Substituting Objects from Consciousness: A Review of Object Substitution Masking

Stephanie C. Goodhew^{a,b}, Jay Pratt^b, Paul E. Dux^c and Susanne Ferber^{b,d}

^a*Research School of Psychology, The Australian National University, Australia*

^b*Department of Psychology, University of Toronto, Canada*

^c*School of Psychology, The University of Queensland, Australia*

^d*Rotman Research Institute, Canada*

Corresponding Author: Stephanie C. Goodhew

Address: Research School of Psychology (Building 39)

The Australian National University,

Canberra, ACT, 0200, Australia

Email: s.c.goodhew@gmail.com

Phone: +61 2 6125 5545

Abstract

Object substitution masking (OSM) occurs when a sparse (e.g., four dot), temporally-trailing mask obscures the visibility of a briefly-presented target. Here, we review theories of OSM: Those which propose that OSM reflects the interplay between feedforward and feedback/reentrant neural processes, those which predict that feedforward processing alone gives rise to the phenomenon and theories that focus on cognitive explanations, such as object updating. We discuss how each of these theories accommodates key findings from the OSM literature. In addition, we examine the relationship between OSM and other visual-cognitive phenomena, including object correspondence through occlusion, change blindness, metacontrast masking, backward masking, and visual short-term memory. Finally, we will examine the level of processing at which OSM impairs target perception. Collectively, OSM appears to reflect the conditions under which the brain confuses two visual events for one when they are encoded with low spatiotemporal resolution, due to processing resources being otherwise occupied.

Keywords: object substitution masking; attention; re-entrant processing; consciousness;

When observing a visual scene, at any point in time, we are only conscious of a small fraction of the available information. Given this limitation, it is important to understand the mechanisms that determine what and when we become conscious of stimuli, and the nature of processing that occurs in the absence of explicit awareness. There exists a long and rich history in experimental psychology of using visual masking to explore such issues. Visual masking refers to conditions where the visibility of one to be reported visual stimulus (the target) is obscured by the presentation of another stimulus (the mask) that appears in close spatiotemporal proximity and does not require report. Object substitution masking (OSM) is a recent discovery in the field of visual masking, and is an ideal tool for exploring questions about visual awareness because it selectively impairs the extent to which an individual becomes conscious of a visual stimulus without the extent of image-level degradation that other forms of visual masking induce. This review, therefore, focusses on the determinants and consequences of visual awareness in OSM.

Here we integrate a diverse range of literature to discern what OSM can tell us about perceptual consciousness and visual cognition in general. Specifically, we discuss the role of feedforward versus feedback processing in giving rise to OSM, whether the phenomenon is better characterised as reflecting object updating or object substitution, the similarities and differences between OSM, and other forms of masking and OSM's relationship to other visual-cognitive phenomena. In addition, we review evidence regarding the role of attention in masking. Our aim is to address controversies in the OSM literature in depth, offer insights into how some of them may already be resolved, and how the field might go about resolving those that remain. We begin by describing the emergence of OSM from the visual masking literature.

Visual masking

It was first documented in the late nineteenth century that the presentation of a stimulus (the mask) has the potential to interfere with the perception of a prior target stimulus when they shared a common spatial location, and that this impairment provides a metric of the time taken to recognise the target item (Baxt, 1871; translated in Baxt, 1982). This finding has inspired considerable research interest, chiefly as a tool for exploring the mechanisms that determine which stimuli enter perceptual consciousness and give rise to our rich visual experience. In the early twentieth century, it was discovered that masking impairments extended to conditions where the target and mask occupied distinct (but usually contiguous) spatial locations if they had similar contours – which came to be known as metacontrast masking (Alpern, 1952, 1953; Breitmeyer & Ögmen, 2006; Werner, 1935). Strikingly, metacontrast masking produces a U-shaped masking function, with strongest masking typically occurring when the stimulus onset asynchrony (SOA) between the target and mask is between 50-100ms, with weaker or absent masking at shorter and longer SOAs (Breitmeyer & Ögmen, 2006). This counterintuitive departure from monotonicity has been the impetus for many investigations into masking and has been the topic of considerable theorising (Breitmeyer & Ganz, 1976; Bridgeman, 1971, 1977; Francis, 1997; Kahneman, 1967; Weisstein, 1968).

