to be part of different neural pathways for processing different stimulus properties (Grill-Spector & Malach, 2004). There are several criterions for defining a visual field map (for a review, see Wandell et al., 2007). Each field map contains only one representation of each point in the visual field. If two parts of the cortex represents the same part of the visual field, they are considered separate visual field maps. Second, a visual field map should represent a substantial part of the visual field in an ordered fashion in respect to both polar angle and eccentricity. Third, visual field maps should be consistent across individuals. Visual field

maps have by definition a contiguous representation of the visual field, but there are some important discontinuities. Already mentioned is the hemifield division along the horizontal meridian with each visual hemifield represented in the contralateral visual cortex. V1 has a continuous map of the contralateral visual hemifield, while V2 and V3 is divided into two quarterfield maps along the horizontal meridian, with upper quarterfield maps located ventral to V1, and the lower quarterfield maps dorsal to V1. The border between V1 and the dorsal and ventral V2 (V2d and V2v) is demarked along the lower and upper vertical meridian, respectively. The same goes for the border between V2d and V2v, and V3d and V3v, respectively. In the ventral visual cortex, adjacent to V3v and demarked along the upper horizontal meridian, the angular representation of V4 spans the horizontal meridian and continues almost down to the lower horizontal meridian. In the dorsal visual cortex, adjacent to the V3d, V3a has a complete hemifield map starting at the lower vertical meridian, spanning the horizontal meridian and ending at the upper vertical meridian.

There are additional maps to those described here, but we restricted our analysis to these regions, as we were interested primarily in V1, and the early visual areas V2, V3, V4, and V3a. These visual field maps are revealed by fMRI using a phase encoded mapping technique, also known as the traveling-wave method (Engel, Glover, & Wandell, 1997). A flickering wedge stimulus rotates around a central fixation point, and codes the angular (polar) component of the retinotopic map. A flickering ring that expands outwards from a central fixation codes the radial (eccentricity) component. The BOLD signal induced by the stimuli travels along the cortical surface, and this traveling wave of activation over time is used to map the eccentricity and polar angle of the visual field map in the visual cortex. At the meridians there is a reversal in polar angle, and these reversals mark the border between adjacent visual field maps in the visual cortex, as described above.

3.2 Retinotopic mapping procedure

We performed a retinotopic mapping experiment as described by Slotnick & Yantis (2003) to reveal early visual areas to V1, V2,V3 and V4/V3a in each of the participants' visual cortex. We used simultaneous stimulation of both hemifields during polar angle mapping for more efficient sampling, which is a suitable approach for the areas up to V4/V3a. For mapping the polar coordinates, participants were instructed to fixate as two wedges of flickering checkerboards (Figure 3A) with a contrast reversal of 10 Hz started to unfold (each

unfolding step increasing the wedge with 11.25 degrees) simultaneously at the upper right and lower left vertical meridian. After reaching its maximum extent at 45 degrees, the wedge rotated clockwise about the fixation point, in steps of 11.25 degrees pr TR, before folding in at the lower right and upper left vertical meridian. For the eccentricity mapping participants fixated as concentric rings of flickering checkerboards (Figure 3B) with a contrast reversal of 10 Hz started to grow outwards from the center in steps of 0.5 degrees, before reaching its maximum extent at 2 degrees. The rings then started expanding radially in steps of 0.5 degrees pr TR, before folding in at the edge of the screen. For both polar angle and eccentricity, one cycle was completed in 40s (20 TRs), and each session consisted of 10 full rotation cycles. To minimize eye movements and to help keeping fixation we avoided stimulation of the most central part of the visual field (using a central mask with a radius of 1 degree), thus the foveal representations in cortex are not mapped. The retinotopic maps were revealed by selecting the first stimulus position in a cycle as reference, and cross-correlate it with the number of TRs (lags) for one complete cycle. This determined the most effective stimulus-position (lag) for each voxel, and significantly activated voxels (cutoff at $r > .20$) were assigned a pseudo-color corresponding to the lag in which it had the highest correlation. The cross-correlation maps for both polar angle (Figure 3C) and eccentricity (Figure 3D) were then projected on the flattened cortical surface. Based on these maps, and following the descriptions by Wandell and colleagues (2007), borders between the visual areas were identified visually, and visual field maps drawn manually for each subject (Figure 5A, B & C) and saved as patches of interest (POI).

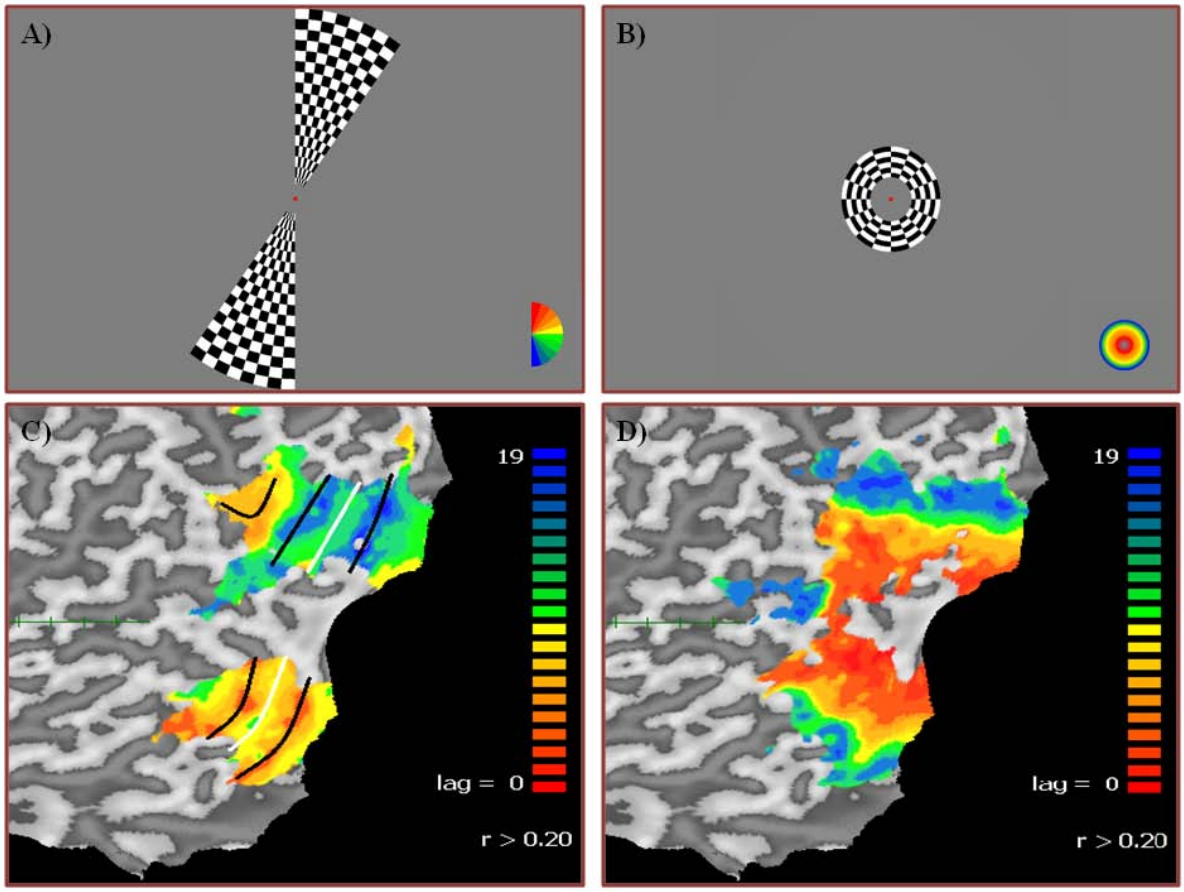

Figure 3 - Retinotopic mapping session

A) Example of the stimuli used to map polar angle. Wedges of flickering checkerboards rotated clockwise, creating a wave of activity traveling along the retinotopically organized visual cortex. The color wheel in the bottom right corner shows the colors assigned to the different polar coordinates in the left visual cortex (not shown during experiment). **B)** Example of the stimuli used to map eccentricity. Concentric rings of flickering checkerboards expanded outwards, creating a wave of activity in the visual cortex corresponding to the stimulated eccentricity coordinates. The color wheel in the bottom right corner shows the color assigned for the optimal stimulus positions for voxels in the visual cortex (not shown during experiment). **C)** Cross-correlation map (threshold r > .20) for the polar angle session, projected on the flattened cortical surface for Participant 1. Each lag is represented by an arbitrary color, coding the optimal stimulus position in polar coordinates for each part of the left visual cortex, starting with red at 12 o'clock (lag 0) to blue at 6'o clock (lag 20). Black lines correspond to phase reversals at the vertical meridian, white lines correspond to phase reversals at the horizontal meridian. **D)** Cross-correlation map (threshold r > .20) for the eccentricity mapping projected on the flattened cortical surface for Participant 1. Each color represents the optimal stimulus position in eccentricity coordinates, of from central (red) to periphery (blue). Since the area round the fixation is not stimulated, the foveal representations are not mapped.

3.3 ROI-localizer

In addition to retinotopic mapping, all participants went through a localizer session to identify regions of interest (ROIs). The basic procedure had a structure adapted from Kraft and colleagues (2005). The flickering dartboard patches were presented at each position of interest in the following sequence: upper left quadrant, lower right quadrant, upper right quadrant, lower left quadrant (Figure 4). Each position were stimulated for a period of 14,000 ms, with a flickering rate of 9 Hz. A testing block consisted of 4 presentations at each single position. Before each testing block, and after the last block in the session, there was a rest period without stimulation for 28,000 ms. There were 3 testing blocks in a session, i.e. each position was stimulated 12 times. A central fixation cross was present at all times and the participants were instructed to focus on this throughout the session.