Over a decade ago, Di Lollo, Enns and Rensink (2000) discovered a form of masking that displayed properties that did not fit with conventional accounts: A mask that consisted of only four dots could impair the visibility of a target with which it shared little physical similarity and never spatially overlapped. This masking was characterised by the target and mask having a common-onset, and the key variable was duration of the mask stimulus trailing target offset (mask duration), rather than the SOA between the target and mask (Di Lollo, et al., 2000; Enns & Di Lollo, 2000). This impairment of target visibility was termed

OSM (Enns & Di Lollo, 1997), and it reflected an exciting new development emerging from a rich tradition of visual masking research in experimental psychology and cognitive neuroscience. OSM is clearly distinct from backward masking by pattern, structure, or noise, where the mask spatially overlaps the target after a given SOA (Breitmeyer & Ögmen, 2006), since the four dots never even partly cover the target. OSM is more similar to metacontrast masking and, initially, there was debate about whether it was an original form of masking or merely a variant of metacontrast (Breitmeyer & Ögmen, 2000). The evidence since then, however, has demonstrated that while there are some similarities, metacontrast and OSM are experimentally dissociable. For example, Chakravarthi and Cavanagh (2009) demonstrated that ‘crowding’, the deficit in identifying a target in the periphery owing to overly-close adjacent flankers (Pelli & Tillman, 2008), could be prevented when metacontrast masks were applied to the flankers but was not when four-dot masks were applied. This suggests that metacontrast masking has an earlier locus of suppression than crowding, whereas OSM has a later locus of suppression (Chakravarthi & Cavanagh, 2009). These similarities and differences between OSM and metacontrast masking and the functional properties of the visual system which they tap will be discussed in greater detail subsequently in this review. Presently, however, we will focus on a description of the properties that define OSM.

Figure 1. An illustration of the standard object substitution masking paradigm. The task is to identify whether the gap in the target ‘C’ (the object inside the four dots) is on the left or the right side. Target perception is impaired when the offset of the four-dot mask is delayed (above) relative to when the mask offsets simultaneously with the target (below).

Since its discovery, it has been argued that one of the defining hallmarks of OSM is that the impairment in report depends on attentional resources being occupied during the presentation of the target array. Typically masking is not obtained unless distractors are presented simultaneously with the target (Di Lollo, et al., 2000) or unless attentional resources are occupied with another demanding task (e.g., mental arithmetic) (Dux, Visser, Goodhew, & Lipp, 2010). Masking is attenuated or eliminated when attention is pre-cued to the location of the target in the array (Di Lollo, et al., 2000; Neill, Hutchison, & Graves, 2002). Similarly, attention to the mask plays a role in producing masking, as deliberate

attention to the mask exacerbates OSM (Tata & Giaschi, 2004). However, like the distribution of attention during target exposure, it is not necessary in order to produce masking (Neill, et al., 2002).

Recently, however, it has been argued that the limited masking obtained with the presentation of a target stimulus without distractors is the result of a ceiling effect constraining performance in this condition, and the purported hallmark interaction between the number of distractors and mask duration (Di Lollo et al., 2000) is an artefact of this constraint (Argyropoulos, Gellatly, Pilling, & Carter, in press). These authors demonstrated that when performance is shifted away from ceiling, although both set-size (number of stimuli in the target array, target and distractors) and mask duration impaired target perception, their effects were additive rather than interactive. Argyropoulos et al. also obtained significant masking when the target was the only item in the target array (no distractors; set-size one). However, even in this condition, the target could appear in any one of four possible locations, meaning that there was still an element of uncertainty about the spatial location of the target, and consequently attentional resources would have to be distributed across the potential target locations. Therefore, while it appears that OSM is not necessarily defined by an interaction between set-size and mask duration, there is still no conclusive evidence against the idea that OSM depends on the prevention of focussed attention on the target.

Collectively, the results above demonstrate that there are several key physical characteristics that define OSM: 1) Attention is distributed during target exposure; 2) target exposure is relatively brief ($\leq 100\text{ms}$); 3) the mask continues to remain physically present after target offset (or at the very least is visible after target offset; (Enns & Di Lollo, 1997; Lleras & Moore, 2003)); and 4) it is characterised by a loss of visual awareness of the target. In this next section, we specify what we mean by ‘awareness’, describe some of the

prominent theories put forward to explain OSM, and then analyse their ability to account for the key findings in the literature.