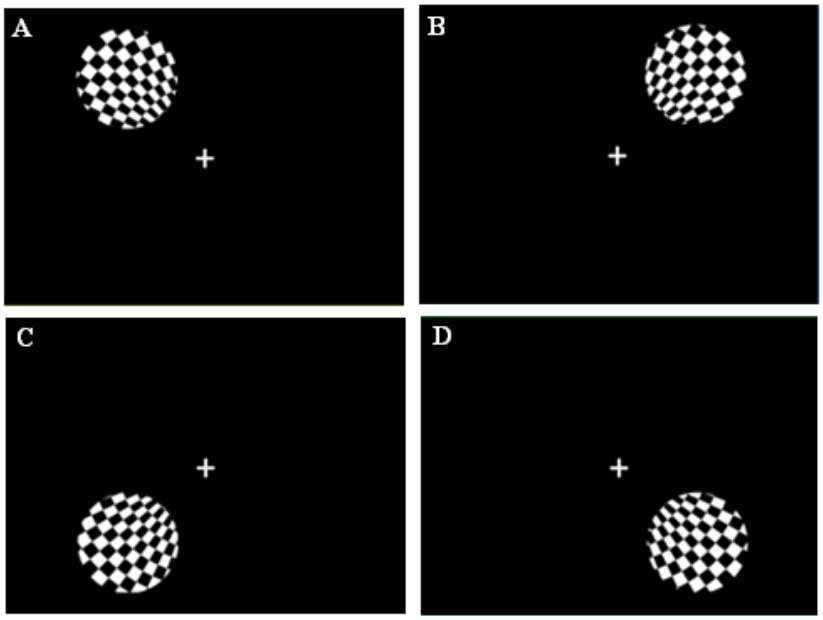

Figure 4 - ROI localizer session

Regions of interest were identified during a separate localizer session, where we stimulated the four stimulus positions used in the fMRI-experiment using flickering checkerboards. The positions were stimulated sequentially in a sequence corresponding to this figure; (A) Position 1: Upper left quadrant, (B) Position 2: Upper right quadrant, (C) Position 3: Lower right quadrant, (D) Position 4: Lower left quadrant. Voxels that responded to one quadrant in contrast to stimulation in the three other quadrants ($p < 0.01$, FDR, voxel threshold 5) were used to define regions of interest in the different visual areas.

We used a general linear model (GLM) to detect voxels that responded stronger to stimulation in one quadrant in contrast to stimulation in the other three quadrants. Voxels that

passed a statistical threshold of $p < 0.01$ (with a correction for multiple comparisons using false discovery rate (FDR), and a cluster threshold of 5 voxels) were projected on the flattened visual cortex for each subject (Figure 5, D, E & F). Regions of interest were defined for each visual area based on the overlap between the individual POIs and the voxel clusters for each position that survived the FDR threshold.

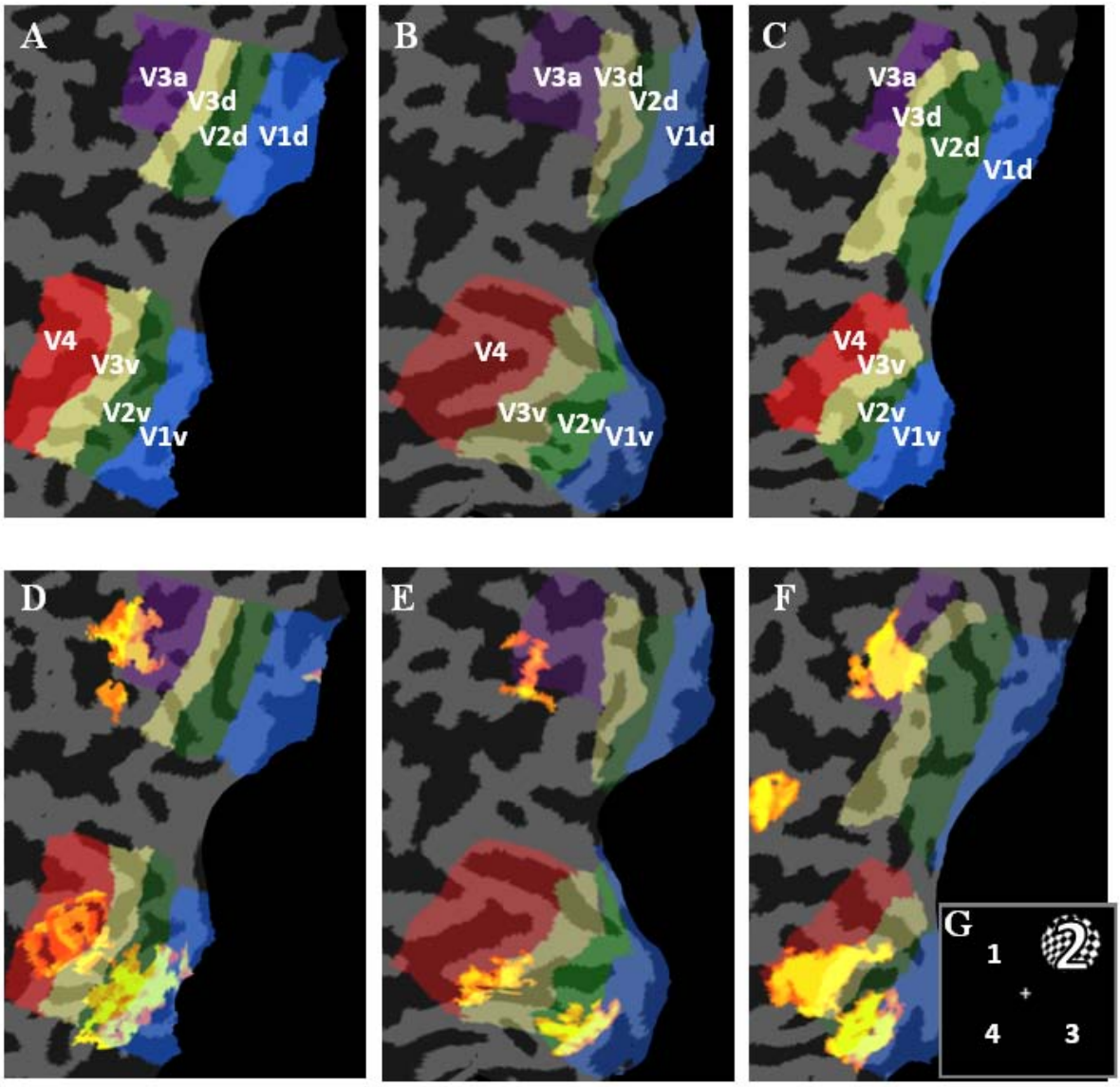

Figure 5 - Visual field maps and ROI localizers

The upper row shows visual field maps defined on the basis of the retinotopic mapping , displayed on the flattened representation of the left visual cortex for Participant 1, 2 and 3 (**A**, **B** and **C**, respectively). The encoding position for the upper right visual quadrant identified during the localizer session (Position 2, as illustrated in **G**), for Participant 1, 2 and 3 are shown on the flattened occipital cortex of the left hemisphere in

figure **D**, **E** and **F**, respectively. Activation shown is for the contrast Pos2>Pos1,Pos 3 & Pos4, $p < 0.01$ (FDR).

Average ROI-size, nr of voxels

	Visual area				
	V1	V2	V3	V4	V3a
Participant 1	92	104	44	27	40
Participant 2	117	88	56	15	36
Participant 3	88	133	92	29	47

Figure 6 - Average ROI-sizes

Because the ventrally (V4), and dorsally (V3a) confined visual areas both contain a full hemifield representation, each position were represented by five unique ROIs. Since the ROIs derived from the FDR < 0.01 criterion varied in size (Figure 6) for the different visual areas, we exported the time series to the BVQXtools (v.0.8) extension for MatLab, and calculated two additional sets of ROIs representing the 50 and 100 most spatially selective voxels for each position in each visual area. These three sets of ROIs where used in the further analysis of the two fMRI-experiments.

4 Analysis

4.1 Psychophysical data analysis

All behavioral data analysis were conducted on single subject level using SPSS (version 15; SPSS Inc., USA). The estimated difference thresholds, 75% hit rate at an ISI of 3,000 ms and 85% hit rate at an ISI of 9,000 ms, were averaged within the conditions and used as percentage change in each participants' F1-F2 and S1-S2 discrimination task, respectively.

For the two psychophysical experiments, paired samples t-tests were used to compare accuracy scores from the F1-F2 task over conditions. The two tests of interest were: Accuracy when Mask/Sample Grating Ratio = 2 vs. Accuracy when Mask/Sample Grating Ratio = 1, and Accuracy when Mask/Sample Grating Ratio = 0.5 vs. Accuracy when Mask/Sample Grating Ratio = 1.

4.2 fMRI data analysis

We performed statistical analysis at single subject level based on the General Linear Model as implemented in BrainVoyager QX v2.2 (Brain Innovation, Maastricht, The Netherlands; Goebel, et al., 2006). Both experiments were analyzed using two models; the first model did not separate between error trials on the F1-F2 task, while the second model only included correct trials in the analysis. The trials were modeled using a box-car function, which was then convolved with the canonical double-gamma HRF function, as implemented in Brain Voyager QX. We chose to model trials as blocks instead of using an event-related approach because stimuli within the trials were not jittered, thus the individual contributions of the different stimuli on the BOLD-response within a trial cannot be separated (see discussion for an elaboration on this decision).

For fMRI Experiment 1, the first model was specified using 8 regressors; the two masking conditions (Mask/Sample Ratio = 1, Mask/Sample Ratio ≠ 1) across the four stimulus positions, with onsets at stimulus S1, and a duration of 9.5 seconds, thus lasting to the offset of S2. For experiment 2, this model was created in two versions, one representing the F1-F2-task in the upper visual field, and the other representing the F1-F2-task in the lower

visual field. In each of the versions 8 regressors were specified; 4 regressors started with the onset of a sample stimulus, F1, and lasted until the offset of stimulus F2, thus having a duration of 3.5 seconds, and were separated over the two masking conditions and two positions. In addition the model contained 2 regressors modeled as events representing the onset of the mask stimulus (S1) at each position, and two similar regressors representing onset of the last stimulus (S2). The second model was similar to the first, except that errors trials were represented with separate regressors. Thus, the second set of models representing Experiment 1 and Experiment 2 contained 16 and 12 regressors, respectively.

For Experiment 1, ROI-time series were combined applying a weighted average of all 4 positions within a visual area, producing a single time series consisting of 3560 data points per visual area. For Experiment 2, due to the stimulation of two different positions in a trial (one in the upper and one in the lower visual field), the time series from the visual areas representing the upper and the lower visual field were averaged and analyzed separately. A set of *t*-contrasts were defined *a priori* for each experiment and each model, testing whether the BOLD response to a full trial (Experiment 1), or to stimulus F1 and F2 (Experiment 2), was lower in the memory masking condition (MSR $\neq$ 1) compared to the MSR = 1 condition. When testing the second model, only correct response trials were included in the contrast. Univariate statistical testing was performed separately on the three sets of localizer-derived ROIs for each visual area. Since the statistical analyses were performed on single time courses, no corrections for multiple comparisons were necessary.