Awareness

For the present purposes, when we discuss visual awareness or consciousness (used interchangeably in this review) in relation to masking, we specifically refer the ability to report either the stimulus' presence, or identity (depending on the task). This reportability definition of visual awareness is what Block (1995, 2011) has referred to as *access awareness*. Block argues that this is an incomplete definition of consciousness because there is also the possibility for *phenomenal consciousness* – the subjective sense of being aware of something without being able to explicitly report it. Note that this is different from the failure of visual awareness that we can measure in masking – being able to report stimulus presence (detection) without being able to accurately report identity. Instead, phenomenal consciousness is the purported situation of being aware without even being able to report that one is aware of something at any level. So when we present the findings from visual masking studies that use reportability as a gauge of visual awareness, and infer that failures of reportability mean failures of conscious processing, there remains the possibility that participants could have this phenomenal consciousness state of the visual stimulus.

The notion of a present but unreportable awareness is, however, not testable. We can neither rely on reportability, nor establish an implicit measure. For example, in order to develop a physiological signature of phenomenal consciousness, it would first be necessary to establish its presence versus absence and then correlate this with an accompanying physiological response. But this would require measuring phenomenal consciousness to begin with, but since explicit report does not suffice as a measure, this approach gets us no further ahead. In fact, it is not clear that there is any appropriate reference against which to gauge

phenomenal consciousness (Kouider, de Gardelle, Sackur, & Dupoux, 2010; Kouider, Sackur, & de Gardelle, 2012). For these reasons, we will focus exclusively on access awareness.

The majority of OSM studies use identification as a measure of target awareness. For example, in Enns and Di Lollo's (1997) task, participants' were required to perform a two-alternative forced choice discrimination, where the target was a diamond with either its left or right corner cropped off. In this case, if participants could not accurately report the identity of the target, it was inferred that their visual awareness of it was impaired. A few studies, however, have also used a target detection task (e.g., Chen & Treisman, 2009) where participants make a judgement regarding whether a target is present or absent. In this task participants can err in one of two ways; a 'false alarm', in which the target was absent but they indicated they thought it was present, and a 'miss', where the target was present, but the participant responded that it was absent. Misses represent the most severe failures of awareness. If a participant makes an error on a target identification or discrimination task, then it is reasonable to assume that the participant's awareness of the stimulus item was impaired, but it is still possible (even likely) that they had a sense that *something* was fleetingly present. That is, both target discrimination and detection are valid measures of awareness of a masked target – it is just that errors on these tasks they may simply represent different levels or severities of failures of consciousness.

Of course, it is possible to correctly guess the target's identity or presence, even in the absence of an explicit awareness of the target and consequently 'target correct' responses across an experiment reflect a mix of actual target aware trials and some guesses. The errors on such forced-choice metrics, however, have a more clear-cut interpretation: Impaired visual awareness. If one is aware of the target, then they do not need to guess – they can simply respond accurately. Thus, any 'unaware' condition as isolated via forced choice target

identification or detection measures can be more confidently thought to have isolated conditions of unawareness.

One potential issue, however, with detection tasks is the issue of response criteria. That is, some participants might have a liberal threshold for what they deem sufficient sensory evidence to conclude that a target was present, whereas others might have a more conservative approach and only deem a target ‘present’ in light of much stronger sensory evidence. This issue can be mitigated by use of measures of sensitivity (such as d') which take into account response bias and allow it to be disentangled from sensitivity. But such measures only apply to aggregate measures in a condition and do not resolve the problem if trials are sorted based on responses in order to examine the fate of a masked target. An approach to dealing with this issue is to explicitly encourage one type of criterion (e.g., if you have any sense that a stimulus was there, report that it was). But still, for these reasons, some authors prefer a discrimination metric in which participants are forced to make a response about the target which removes any deliberation about when to respond.

Another potential criticism of the reportability definition of awareness, however, is that participants may have been fleetingly aware of a visual stimulus, but due to forgetting are subsequent unable to report it. This seems unlikely in visual masking paradigms, especially when the task requires report of the stimulus identity/presence within a fraction of a second of the stimulus itself. However, it remains a possibility, but it is also one that is impossible to empirically gauge, and thus, like phenomenal consciousness, this possibility appears outside the scope of investigation. In this next section, we will describe the theories put forward to explain the failure of visual awareness in OSM.

Theories of OSM

Feedback versus feedforward-only models and OSM

It was once thought that visual perception conformed exclusively to a feedforward hierarchy of processing in which areas encoded increasingly complex stimulus properties (Hubel, 1963). However, it is now widely recognised that in addition to this, the brain has extensive re-entrant connections between regions, implying that perception does not exclusively conform to a hierarchy of analysis (Bullier, 2001; Zeki, 1993). Instead, it has been suggested that conscious perception is achieved via re-entrant processing (Cudeiro & Sillito, 2006; Lamme & Roelfsema, 2000; Pascual-Leone & Walsh, 2001; Zeki, 2001). In this section we will discuss the re-entrant processing account that was first offered to explain OSM (Di Lollo et al., 2000), as well as other models of re-entrant processing and their applicability to the phenomenon. Finally, we examine the ability of exclusively feedforward models to account for OSM.

Mumford (1991) proposed one of the earliest re-entrant models of object perception. A unique aspect of this account was that it posited that the thalamus is the destination for feedback processing and the locus for integrating input from different cortical areas, thus representing the current cohesive contents of consciousness. In this framework the term ‘active blackboard’ is used to describe how the thalamus maintains the current best reconstruction of the visual world. Specifically, perceptual processing is hypothesised to begin with a feedforward sweep through early areas that registers stimulus feature information, which then continues on to higher cortical regions (e.g., inferior temporal [IT] cortex) where it activates multiple abstract representations (hypotheses) regarding object identity. From there, recurrent processing proceeds down a pathway from the cortex to the thalamus where representations are compared in parallel against sensory information still reverberating from the feedforward sweep. The model employs a “relaxation algorithm”

which chooses the hypothesis that is the closest match for the sensory information. Conscious perception of this representation then ensues (Di Lollo, 2012; Mumford, 1991).

Mumford's (1991) model was developed as a general theory of object perception and predates the discovery of OSM. As a model of object perception, it has its merits, and it is noteworthy for being ahead of its time in terms of its emphasis on re-entrant processing. Moreover, there is supporting physiological evidence, for example, activity in the (subcortical) lateral geniculate nucleus (LGN) strongly correlates with the contents of visual awareness (Wunderlich, Schneider, & Kastner, 2005). But, fMRI has revealed that OSM activates frontal regions and primary visual cortex, and the magnitude of the BOLD response changes in these areas correlated with effectiveness of OSM (Weidner, Shah, & Fink, 2006). Moreover, it is not just the physical properties of OSM that activate these regions, but rather the process of object substitution itself, since an OSM display in which the target location is cued did not elicit the same magnitude of activations as when target location was unpredictable. But it is neither just brief nor difficult-to-see targets that activate these regions, since these effects were more marked in OSM as compared to backward pattern masking. They were still present for pattern masking, and this likely reflects a residual role of object substitution in such forms of masking (Weidner, et al., 2006). This suggests that V1 is an important locus of sensory input against which perceptual hypotheses are compared, which is not accounted for by Mumford's (1991) model.

Lamme and Roelfsema (2000) popularised the notion of re-entrant processing and its critical importance in conscious object perception. According to this view, recurrent processing in early visual areas is a necessary condition for visual awareness. Whereas Mumford (1991) suggested that areas of the cortex have static functions and the thalamus is responsible for dynamic representations, according to Lamme and Roelfsema (2000) lateral and feedback connections allow cortical neurons to contribute different analyses of a visual

scene at different moments in time. Specifically, according to this model, the feedforward sweep reaches the highest levels of cortical processing within about 100ms of stimulus onset. Recurrent processing typically exerts its influence at relatively long latencies and is necessary for refining stimulus representations and rendering them accessible to visual awareness. Within this framework, a backward mask is thought to disrupt recurrent interactions between cortical areas by creating stimulus responses to the mask in lower areas that clash with those in higher areas. In support of this, backward masking has been found to leave the feedforward sweep of processing intact while selectively preventing feedback processing from extrastriate areas to V1 (Fahrenfort, Scholte, & Lamme, 2007; Lamme, Zipser, & Spekreijse, 2002).

Again, this model was developed as a general framework for visual processing rather than an explanation for OSM. However, it is consistent with the finding that an occipital P2 component is elicited about 220 ms after stimulus onset in OSM, and this component was significantly correlated with target reportability (Kotsoni, Csibra, Mareschal, & Johnson, 2007). Although the spatial resolution of ERPs is limited, the origin of this component and its timing is consistent with the notion of re-activation of primary visual cortex in order to refine the target representation, and that the greater this re-activation, the more likely the visual system will discard the perceptual representation of the target in favour of the mask-alone stimulus (Kotsoni, et al., 2007). However, while Lamme's model provides a conceptual framework for understanding the dynamics of recurrent processing and its role in conscious vision, there are some aspects unique to OSM, as opposed to backward masking, which it does not specifically address, including the role of the distribution of attention and the effect of trailing mask duration. Di Lollo et al.'s (2000) model, however, was designed to do this.

According to the re-entrant processing account of object substitution, OSM results from a conflict between abstract representations in higher level cortical areas and transient sensory input to V1 (Di Lollo, 2010, 2012; Di Lollo, et al., 2000; Enns & Di Lollo, 1997).