For visualization of the data and to assess the temporal aspects of the BOLD responses across masked and non-masked trials, event related averaging (AVG) plots were generated for each experimental run, separated across conditions and visual areas V1-V4/V3a for each participant. This was based on extracted time courses for each of the ROIs derived from the localizer runs when applying the FDR < 0.01 threshold. The baseline was calculated as the average of the intensity values at the onset of S1 (Experiment 1), or F1 (Experiment 2), and the two preceding TRs in a run. The calculation of the signal change was performed with this average value following the formula: *percent signal change = value - average baseline for run / average baseline for run*. This yielded one AVG plot pr. position pr. experimental run pr. visual area for each participant. All positions and experimental runs where then collapsed, resulting in an average AVG plot for each visual area in each of the three participants.

5 Results

5.1 Behavioral results

The estimated 75% discrimination thresholds calculated from the QUEST procedure at an ISI of 3,000 ms were: ±11.7% (2.0%) for Participant 1; ±11.3% (2.0%) for participant 2; and ±12.7% (3.5%) for participant 3. At an ISI of 9,000 ms the estimated 85% discrimination thresholds were: ±17.7% (2,9%) for Participant 1; ±18.7% (0,8%) for Participant 2.; and ±20.0% (1,7%) for Participant 3 (each value is the average over 3 sessions; standard error of the mean (SEM) in parenthesis). These calculated thresholds were used as differences in the two tasks in the experiments.

The results from psychophysical Experiment 1, where all stimuli in a trial were presented at the same position, are shown in Figure 7B. The hit rate values show averages over 3 sessions. Visual evidence is significant at the $t(2) < 2.92$ level ($p < 0.05$, one-tailed p-value), indicating a clear memory masking effect on the F1-F2 discrimination task from both higher and lower masks (S1), relative to the sample stimulus (F1). The behavioral results from fMRI Experiment 1 (similar to psychophysical Experiment 1) is shown in Figure 7C, and replicates the findings from Psychophysical Experiment 1. Visual evidence is significant at the $t(3) < 2.35$ level ($p < 0.05$, one-tailed p-value). The hit rate values are averages of accuracy scores on the F1-F2 discrimination task over 4 sessions. Hit rates in the masking condition Mask Sample grating Ratio (MSR) $\neq 1$ (MSR = 2 and MSR = 0.5 collapsed) were significantly lower in all participants than the hit rates in the MSR = 1 condition

The results from Psychophysical Experiment 2, in which the masking stimulus S1 and the sample stimulus F1 were presented in opposite visual quadrants, are presented in Figure 8B. The hit rate values are averages over 3 sessions. No significant memory masking effect was observed on the accuracy of the F1-F2 task for Participant 1 and Participant 3, nor for Participant 2 when the mask had a lower spatial frequency than the sample stimulus (MSR = 0.5). Participant 2 did however show a significantly lower accuracy on the F1-F2 task $t(2) < 2.92$ level ($p < 0.05$, one-tailed p-value) when a mask of higher spatial frequency preceded the sample stimulus, compared to when S1 and F1 had identical spatial frequencies, however the masking effect is weaker compared to the same Mask/Sample Grating Ratio in experiment 1. The behavioral results from fMRI Experiment 2 (similar to psychophysical Experiment 2) is

shown in Figure 8C, and replicates the results from Psychophysical experiment 1, showing no masking effect for Participant 2 and 3, and a significant but weaker masking effect for Participant 3($t(3) < 2.35$, $p < 0.05$, one-tailed p-value).

One might question the use of one-tailed t-tests, however it can be justified by our priori assumptions about the direction of the mask's effect on the discrimination thresholds, since we are replicating a well documented effect. Also, it is unusual to do statistical analysis at all on psychophysical N=1 designs, as the visual evidence is usually considered sufficient. For example Lalonde & Chaudhuri (2002), and Magnussen & colleagues (1991) using similar approaches as ours, did not present statistical evidence in addition to the visual evidence.

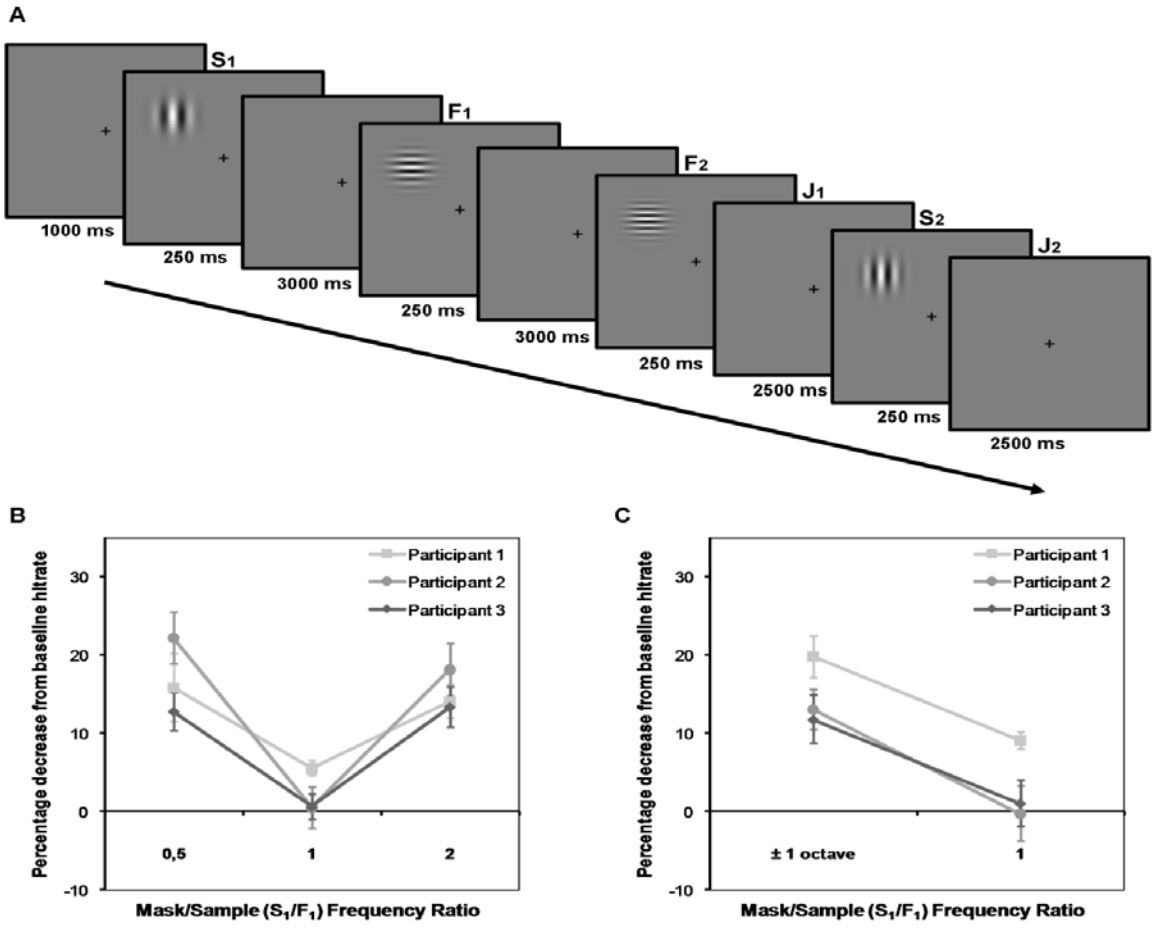

Figure 7 - Behavioral results for Experiment 1

(A) Overview of stimulus presentation in experiment 1, in which all stimuli were presented to the same position (**B**) Behavioral results from psychophysical experiment 1. Data points represent the individual hit rate on the F1-F2 discrimination task for the different Mask/Sample Grating Ratio conditions. Each data point is an average from 3 sessions. Error bars represents the standard error of the mean (SEM) over the 3 sessions. A clear memory masking effect on the F1-F2 discrimination task is seen both when the mask has a higher and a lower spatial

frequency relative to the sample stimulus (**C**) Behavioral results from fMRI experiment 1, here conditions with Mask/Sample Grating Ratio = 0.5 and 2 are collapsed. The findings replicate those seen in Psychophysical Experiment 1.

.

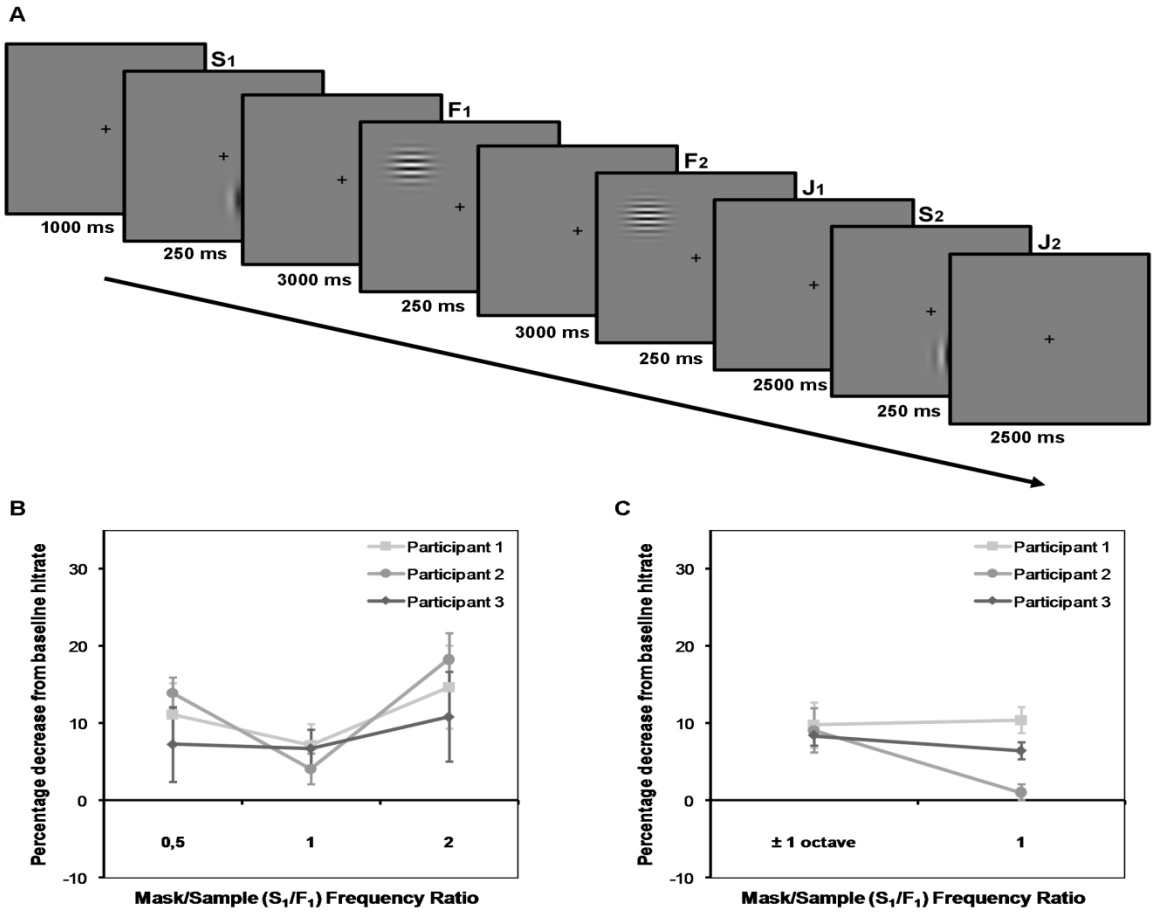

Figure 8 - Behavioral results for Experiment 2

(A) Overview of stimulus presentation in experiment 2, where the mask and sample were presented to different positions **(B)** Behavioral results from psychophysical experiment 2. Data points represent the individual hit rate on the F1-F2 discrimination task for the different Mask/Sample Grating Ratio conditions. Each data point is an average from 3 sessions. Error bars represents the standard error of the mean (SEM) over the 3 sessions. No significant masking effect was found, except for in Participant 2 when the mask had a higher spatial frequency **(C)** Behavioral results from fMRI experiment 2, here conditions with Mask/Sample Grating Ratio = 0.5 and 2 are collapsed. The behavioral results from fMRI Experiment 2 replicates the findings from Psychophysical Experiment 2, showing no or weaker memory masking effect.

5.2 fMRI-experiments

All fMRI data analyses were conducted on intensity time courses extracted from each participant's individual ROIs in early visual areas V1, V2, V3, V4 and V3a.

For Experiment 1, the contrast of interest was between the measured BOLD amplitudes for the full train of stimuli in a trial in the two main conditions: MSR = 1 and MSR ≠ 1. In V1, with localizer-derived V1 ROIs thresholded at FDR < 0.01, all participants showed a significantly lower response in the MSR ≠ 1 condition when all trials were included in the analysis (Model 1: Participant 1, $t = 3.0$, $p < .01$; Participant 2, $t = 2.2$, $p < .05$; Participant 3, $t = 5.3$, $p < .01$). Event-related averages representing the V1 time series of the two conditions are depicted in Figure 9 (B,E,H). Similar results were also found when only trials in which the participants successfully discriminated F2 from F1 were included in the analysis (Model 2: Participant 1, $t = 2.7$, $p < .01$; Participant 2, $t = 2.6$, $p < .01$; Participant 3, $t = 4.7$, $p < .01$). Finally, the analyses of both models produced the same significant results in all participants when the V1 ROIs analyzed consisted of the 50 or 100 most spatially selective voxels for each position. The other visual areas (V2-V3a) did not show a similar consistent differential pattern of activity across participants. No significant differences were found in Participant 1 for any of the models across the different ROI definitions. Participant 2 showed significant masking effects in V2 for both models when the V2 ROIs were defined as the 50 most spatially selective voxels (Model 1 & Model 2: $t > 2.8$, $p < .01$). Participant 3 showed significant masking effects in V4 (Model 1 & Model 2: $t > 3.0$, $p < .01$) and V3a (Model 1 & Model 2: $t > 4.0$, $p < .01$) for the FDR < 0.01 ROI definition, and similar effects were observed when analyzing the 50 or 100 most spatially selective voxels for each position. Importantly, no significant effects were found in the analyses of Model 2, in which only correct response trials were included, that were not found in the analyses of Model 1. Event-related averages representing the time series of the two conditions in visual areas V2-V3a are depicted in Figure 10 (B,E,H).

For Experiment 2, the contrast of interest was between the measured BOLD amplitudes to stimulus F1 and F2 in the two main conditions: MSR = 1 and MSR ≠ 1. Due to stimulation across the horizontal meridian in a trial in Experiment 2, the ROIs coding for the upper and lower visual field were analyzed separately (see the *Methods* part). In V1, none of the participants showed any significant differences between the conditions in any of the

models, nor with any of the different criterions for defining the V1 ROI. Event-related averages representing the V1 BOLD time series after the onset of stimulus F1 are depicted in Figure 9 (C,F,I). As with Experiment 1, no consistent differential pattern of activity across participants was found in the analysis of the visual areas V2-V3a. Participant 1 did not show any significant differences across models or ROI definitions, neither in ROIs coding for the upper visual field, nor for the lower visual field. Participant 2 showed a significantly lower pattern of activity in the MSR ≠ 1 condition in parts of V4 representing the lower visual field (Model 1 & Model 2: $t > 2.4$, $p < .05$), but the opposite pattern of activity was evident as a trend in the ROIs coding for the upper visual field (Model 1 & Model 2: $t < -1.7$, $p < .08$). Similar opposing trends were observed in the V3a ROIs of Participant 2. Participant 3 did however show a consistent pattern of lower activity in V4 for the MSR ≠ 1 condition; significant in the ROIs coding for the upper visual field (Model 2: $t > 2.3$, $p < .05$), and visible as a trend in the ROIs coding for the lower visual field (Model 2: $t > 1.6$, $p < .10$). This trend was consistent when the V4 ROIs were defined as the 50 most spatially selective voxels for a position, and for ROIs defined based on a FDR < 0.01 criterion, but was only seen in the analysis of Model 2, i.e. when only correct responses trials were included as regressors of interest. An effect was also found in Participant 3 in the V3a ROIs representing the upper visual field (Model1 & 2: $t > 2.4$, $p < .05$), for all ROI definitions, but no corresponding effect was found when stimuli were presented to the lower visual field. Event-related averages representing the time series in visual areas V2-V3a after the onset of stimulus F1 are depicted in Figure 10 (C,F,I).

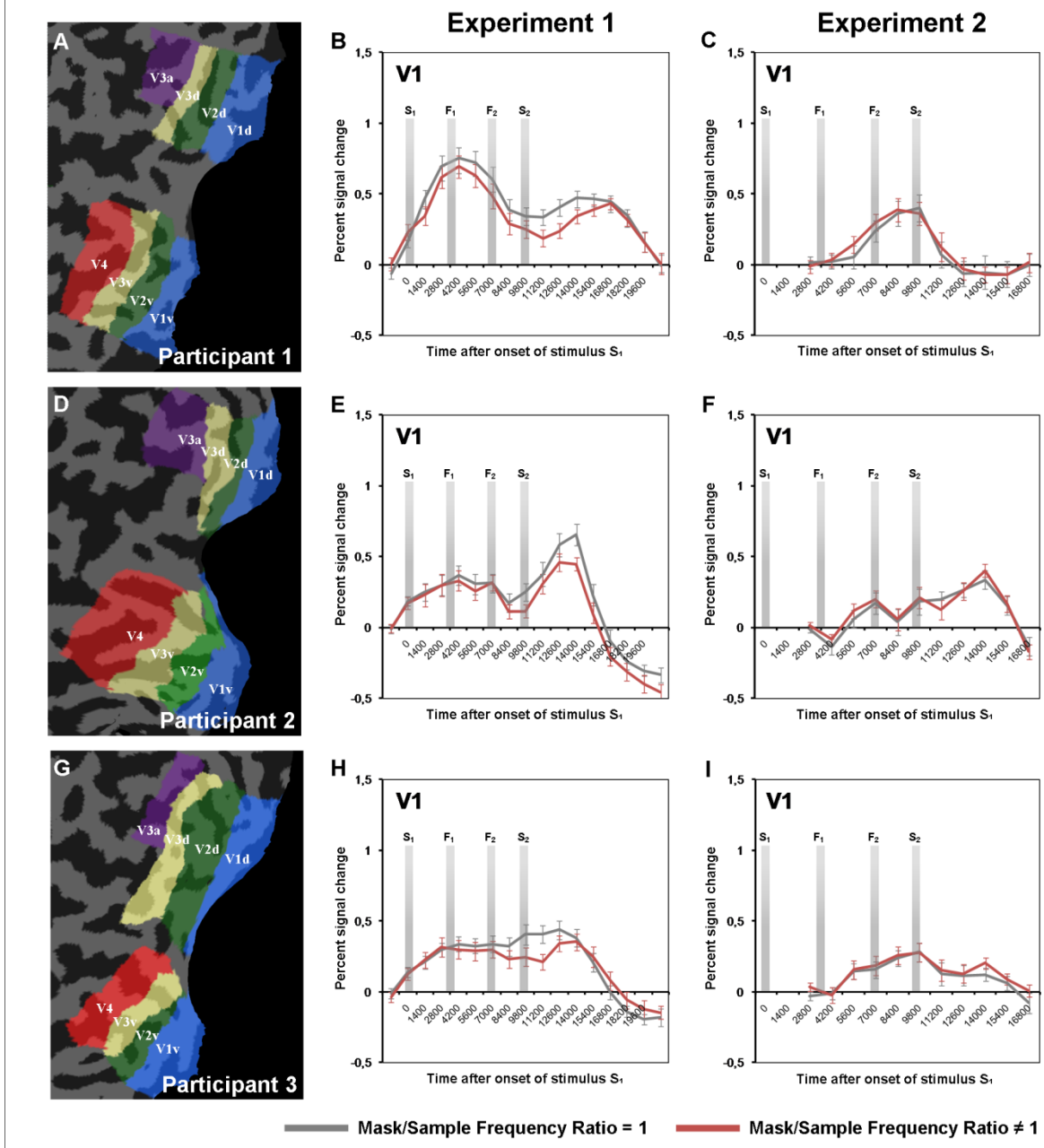

Figure 9 - Event related averaging - V1

(A, D, G) All analysis were done separately for the different visual areas, here shown on the flattened visual cortical surface for Participant 1, 2 and 3, respectively. **(B, E, H)** Event related averaging in visual area V1 for Experiment 1, for Participant 1, 2 and 3, respectively **(C, F, I)** Event related averaging in visual area V1 for Experiment 2. The Y-axis represents %-signal change from baseline, which was calculated as the average of the TR at onset of S1 and the two previous TRs. The X-axis represents time, with gray bars illustrating onsets for the four stimuli within the trials. The gray and red line represent the Mask/Sample Grating Ratio (MSR) conditions MSR=1 and ≠1, respectively. Each data point is an average of all trials in the given conditions in all 4 positions and all experimental runs. Error bars represents the standard error of the mean (SEM).