Under normal viewing conditions, in the first cycle of processing, input from the target array is encoded in V1, then proceeds to higher levels where tentative representations (or ‘perceptual hypotheses’) about object identity are generated. However, such perceptual hypotheses require comparison against the high resolution and spatially-specific sensory information in V1 because these high-level representations are coarse or incomplete (low resolution), sensitivity to the location of the target stimulus is reduced due to the large receptive field size of neurons in higher areas, and ambiguity about object identity is created when multiple perceptual hypotheses are activated by the feedforward sweep. The circuit then checks, via ongoing re-entrant processing, for correlations between the descending codes representing multiple perceptual hypotheses, and the ongoing pattern of sensory activity. Perceptual hypotheses with low correlations are discarded, and ultimately, the representation with the highest correlation wins the competition for consciousness. Given the brief presentation of the target in OSM, when the descending signals arrive at V1, there is a mismatch between the perceptual hypothesis that represents the target and mask and the sensory activity representing the mask alone. This yields a low correlation for the perceptual hypothesis containing the target, which is then discarded, and the hypothesis reflecting the mask-alone is ultimately consciously perceived (Di Lollo, 2010, 2012; Di Lollo, et al., 2000; Enns & Di Lollo, 1997, 2000).

This model predicts that OSM depends on set-size, owing to the number of iterations required to locate the target in the search array. If fewer iterations are required, then there are a greater number of re-entrant loops during which stimulus information about the target will still be lingering. In contrast, when a greater number of iterations are required to identify the target amongst distractors (i.e., at larger set-sizes), then the descending signals representing the target are more likely to overlap with the mask only sensory information, leading to a

lower correlation and thus the selection of the perceptual hypothesis consistent with just the mask being presented (Di Lollo, 2010; Di Lollo, et al., 2000).

Di Lollo et al. (2000) developed and tested a computational model of object substitution (CMOS) in which a cortical hypercolumn in striate cortex (V1) connected to a corresponding region in an extrastriate visual area. This model posits iterative exchanges between fine-grained high spatial resolution sensory input information and more abstract pattern information that is lacking in spatial resolution which ultimately results in conscious perception. Specifically, it stipulates three main stages of processing. Firstly, there is the input layer (*I*), which, like V1, has small receptive fields, allowing for detailed and spatially-precise coding of information. Secondly, there is an intermediate layer called the working space (*W*), thought to be part of the striate cortex, which has greater spatial resolution than *I*. Thirdly, there is the final pattern layer (*P*), purported to be in extrastriate cortex, which has large receptive field sizes, rendering it sensitive to the overarching pattern. At the onset of a stimulus, it is first encoded at *I*, then transferred to *W*, and a weighted sum of *W* and *I* are sent to *P*. *P* then outputs back to *W*, via this transfer, the pattern codes from *P* are translated into the pixel codes that predominate at *W*. Now *W* contains the pattern information that was represented in *P*, but it is in a form (pixel-based) that allows for direct comparison with the information at *I*. Then the contents of *W* are compared with the contents of *I*. Here, a “hill-climbing algorithm” searches for the highest correlation between the pattern information from *P* that has been translated into *W* on the one hand and the sensory information in *I* on the other. This is achieved over successive iterations. Importantly, the fact that *P* receives input that is a weighted sum of the contents *I* and *W* (which is being fed back the representation from *P*) means that the patterns represented in *P* change more gradually than the rapid changes in sensory input that occur in response to physical stimulation changes (Di Lollo, et al., 2000).

CMOS performed reasonably well with r^2 values between .78 and .98 being obtained when comparing the fits of the model to the psychophysical results (Di Lollo et al., 2000). However, Pöder (in press) has since pointed out that CMOS is not actually a true instantiation of re-entrant processing, but instead, is essentially an attentional-gating model. More critically, CMOS is limited in its ability to explain some important findings. Firstly, this model stipulates that set-size and mask duration interact, but when performance is unconstrained by ceiling effects, this interaction does not occur (Argyropoulos, et al., in press; Pöder, in press). Secondly, it does not readily explain why prolonged exposure of placeholder stimuli in the target array can reduce masking (Guest, Gellatly, & Pilling, 2012).