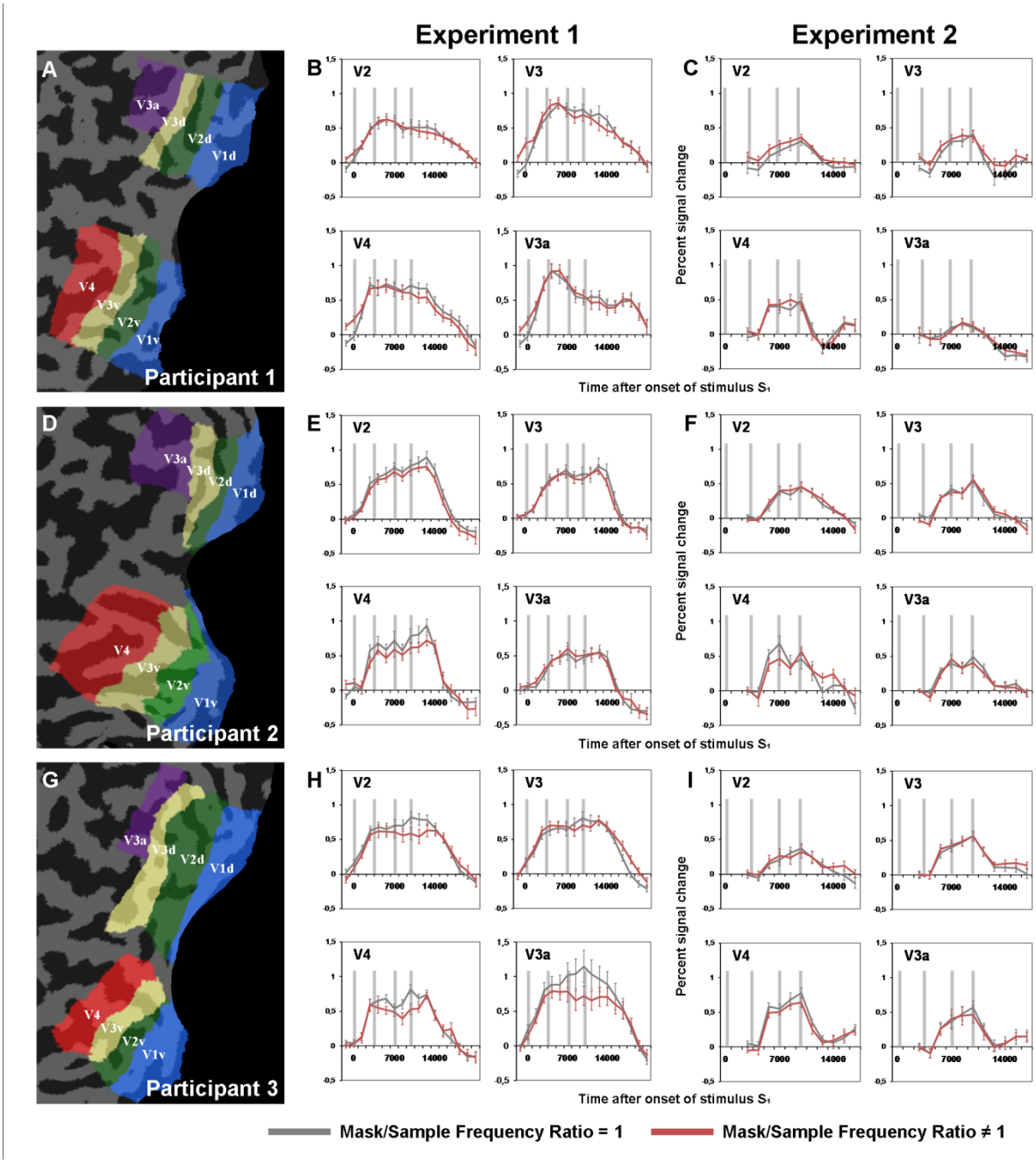

Figure 10 - Event related averaging - V2-V4/V3a

(A, D, G) All analysis were done separately for the different visual areas, here shown on the flattened visual cortical surface for Participant 1, 2 and 3, respectively. **(B, E, H)** Event related averaging for visual areas V2-V4/V3a in experiment 1. **(C, F, I)** Event related averaging for visual areas V2-V4/V3a in experiment 2. The Y-axis represents %-signal change from baseline, which was calculated as the average of the TR at onset of S1 and the two previous TRs. The X-axis represents time, with gray bars illustrating onsets for the four stimuli within the trials. The gray and red line represent the Mask/Sample Grating Ratio (MSR) conditions MSR=1 and ≠1, respectively. Each data point is an average of all trials in the given conditions in all 4 positions and all experimental runs. Error bars represents the standard error of the mean (SEM).

6 Discussion

Our primary finding is that V1-activity is modulated by the memory masking effect. In trials in which the mask differed from the sample with ±1 octave, the discrimination thresholds were elevated on the behavioral level in all participants, consistent with earlier studies on memory masking of spatial frequencies (Bennett & Cortese, 1996; Lalonde & Chaudhuri, 2002; Magnussen et al., 1991). This behavioral effect on accuracy in the delayed discrimination task is reflected in the BOLD-response in V1-voxels representing the retinotopic area coding for the stimulus position, producing a weaker signal across the trial when the mask differs from the sample compared to trials where the spatial frequency of the mask and sample is the same. As can be seen in Figure 9 (B,E,H) this effect starts to occur (i.e. the time series separate reliably) 5 to 6 seconds after the presentation of the sample stimulus (F1), an observation that fits well with the temporal delay inherent in the BOLD response (Logothetis & Wandell, 2004). It is important to stress that the two conditions only varied in the spatial frequency ratio between the mask and the sample stimulus (MSR = 1, MSR = 0.5, or MSR = 2), and that the participants were unaware of this relationship until the presentation of the sample stimulus. Our interpretation of this finding is that neurons at the level of V1 is recruited in visual working memory for spatial frequencies, and that the reduction in the BOLD response to trials in which the mask differs from the sample is a result of cross-channel inhibition between channels coding a narrow range along the spatial frequency dimension, as suggested by recent models of low-level visual working memory (Magnussen, 2000; Magnussen et al., 2009; Pasternak & Greenlee, 2005). This in accordance with the findings by Baumann and colleagues (2008) that V1 is activated when representations for spatial frequencies are extracted across orientations during delayed discrimination, and also in agreement with the studies reporting dimension-specific activation patterns for remembered features in V1 in delayed discrimination tasks using MVPA (Harrison & Tong, 2009; Serences et al., 2009). These studies did not find any corresponding memory related activation when analyzing BOLD amplitudes during the memory delays. Our findings support the idea that the reported lack of correspondence between the findings using univariate and multivariate analysis, results from an increase in the neural population coding the values of the remembered dimension and a simultaneous decrease in the neural population coding for different values along the dimension, thereby canceling each other out on the voxel level. In our study these interactions manifest as a relative weakening of the BOLD signal in

masked trials. In effect, the two conditions in our memory masking paradigm produce an amplitude contrast which is measurable using univariate analysis approaches – resulting from the sample stimulus activating neurons from a suppressed population in one condition (MSR $\neq 1$), and neurons from a population unaffected by the mask stimulus in the other condition (MSR = 1). In addition to the masking effects in V1, we also found masking effects in V2 in one of the participants, and in V4 and V3a in a second participant. These effects are not consistent across the participants, and so it is difficult to interpret their significance with respect to the memory masking effect. Inconsistencies concerning the involvement of extrastriate areas in low-level memory has also been reported in earlier studies (e.g. Harrison & Tong, 2009), and might reflect different strategies that participants employ during the task. The inconsistent results might also depend on the position-selectivity of voxels included in the ROI-analysis. For some of the extrastriate areas the ROIs defined on the basis of the FDR threshold were quite small, perhaps leading to a higher susceptibility to noise in these areas compared to the larger V1-ROIs. We therefore analyzed both models for the two experiments using ROIs based on the most selective 50 and 100 voxels in each visual area. V1 was the only area which showed a consistent masking effect in all three participants in both models and across the different ROI-sets.

Our second finding concerns the spatial extent of the memory representations for spatial frequencies. On a behavioral level masking produces significant decreases in accuracy when the mask and sample are presented to the same retinotopic area. However, when the mask and sample were presented to different areas, these effects are weaker or absent. This supports the idea that low-level memory representations of spatial frequencies are retained in retinotopically organized areas, and that masking occurs from lateral interactions between local neurons, as one would expect if each part of the visual field is analyzed by a full set of spatial frequency channels in the retinotopically organized visual cortex. Experiment 2 shows that there is still some residual effect of masking, indicating that processes that are not retinotopically confined might also be affected by the masking manipulation, but the lack of significant differences in V1 BOLD-activity in Experiment 2 support the interpretation that retinotopically organized areas are involved in retention of low-level features, and that masking mainly is a local effect of lateral inhibition. The reported results by Ester and colleagues (2009) suggest that sensory recruitment is global rather than retinotopically confined. However, there is an important difference with respect to the specificity of the representations that are probed in the current study and those decoded using MVPA. A

somewhat problematic aspect of using decoding to study low-level memory representations is the low differential resolution of the classification procedures, limited by the number of categories the classifier algorithms have been trained on. Consequentially, the decoding in the abovementioned study was performed on trial differences that were easy to categorize; between trials where orientations of gratings differed with 90 degrees. However, within trials the difference between sample and test was much smaller; the discrimination task was performed on the psychophysical 75% discrimination threshold. It is therefore difficult to interpret whether the decoded content reflect the high fidelity memory representations that underlies these discriminations, or other sensory recruitment processes such as feature-based attention and visual imagery (Kosslyn et al., 2001; Slotnick et al., 2005). For example: feature-based attention has been shown to effect global modulations of neuronal populations in the visual cortex, even in absence of visual stimuli (Saenz, Buracas, & Boynton, 2002; Serences & Boynton, 2007) There might therefore be different processes involved at the global and retinotopic level in the delayed discrimination tasks; at the global level neurons tuned to the attended dimension are modulated to increase sensitivity for the task-relevant aspect of the stimuli, while the high fidelity memory representation is confined retinotopically. Ong and colleagues (2009) showed that separation elevates thresholds for delayed discrimination of direction of motion. They investigated different degrees of separation at different eccentricities, and found that there is a critical separation of stimuli where discrimination performance drops. Importantly, this drop appears not to be gradual, but to occur when the distance between test and sample exceeds the receptive field sizes of V5 neurons at a given eccentricity. Similar results were reported by Zaksas and colleagues (2001). Both suggest that the transfer of information regarding the sample stimulus to new units of neurons with appropriate receptive fields coding the test stimulus introduces noise in the population-coding of the retained information, thereby reducing discrimination performance. The same conclusion can be drawn regarding orientations and spatial frequencies. Danilova and Mollon (2003) report discrimination thresholds for delayed discrimination of these features when sample and test are spatially separated, that are elevated compared to those reported by studies presenting the sample and test to the same location (e.g. Magnussen et al., 1991). Our result also suggest that high fidelity representations for spatial frequencies are confined locally rather than globally, since the masking effect is attenuated or absent when spatially separating mask and sample, and is also reflected in the

measured BOLD responses in V1, where we do not find significant differences between masking conditions when mask and sample are presented to different retinotopic positions.

A challenge in brain imaging studies concerning perceptual memory is to disentangle attentional and memory related processes. Studies have shown that attention modulates activity in the visual cortex, even at the level of V1 (Offen et al., 2009; Posner & Gilbert, 1999). It has also been suggested that visual attention and short-term memory are closely linked in their neural underpinnings, where attention is thought to keep memory representations alive via top-down signals that modulate neurons coding for behaviorally relevant stimulus properties, and acting as a rehearsal mechanism (Awh & Jonides, 2001; S. K. Ungerleider & Leslie, 2000). Visual attention seems also to be involved in the maintenance of representations in low-level working memory. In a delayed discrimination task for spatial frequencies, Lalonde and Chaudhuri (2002) presented a masker stimulus in the beginning of a trial, rather than between the sample and test stimulus. When the ISI between mask and sample was 3 seconds, they found masking effects consistent with previous findings. But when the ISI was increased to 10 seconds, the masking effect disappeared. In a follow-up experiment, the participants were now required to remember the mask, and compare it to a fourth grating at the end of each trial. Now the mask interfered with the task also in the 10 second ISI condition. This suggests that attentional processes are involved in the coding of the remembered stimulus features. A third experiment showed that when the discrimination task of the masker and mask test was changed from spatial frequencies to orientations, the masker stimulus no longer interfered with the main discrimination task at either ISI conditions. Since differences in orientations has been shown not to affect memory masking for spatial frequencies, it seems that attention to a specific dimension of the stimulus is required for encoding it in memory, and at the same time prevents the encoding other task-irrelevant dimensions. Even though attentional processes are involved in the encoding of stimuli in our experiment, we are looking at differential BOLD-responses to the masked and unmasked trials, which we interpret as reflecting interference related to mnemonic processes. One could argue that this differential activation stems from attentional effects, i.e. that the participants are able to detect when the mask and sample share the same spatial frequency and then devote more attentional resources in the encoding of the sample stimulus. However, we only find a consistent memory masking effect in V1 in experiment 1. Even though the spatial separation of stimuli increases discrimination thresholds, subjects are still able to discriminate spatial frequencies with relatively high precision across distances (Danilova & Mollon, 2003). The

differences between mask and sample in the masked trials in our study were above this threshold, so a conscious detection of their shared or differing spatial frequencies should be equally possible in both our experiments. This argues against a differential attentional effect across the masked and unmasked trials explaining our results.

The mask and sample stimuli within a trial were always presented with orthogonal orientations. The reported masking effect on spatial frequencies occurs, as reported by earlier studies (Bradley & Skottun, 1984; Magnussen et al., 1998) across differences in another dimension. The difference in orientation was mainly used to avoid effects of repetition priming in trials where the spatial frequencies were the same (Mask/Sample Grating Ratio = 1), since repetition priming is known to produce increased BOLD-signals in V1 for the second presentation of the primed stimulus (Slotnick & Schacter, 2004, 2006). In our study this would show up as increased activity in the trials where mask and sample shared the same spatial frequency, however because of the difference of orientations it is unlikely that the effects we observe reflect repetition priming. Supporting this, the abovementioned studies reported that the BOLD-increase from repetition priming did not show any relationship with behavioral performance, unlike in our study where the modulation of the BOLD-signal correlates with discrimination accuracy. At short intervals, priming effects might be replaced with the effect of neural adaptation; a decrease in the response of neural populations coding for the repeated stimulus (Krekelberg, Boynton, & van Wezel, 2006). Adaptation has been shown to affect BOLD-signal in the visual cortex for successive presentation of sinusoidal grating patterns (Boynton & Finney, 2003), however only at short ISIs. At ISIs of above 3 seconds the effect was shown to be minimal. Also, our short presentation times (250ms) should ensure that adaptation effects are minimal, as presentation times under 1000 ms has been shown to produce only negligible adaptation effects as measured by fMRI in early visual areas (Fang, Murray, Kersten, & He, 2005). Anyhow, the effects of neural adaptation would be the opposite of what is observed in our study, where the BOLD response is weakened when the mask differs from the sample stimulus.

Our approach of analyzing the trials as blocks might not be optimal, as it makes us unable to make assumptions regarding different stages of the memory processes by analyzing responses to individual stimuli within a trial. The standard approach to event-related designs is to randomly intermix the contrasted events, in order to obtain separate BOLD response

estimates for events spaced closely in time (Ruge, Goschke, & Braver, 2009). However, since the memory masking paradigm requires a fixed order of the stimuli (a so-called compound trial), and due to the sluggishness of the BOLD-response, the measured BOLD signal change to one event within the trial cannot be separated from preceding events in the trial. There are different approaches to solve this problem. One is to separate trials using variable intervals with long durations. The second approach, which allow for shorter ISIs, is to include partial trials (Motes & Rypma, 2009; Ruge et al., 2009). In our experiment the trial is comprised of 4 stimuli, so three different sets of partial trials would have to be included to obtain separate estimation for all events. Implementing any of these methods is difficult to do mainly for three reasons; first, it introduces uncertainty to the task. Factors of uncertainty is known to raise delayed discrimination thresholds (Magnussen, Greenlee, & Thomas, 1996). Second, it would require an enormous number of additional partial trials, making an already time consuming study too time consuming to be compatible with an MR-setting. Third, partial trials might produce unwanted effects, e.g. change subject's strategies since it's not necessary to remember the stimuli after all in a large portion of the trials (20-30% of the trials have to be partial, and we would have to multiply this number by 3). Because of these constraints, our experiment is not suitable to answer questions related to the role of V1 at different stages in the memory process. Since the interacting representations within a trial is separated in time, we interpret the BOLD signal changes as relating to memory processes, but we are unable to differentiate between retention and retrieval processes. Magnussen and colleagues (2009) have proposed that the memory trace resides at a higher stage in the processing hierarchy during the retention interval, recruiting V1-neurons during retrieval of the memory representation. This would suggest that the attenuated V1 BOLD-response we observe is a secondary effect related to processes in extrastriate areas. However, we did not find any consistent memory masking effect in extrastriate visual areas. Here our results concerning spatial frequencies parallel those reported by Serences and colleagues (2009) which fail to reliably decode the features orientation and color in extrastriate areas during delay intervals. Bennet and Cortese (1996) showed that the mechanism of memory masking is selective to perceived rather than retinal frequency. One can speculate that there are processes involved in the computation of size and shape constancies (e.g. V4; Dobbins, Jeo, Fiser, & Allman, 1998; Schiller, 1995) - which then recruits appropriate spatial frequency coding neurons in V1 via feedback mechanisms - that are not subject to the cross-channel interaction in the memory masking phenomenon. There are indeed studies indicating that retinotopic representations in

V1 can be modulated by top-down signals related to perceived size (Fang, Boyaci, Kersten, & Murray, 2008; Murray, Boyaci, & Kersten, 2006). Our experiment would then fail to detect such processes in the extrastriate cortex.

6.1 The neural mechanism of memory masking

The traditional account of the memory masking effect has been through lateral inhibition between feature specific channels, each representing a range of values along a continuum for a given dimension. Such feature specific channels for low-level features were first hypothesized on the basis of studies showing selective adaptations to spatial frequencies (Blakemore & Campbell, 1969; Campbell & Robson, 1968) and orientation (Bradley, Switkes, & De Valois, 1988). Physiological studies of non-human primates and cats later discovered that V1-neurons sensitive to different spatial frequencies and orientations are arranged in the primary visual cortex in a systematic fashion, in which each part of the visual field is analyzed by a complete set of such channels (reviewed by De Valois & De Valois, 1990). Evidence for cross-channel inhibition came from studies on cats were it was demonstrated that non-optimal spatial frequencies (DeValois & Tootell, 1983) and orientations (Morrone, Burr, & Maffei, 1982) reduced activity in V1-neurons coding for specific frequencies and orientations, and this inhibition was later shown to be GABA-mediated (Morrone, Burr, & Speed, 1987). In a study using human participants, Greenlee and Magnussen (1988) demonstrated cross-channel inhibition using a sequential adaptation paradigm; participants were adapted to high contrast stimuli with a spatial frequency of 4 c/deg, which was temporally interleaved with another adapting grating that either shared or differed in spatial frequency. The two adapting stimuli were shown sequentially at a rate of 0.5 Hz for 8 minutes. Detection thresholds for a low-contrast stimulus with a spatial frequency of 4 c/deg measured after adaptation were shown to be elevated, and the largest increase in detection threshold was seen when both interleaved adapting stimuli shared the same spatial frequency. Minimum elevation was observed when the second adapting stimuli differed from the first with 1½ octave. The attenuated adaptation when the interleaved adapting stimuli differed was interpreted by the authors as cross-channel inhibition; neurons coding for different frequency ranges inhibit each other and prevent the build-up of the adaptation after-effect.