Pöder (in press) developed a model which did not incorporate re-entrant processing and was designed to improve on the limitations of CMOS. This model assumes a pre-attentive stage of processing that covers the target array, in addition to the basic mechanism of CMOS. Specifically, it assumes temporal integration of the target and mask signals, the strength of which is modulated by attention. When the two stimuli have a simultaneous offset, then the signals for both stimuli are likely to be preserved up until the object perception level. However, with the trailing mask-alone stimulus, the mask signals are present for a longer integration window, degrading the target representation and strengthening the mask representation, thereby impairing visibility of the target. After some time, when attention is focussed on the target location, the target is no longer accessible. The preattentive stage, however, means that some target information was acquired, and thus performance never declines to chance. Pöder's model provides an almost perfect fit for the behavioural data he modelled: r^2 values of .99 were obtained when compared with Di Lollo et al.'s (2000) behavioural data and .98 when compared with Luiga and Bachmann's (2008) results, pooled across polarity conditions.

The core advantage of Pöder's theory is that it dovetails nicely with the important role of attention in OSM (Dux, et al., 2010) while not stipulating an *interaction* between set-size and mask duration, which is a stumbling block for CMOS (Argyropoulos, et al., in press). That said, it still does not provide a complete account for the full gamut of OSM findings. For example, there are many object-updating effects in OSM (e.g., Hirose et al., 2007), which we will discuss in the next section, on which this model is silent. Similarly, in its current form, this model does not predict the non-monotonicity which OSM functions can produce (Di Lollo, et al., 2000; Goodhew, Dux, Lipp, & Visser, 2012; Goodhew, Visser, Lipp, & Dux, 2011a): That is, with prolonged mask exposure (e.g., 640 ms), there can be an improvement in target identification accuracy relative to intermediate mask durations (e.g., 240 ms) yielding a U-shaped function of masking across mask exposure (Goodhew, et al., 2012; Goodhew, et al., 2011a). It should be noted, however, that recovery does not universally occur. In one case when an outline square mask was used and the task for participants was to detect the presence of an unbroken ring target amongst broken ring distractors, statistically significant recovery was not obtained (Tata & Giaschi, 2004). But there are considerable individual differences in the temporal dynamics of the OSM recovery function (Goodhew, et al., 2012), and possibly averaging across participants in the Tata and Giaschi study diluted the recovery effect. Consistent with this idea, there are trends towards recovery with prolonged mask exposure in some of Tata and Giaschi's (2004) figures (see Figure 2, 8 mask condition; see Figure 4, no preview condition; Figure 6, no preview condition). Further evidence for the existence of recovery in OSM is that Di Lollo et al. (2000) also found such non-monotonicity with delayed offset of an annulus mask around a target with concurrent distractors in a trained observers, with this pattern most pronounced at smaller set-sizes (i.e., set-size 1). Perhaps, therefore, recovery is observed under particular conditions of perceptual degradation of the target, which may reflect the interplay of target

exposure duration, concurrent perceptual load, mask density/similarity to the target, and observer individual differences (e.g., working memory capacity). However, Pöder's model does not make any provision for the recovery effect.

There is a model that predicts non-monotonicity of OSM functions, as Bridgeman (2007) modified a single-layer model that relied on distributed stimulus representations to account for metacontrast masking (Bridgeman, 1971) in order to explain OSM. The model employed lateral inhibition between adjacent neurons at recurrent 30ms intervals, which is the latency of reciprocal lateral inhibition in the cat LGN (Singer & Creutzfeldt, 1970). The role of attention in masking is simulated in the model by varying the number of iterations over which neural net activity is collated. That is, fewer recurrent iterations occur for strongly attended items and a larger number for less strongly attended. This is a more plausible instantiation of attention than that employed in earlier attempts to adapt models of metacontrast masking to explain OSM, where attention was modelled via mask intensity (Francis & Hermens, 2002). While it is true that focussed attention generally reduces the effect of the trailing mask, mask intensity does not seem to capture what is meant by the nuanced construct that is 'attention', and, furthermore, it has been established that target perceptibility is largely unaffected by the relative intensity of the target versus mask (Neill, et al., 2002).

Strong evidence in favour of Bridgeman's (2007) model is that it produces U-shaped masking functions like those seen in recovery from OSM. The explanation for this pattern of masking is based on the timing of mask offset and target signals and their interaction within the model. That is, recovery occurs when the target is most closely attended (fewer iterations), preventing the mask offset transient from interfering with the target signal. But, when focussed attention is prevented (more iterations), the target and mask representations merge and produce masking. This neatly explains why Di Lollo et al. (2000) obtained non-

monotonic masking across mask duration at set-size one (where there was still spatial uncertainty about target location): Here the target was more strongly attended relative to larger set-sizes. This theory, however, does not as readily explain why Goodhew et al (2012; 2011a) obtained recovery at larger set-sizes, nor why recovery can occur with prolonged mask exposure in the absence of a mask offset prior to response delay or simply with delayed responses.

Many of the models discussed above assume that feedforward and/or recurrent processing between visual and higher-level cortical regions is an essential component of conscious perception. Another prevailing view in cognitive neuroscience stipulates that the dorsal and ventral processing streams are differentially implicated in conscious and unconscious perception (Goodale, 2008; Goodale & Milner, 1992; Goodale & Westwood, 2004). That is, the dorsal stream governs accurate visually-guided motor planning behaviour, which can be executed in the absence of conscious visual awareness. The ventral stream, in contrast, underlies conscious object perception. The classic example of this dissociation was provided by patient D.F. who suffered visual form agnosia after ventral occipital lesion: She was unable to recognise objects and yet her object-directed grasping remained intact (James, Culham, Humphreys, Milner, & Goodale, 2003). This same dissociation between visuomotor action and conscious perception has been demonstrated in OSM. That is, it has been found that normal motor reaching performance occurs even in the absence of conscious perception of the size of the target object (Binsted, Brownwell, Vorontsova, Heath, & Saucier, 2007; Heath, Neely, Yakimishyn, & Binsted, 2008).

The dorsal/ventral unconscious/conscious distinction, however, is not as absolute as initially conceived. In some of the earlier models of metacontrast masking, this interaction between the two processing channels was emphasised, especially how the magnocellular (transient) channel, which predominately innervates the dorsal stream, (Shapley, 1990),

disrupts processing of target-related information along the parvocellular (sustained) channel (which connects to the ventral stream) (Breitmeyer & Ganz, 1976). More recent models, however, have established that magnocellular input can have a facilitatory effect on object perception (Bar et al., 2006).

Breitmeyer and Ganz (1976) proposed the most influential and, for its time, the most physiologically plausible theory of metacontrast masking: The sustained-transient dual-channel account. The feedforward, dual-channel architecture explains masking as a function of time between target and mask onsets and/or offsets. Both offsets and onsets activate the faster transient channel, which inhibits the slower sustained channel carrying information about stimulus identity (the transient/sustained maps onto the magnocellular/parvocellular distinction; Sherman, Wilson, Kaas, & Webb, 1976). In metacontrast masking, the sustained activity representing the target is inhibited by the rapid mask-induced transient when the mask onsets 50-100 ms after the target. At shorter SOAs, the mask-induced transient occurs prior to the target-related sustained response, and so does not interfere with it. Similarly, at longer SOAs, the target-related sustained response is consolidated and thus impervious to the effect of the mask. While this model was initially designed to account for the U-shaped masking function that characterises metacontrast masking, it is also relevant to OSM. In OSM, there is typically constant (0 ms) target-mask SOA, and the one factor that varies is trailing mask duration, which concomitantly varies the temporal separation of mask offset from the target. This model, therefore, would predict a U-shaped function of OSM across mask duration (Breitmeyer & Ganz, 1976; Breitmeyer & Kersey, 1981; Breitmeyer & Ögmen, 2006), which has been found (Goodhew, et al., 2011a). But, the U-shaped pattern also occurs in the absence of mask offset or just with a response delay after mask offset (Goodhew, et al., 2012), which cannot be explained within the sustained-transient framework.

More recently, Tapia and Breitmeyer (2011) compared behavioural responses to known response properties of magnocellular and parvocellular neurons. Observers were presented with either a visible or invisible (rendered so by a metacontrast mask) prime arrow (pointing left or right) at different contrasts, followed by a probe arrow. Their task was to identify whether the probe arrow was pointing left or right as quickly and accurately as possible. Prime visibility was assessed in a separate identification block, and they found that the function of the effect of the prime across contrast was best approximated by the known properties of magnocellular neurons for consciously perceived visible primes, whereas the effect of the prime was best explained by the known properties of parvocellular neurons for the unconsciously perceived primes. These authors suggested that the role of the magnocellular channel in conscious vision is to generate feedback to IT areas that is necessary for rendering stimuli accessible to visual awareness.