The first memory masking experiment (Magnussen et al., 1991) was interpreted in light of this evidence, and several studies replicating the effect have given results consistent with the interpretation of cross-channel inhibition as the mechanism involved in the memory interference (Bennett & Cortese, 1996; Lalonde & Chaudhuri, 2002; McKeefry et al., 2007). The interpretation was based on the assumption that masking can occur for above-threshold stimuli, which now has been confirmed (Meese & Hess, 2004). There is an ongoing debate as to which cross-channel inhibition is the mechanism behind masking, and because of failure to find evidence for the suppression effect in single-cell studies, some have proposed alternative models (Priebe & Ferster, 2008). Recent studies have however found evidence for independent population coding for orientation with cross-orientation suppression in non-human primates (MacEvoy, Tucker, & Fitzpatrick, 2009). Using single cell recording and optical imaging to record neural responses to superimposed gratings, they found that this suppression was confined to cortical locations that had receptive field sizes encompassing both components of the super-imposed grating. Similar findings have been reported for cats, in addition to indirect evidence from studies measuring visual evoked potentials (VEPs) in human participants (Busse, Wade, & Carandini, 2009). The interpretation of the memory masking effect based on a local cross-channel inhibition process thus holds. Studies have also shown that decrease in BOLD-signal correlates with decrease of neuronal activity in the visual cortex of non-human primates (Scmuel et al., 2006), suggesting that the BOLD-differences we observe are the result of lateral inhibition in the memory masking trials.

6.2 Sensory recruitment - a philosophical note

The idea of sensory recruitment during working memory is interesting both in the perspective of understanding the basic functions of the brain's perceptual and memory systems, but also in the philosophical sense, concerning how specialized subsystems in the brain tuned to different stimulus features give rise to a unified percept, a dilemma known as the binding problem (Crick & Koch, 2003). Some have been looking for the place in the brain where "it all comes together", what the philosopher in cognitive sciences Dennett (1991) has called the Cartesian theater (after Descartes, who proposed the pineal gland as the mediator between the material brain, and the soul). The idea that somewhere in the brain there is an internal viewer, watching the show, is a fallacy because it just postpones the problem of how the perceptual processes give rise to our percepts, leading to an infinite regress of internal

viewers. We need to understand the brain in terms of distributed processes, without a centralized "supreme command" where all information is sent and integrated. From a Darwinian perspective the primary function of perception is guiding behavior (Shettleworth, 2010). Neural tissue is metabolically expensive, so the evolution of the brain has been subject to conflicting pressures; the need to minimize metabolic cost, and the importance of generating adaptive behavior (Niven & Laughlin, 2008). When taken into consideration, we should not expect the brain to transfer information that has already been discriminated and represented by one system, to a second system for re-representation (Dennett, 1991). This is a global assumption as to how the brain represents information, and it says that as soon as a discrimination of stimuli gives actionable information, it will be used to guide behavior, without spending additional time or brain resources to do the representing twice (for the pleasure of an internal viewer). In line with this perspective, and as a specific hypothesis regarding how perceptual information is represented in the brain, the sensory recruitment hypothesis proposes that the systems evolved for accomplishing on-line perceptual representing, also are involved in representing information about percepts in the absence of the sensory input that gave rise to them (Postle, 2006). This does not mean that percepts and memory representations cannot be abstracted to a degree in which perceptual details are not present, this is certainly often the case, for example in the phenomenon of gist perception (Wolfe, 1998). Further, it does not imply that perception and memory represent identical processes. But it implies that memory representations are linked in some way to the systems involved in the perceptual coding of the information that is to be retained, without the need of re-representation in a separate higher level memory system.

6.3 Concluding remarks

We have conducted the first fMRI-study on the memory masking phenomenon. The study has provided evidence for sensory recruitment at the level of V1 in working memory for spatial frequency information, a finding which is consistent with the sensory recruitment model of working memory. We thereby complement several recent studies that have decoded feature specific activation patterns in V1 in low-level working memory, by showing the involvement of V1 using a univariate approach. Our experimental design did not allow us to separate the different stages of the memory process. A possible follow-up of this study is therefore to investigate the specific mechanisms of retention and retrieval processes, to

determine the exact stage in which the memory representations interact. The degree of involvement of extrastriate areas in processing of perceived size of spatial frequencies also remains unanswered, and calls for further investigations of the memory masking phenomenon.

References

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences, 5*(3), 119-126.

Awh, E., Jonides, J., Smith, E. E., Buxton, R. B., Frank, L. R., Love, T., et al. (1999). Rehearsal in spatial working memory: Evidence from neuroimaging. *Psychological Science, 10*(5), 433.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829-839.

Bartolomeo, P. (2002). The Relationship Between Visual Perception and Visual Mental Imagery: A Reappraisal of the Neuropsychological Evidence. *Cortex, 38*(3), 357-378.

Baumann, O., Endestad, T., & Magnussen, S. G., M. W. (2008). Delayed discrimination of spatial frequency for gratings of different orientation: behavioral and fMRI evidence for low-level perceptual memory stores in early visual cortex. *Experimental Brain Research, 188*(3), 363-369.

Bennett, P., & Cortese, F. (1996). Masking of spatial frequency in visual memory depends on distal, not retinal, frequency. *Vision Research, 36*(2), 233-238.

Bisley, J. W., Zaksas, D., Droll, J. A., & Pasternak, T. (2004). Activity of neurons in cortical area MT during a memory for motion task. *Journal of Neurophysiology, 91*(1), 286.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology, 203*(1), 237-260.

Boynton, G., & Finney, E. (2003). Orientation-specific adaptation in human visual cortex. *Journal of Neuroscience, 23*(25), 8781.

Bradley, A., & Skottun, B. C. (1984). The effects of large orientation and spatial frequency differences on spatial discriminations. *Vision Research, 24*(12), 1889-1896.

Bradley, A., Switkes, E., & De Valois, K. (1988). Orientation and spatial frequency selectivity of adaptation to color and luminance gratings. *Vision Research, 28*(7), 841-856.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436.

Busse, L., Wade, A., & Carandini, M. (2009). Representation of Concurrent Stimuli by Population Activity in Visual Cortex. *Neuron, 64*(6), 931-942.

Campbell, F., & Robson, J. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology, 197*(3), 551.

Constantinidis, C., Franowicz, M., & Goldman-Rakic, P. (2001). The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nature Neuroscience, 4*(3), 311-316.

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

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences, 7*(9), 415-423.

d'Avossa, G., Tosetti, M., Crespi, S., Biagi, L., Burr, D., & Morrone, M. (2007). Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nature Neuroscience, 10*(2), 249-256.

D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1481), 761.

D'Esposito, M., Postle, B., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. *Experimental Brain Research, 133*(1), 3-11.

Danilova, M., & Mollon, J. (2003). Comparison at a distance. *Perception, 32*(4), 395-414.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial vision*. Oxford: Oxford University Press.

Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little, Brown and Company.

DeValois, K. K., & Tootell, R. B. (1983). Spatial-frequency-specific inhibition in cat striate cortex cells. *The Journal of Physiology, 336*(1), 359.

Dobbins, A., Jeo, R., Fiser, J., & Allman, J. (1998). Distance modulation of neural activity in the visual cortex. *Science, 281*(5376), 552.

Druzgal, T., & D'esposito, M. (2003). Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. *Journal of Cognitive Neuroscience, 15*(6), 771-784.

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex, 7*(2), 181.

Ester, E., Serences, J., & Awh, E. (2009). Spatially global representations in human primary visual cortex during working memory maintenance. *Journal of Neuroscience, 29*(48), 15258.

Fang, F., Boyaci, H., Kersten, D., & Murray, S. (2008). Attention-dependent representation of a size illusion in human V1. *Current Biology, 18*(21), 1707-1712.

Fang, F., Murray, S., Kersten, D., & He, S. (2005). Orientation-tuned fMRI adaptation in human visual cortex. *Journal of Neurophysiology, 94*(6), 4188.

Furmanski, C. S., & Engel, S. A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience, 3*(6), 535-536.

Gardner, J. L., Merriam, E. P., Movshon, J. A., & Heeger, D. J. (2008). Maps of Visual Space in Human Occipital Cortex Are Retinotopic, Not Spatiotopic. *Journal of Neuroscience, 28*(15), 3988-3999.

Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human brain mapping, 27*(5), 392-401.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron, 14*(3), 477-485.

Goldman-Rakic, P. S., & Leung, H. C. (2002). Functional architecture of the dorsolateral prefrontal cortex in monkeys and humans. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 85–95). New York: Oxford University Press, USA.

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

Greenlee, M. W., & Magnussen, S. (1988). Interactions among spatial frequency and orientation channels adapted concurrently. *Vision Research, 28*(12), 1303-1310.

Greenlee, M. W., & Thomas, J. P. (1993). Simultaneous discrimination of the spatial frequency and contrast of periodic stimuli. *Journal of the Optical Society of America. A, Optics and image science, 10*(3), 395-404.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience, 27*(1), 649-677.

Harrison, S., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458*(7238), 632-635.

Haxby, J. V., Petit, L., Ungerleider, L. G., & Courtney, S. M. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *NeuroImage, 11*(5), 380-391.

Irwin, D. E. (1996). Integrating Information across Saccadic Eye Movements. *Current Directions in Psychological Science, 5*(3), 94-100.

Ishai, A., Haxby, J., & Ungerleider, L. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *NeuroImage, 17*(4), 1729-1741.

Issa, N. P., Rosenberg, A., & Husson, T. R. (2008). Models and Measurements of Functional Maps in V1. *Journal of Neurophysiology, 99*(6), 2745-2754.

Kaas, A., Weigelt, S., Roebroeck, A., Kohler, A., & Muckli, L. (2009). Imagery of a moving object: The role of occipital cortex and human MT/V5+. *NeuroImage*.

Kanwisher, N., McDermott, J., & Chun, M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302.

Klein, I., Dubois, J., Mangin, J., Kherif, F., Flandin, G., Poline, J., et al. (2004). Retinotopic organization of visual mental images as revealed by functional magnetic resonance imaging. *Cognitive Brain Research, 22*(1), 26-31.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. [10.1038/35090055]. *Nature Reviews Neuroscience, 2*(9), 635-642.

Kraft, A., Schira, M. M., Hagendorf, H., Schmidt, S., Olma, M., & Brandt, S. A. (2005). fMRI localizer technique: efficient acquisition and functional properties of single retinotopic positions in the human visual cortex. *Neuroimage, 28*(2), 453-463.

Krekelberg, B., Boynton, G., & van Wezel, R. (2006). Adaptation: from single cells to BOLD signals. *Trends in Neurosciences, 29*(5), 250-256.

Lakha, L., & Wright, M. J. (2004). Capacity limitations of visual memory in two-interval comparison of Gabor arrays. *Vision Research, 44*(14), 1707-1716.

Lalonde, J., & Chaudhuri, A. (2002). Task-dependent transfer of perceptual to memory representations during delayed spatial frequency discrimination. *Vision Research, 42*(14), 1759-1769.

Lee, B., & Harris, J. (1996). Contrast transfer characteristics of visual short-term memory. *Vision Research, 36*(14), 2159-2166.

Levy, R., & Goldman-Rakic, P. S. (1999). Association of Storage and Processing Functions in the Dorsolateral Prefrontal Cortex of the Nonhuman Primate. *Journal of Neuroscience, 19*(12), 5149-5158.

Logothetis, N., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*(6843), 150-157.

Logothetis, N., & Wandell, B. (2004). Interpreting the BOLD signal. *Annual Review of Physiology, 66*, 735-769.

MacEvoy, S. P., Tucker, T. R., & Fitzpatrick, D. (2009). A precise form of divisive suppression supports population coding in the primary visual cortex. *Nature Neuroscience, 12*(5), 637-645.

Magnussen, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences, 23*(6), 247-251.

Magnussen, S. (2009). Implicit visual working memory. *Scandinavian Journal of Psychology, 50*(6), 535-542.

Magnussen, S., & Greenlee, M. W. (1992). Retention and disruption of motion information in visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(1), 151-156.

Magnussen, S., & Greenlee, M. W. (1997). Competition and sharing of processing resources in visual discrimination. *Journal of Experimental Psychology. Human Perception and Performance, 23*(6), 1603-1616.

Magnussen, S., & Greenlee, M. W. (1999). The psychophysics of perceptual memory. *Psychological research, 62*(2), 81-92.

Magnussen, S., Greenlee, M. W., Asplund, R., & Dyrnes, S. (1990). Perfect visual short-term memory for periodic patterns. *European Journal of Cognitive Psychology, 2*(4), 345-362.

Magnussen, S., Greenlee, M. W., Asplund, R., & Dyrnes, S. (1991). Stimulus-specific mechanisms of visual short-term memory. *Vision Research, 31*(7-8), 1213-1219.

Magnussen, S., Greenlee, M. W., Baumann, O., & Endestad, T. (2009). Visual perceptual memory – Anno 2008. In L. Bäckman & L. Nyberg (Eds.), *Memory, aging and the brain* (pp. 53–75). Hove, UK: Psychology Press.

Magnussen, S., Greenlee, M. W., & Thomas, J. P. (1996). Parallel processing in visual short-term memory. *Journal of Experimental Psychology. Human Perception and Performance, 22*(1), 202-212.

Magnussen, S., Idås, E., & Myhre, S. H. (1998). Representation of orientation and spatial frequency in perception and memory: A choice reaction-time analysis. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 707-718.

McKeefry, D. J., Burton, M. P., & Vakrou, C. (2007). Speed selectivity in visual short term memory for motion. *Vision Research, 47*(18), 2418-2425.

McKeefry, D. J., Burton, M. P., Vakrou, C., Barrett, B. T., & Morland, A. B. (2008). Induced Deficits in Speed Perception by Transcranial Magnetic Stimulation of Human Cortical Areas V5/MT+ and V3A. *Journal of Neuroscience, 28*(27), 6848-6857.

Meese, T. S., & Hess, R. F. (2004). Low spatial frequencies are suppressively masked across spatial scale, orientation, field position, and eye of origin. *Journal of Vision, 4*(10), 843-859.

Mohr, H., Linder, N., Linden, D., Kaiser, J., & Sireteanu, R. (2009). Orientation-specific adaptation to mentally generated lines in human visual cortex. *NeuroImage, 47*(1), 384-391.

Moro, V., Berlucchi, G., Lerch, J., Tomaiuolo, F., & Aglioti, S. (2008). Selective deficit of mental visual imagery with intact primary visual cortex and visual perception. *Cortex, 44*(2), 109-118.

Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 216*(1204), 335-354.

Morrone, M. C., Burr, D. C., & Speed, H. D. (1987). Cross-orientation inhibition in cat is GABA mediated. *Experimental Brain Research, 67*(3), 635-644.

Motes, M., & Rypma, B. (2009). Working Memory Component Processes: Isolating BOLD Signal-Changes. *NeuroImage, 49*(2), 1933-1941.

Murray, S., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience, 9*(3), 429-434.

Nakamura, K., Matsumoto, K., Mikami, A., & Kubota, K. (1994). Visual response properties of single neurons in the temporal pole of behaving monkeys. *Journal of Neurophysiology, 71*(3), 1206.

Nassi, J., & Callaway, E. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience, 10*(5), 360-372.

Nilsson, T., & Nelson, T. (1981). Delayed monochromatic hue matches indicate characteristics of visual memory. *Journal of Experimental Psychology, 7*(1), 141-150.

Niven, J. E., & Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *Journal of Experimental Biology, 211*(11), 1792-1804.

Norman, K., Polyn, S., Detre, G., & Haxby, J. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences, 10*(9), 424-430.

O'Craven, K., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience, 12*(6), 1013-1023.

Offen, S., Schluppeck, D., & Heeger, D. (2009). The role of early visual cortex in visual short-term memory and visual attention. *Vision Research, 49*(10), 1352-1362.

Ong, W., Hooshvar, N., Zhang, M., & Bisley, J. (2009). Psychophysical evidence for spatiotopic processing in area MT in a short-term memory for motion task. *Journal of Neurophysiology, 102*(4), 2435.

Pasternak, T., Bisley, J. W., & Valkins, D. (2003). Visual information processing in the primate brain. In M. Gallagher & R. J. Nelson (Eds.), *Biological psychology* (pp. 130-185). New York: Wiley.

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience, 6*(2), 97-107.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437-442.

Posner, M. I., & Gilbert, C. D. (1999). Attention and primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 96*(6), 2585-2587.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139*(1), 23-38.

Priebe, N., & Ferster, D. (2008). Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron, 57*(4), 482-497.

Qiu, A., Rosenau, B. J., Greenberg, A. S., Hurdal, M. K., Barta, P., Yantis, S., et al. (2006). Estimating linear cortical magnification in human primary visual cortex via dynamic programming. *NeuroImage, 31*(1), 125-138.

Ranganath, C. (2006). Working memory for visual objects: Complementary roles of inferior temporal, medial temporal, and prefrontal cortex. *Neuroscience, 139*(1), 277-289.

Ranganath, C., Cohen, M. X., Dam, C., & D'Esposito, M. (2004). Inferior Temporal, Prefrontal, and Hippocampal Contributions to Visual Working Memory Maintenance and Associative Memory Retrieval. *Journal of Neuroscience, 24*(16), 3917-3925.

Regan, D. (1985). Storage of spatial-frequency information and spatial-frequency discrimination. *Journal of the Optical Society of America A, 2*(4), 619-621.

Ruge, H., Goschke, T., & Braver, T. (2009). Separating event-related BOLD components within trials: The partial-trial design revisited. *NeuroImage, 47*(2), 501-513.

Saenz, M., Buracas, G., & Boynton, G. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience, 5*(7), 631-632.

Schiller, P. (1995). Effect of lesions in visual cortical area V4 on the recognition of transformed objects. *Nature, 376*, 342–344.

Serences, J., & Boynton, G. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron, 55*(2), 301-312.

Serences, J., Ester, E., Vogel, E., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20*(2), 207.

Shettleworth, S. (2010). *Cognition, evolution, and behavior*. New York: Oxford University Press.

Shmuel, A., Augath, M., Oeltermann, A., & Logothetis, N. (2006). Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nature Neuroscience, 9*(4), 569-577.

Simons, D., & Rensink, R. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences, 9*(1), 16-20.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience, 7*(6), 664-672.

Slotnick, S. D., & Schacter, D. L. (2006). The nature of memory related activity in early visual areas. *Neuropsychologia, 44*(14), 2874-2886.

Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual Mental Imagery Induces Retinotopically Organized Activation of Early Visual Areas. *Cerebral Cortex, 15*(10), 1570-1583.

Slotnick, S. D., & Yantis, S. (2003). Efficient acquisition of human retinotopic maps. *Human brain mapping, 18*(1), 22-29.

Sperling, G. (1960). The information available in brief visual presentation. *Psychological Monographs, 74*(11), 1-29.

Super, H., Spekreijse, H., & Lamme, V. (2001). A neural correlate of working memory in the monkey primary visual cortex. *Science, 293*(5527), 120.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. *Analysis of visual behavior, 549*, 586.

Ungerleider, S. K., & Leslie, G. (2000). Mechanisms of Visual Attention in the Human Cortex. *Annual Review of Neuroscience, 23*(1), 315-341.

Vaughan, J. T., Garwood, M., Collins, C. M., Liu, W., DelaBarre, L., Adriany, G., et al. (2001). 7T vs. 4T: RF power, homogeneity, and signal-to-noise comparison in head images. *Magnetic Resonance in Medicine, 46*(1), 24-30.

Vogels, R., & Orban, G. (1986). Decision processes in visual discrimination of line orientation. *Journal of Experimental Psychology: Human Perception and Performance, 12*(2), 115-132.

Wandell, B., Dumoulin, S., & Brewer, A. (2007). Visual field maps in human cortex. *Neuron, 56*(2), 366-383.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113-120.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology General, 131*(1), 48-64.

Wolfe, J. M. (1998). Visual memory: What do you know about what you saw? *Current Biology, 8*(9), R303-R304.

Zaksas, D., Bisley, J. W., & Pasternak, T. (2001). Motion Information Is Spatially Localized in a Visual Working-Memory Task. *Journal of Neurophysiology, 86*(2), 912-921.

Zaksas, D., & Pasternak, T. (2006). Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *Journal of Neuroscience, 26*(45), 11726.

Zeki, S., & Shipp, S. (1988). The functional logic of cortical connections. *Nature, 335*(6188), 311-317.