A similar idea was previously described in more detail by Bar (2003), who proposed that ‘hypothesis generation’ in object recognition is subserved by a magnocellular input directly into prefrontal cortex, which generates a preliminary ‘guess’ about the identity of an object in advance of the feedforward sweep. This then initiates re-entrant activity from prefrontal cortex to object-related areas in IT, amplifying the sensory input in these regions that is consistent with the perceptual representation, which then facilitates rapid object recognition. In support of this, activity in orbitofrontal cortex preceded (by about 50ms) activity in more classical object-related areas (temporal cortex, including fusiform gyrus and LOC), and moreover, this activity was predictive of subsequent behavioural object recognition reported by the observer (Bar, et al., 2006). Consistent with the magnocellular’s channel preference for low spatial frequency (LSF) content of stimuli (Derrington & Lennie, 1984), both the unfiltered and low-pass spatial frequency filtered images generated equivalent activity in orbitofrontal cortex (OFC) and induced greater synchronization between

feedforward and feedback projections, whereas this was not so for high-pass filtered images. Thus, conscious perception is the not exclusive domain of the ventral cortical stream.

In sum, both the qualitative re-entrant processing account and its purported computational instantiation (CMOS) have difficulty explaining the nuances of OSM. While initial attempts to model OSM without recourse to re-entrant processing fell short of accurately operationalizing important concepts like attention (Francis & Hermens, 2002), more recent developments have provided excellent predictive value in a way that is theoretically sound (Pöder, in press). Bridgeman's (2007) model does one of the best jobs in approximating the behavioural data of OSM, but this is a model of recurrent lateral inhibition rather than feedback between different regions. However, the physiological reality is that the brain has substantial architecture for re-entrant processing, and much evidence points to this being involved in conscious perception (Lamme & Roelfsema, 2000). Re-entrant processing is therefore likely to be involved in some way in rendering the target stimulus accessible to visual awareness, it just may not be via the specific mechanisms proposed in the re-entrant processing account of OSM (Di Lollo et al., 2000). For example, it may be that the visual competition between the target and mask is resolved via recurrent lateral inhibition, and re-entrant processing is subsequently responsible for rendering either just the mask, or both the target and mask accessible to perceptual consciousness. Furthermore, we believe that in the future, models will benefit from synthesising growing knowledge of the role of the magnocellular/dorsal stream in conscious object perception (e.g, Bar, 2003; Bar, et al., 2006). For example, it may be the preattentive sweep proposed by Pöder is a magnocellular-mediated 'first guess' at target object identity.

Object-updating and object substitution

According to the object-updating account, OSM reflects the conditions in which the target and mask are integrated into a single object representation, thereby obscuring the visibility of the target. As we will discuss in this section, there is a wealth of evidence that encouraging the formation of independent object representations for the target and mask reduces OSM. But in order for this to be a compelling explanation, it needs to describe how integrating the target and mask objects would render the target unavailable to awareness, since the target and mask never share a common spatial location. It may be the case that the space inside the four dots is not treated as transparent, but part of the four-dot object in its own right. Evidence for this assertion is that when the four stimuli constituting the mask are arranged in such a way as to induce an illusory contour, masking occurred when this was stimulus appeared immediately adjacent to the target, but not when it appeared elsewhere in the display (Hirose & Osaka, 2009). Furthermore, when a stereoscope is used to induce the perception that the target and mask appear in different depth planes, masking only occurs when the mask is perceived to be in front of the target (Kahan & Lichtman, 2006). This suggests that the target representation is fused not just with four dots, but with an illusory contour created by the four-dot mask, and this obscures the visibility of the target.

In support of the role of object updating in OSM, Lleras and Moore (2003) demonstrated that OSM critically depends on the mask constituting a robust continuing object representation within the timeframe that allows the target to be integrated with it. Their design utilised *apparent motion* - where two static stimuli flashed close in space and time create the percept of motion in the absence of actual motion (Anstis, 1980). Specifically, in their experiments, the target array (presented for 17-34 ms) contained eight 'Cs' (varying in orientation), arranged in a circle, with every item surrounded by four dots. The target was signalled by being a darker shade of grey than the distractors, and participants' were instructed to report the orientation of the target C. In addition to a baseline condition

(standard simultaneous offset condition), there were also three different delayed mask offset conditions. There was a standard delayed mask offset condition; a condition where the masks always offset simultaneously with the target, and then re-appeared at a new spatial location (all masks moved together out into a circle with a larger circumference) after a duration that was conducive to the percept of apparent motion (17-34 ms; short-ISI) - and finally, there was a similar condition where the masks reappeared after a longer delay (216-233 ms; long-ISI) which was not conducive to eliciting apparent motion. Comparing target identification accuracy in each condition against the simultaneous offset baseline, it was found that masking occurred in the standard delayed offset condition and in the apparent motion condition, but not in the condition where apparent motion was not observed. This suggests that when the mask is perceived as a continuing object, it results in OSM, but when the mask appears as a new object that onsets after the target offsets, it does not.

Pilling and Gellatly (2010) recently extended Lleras and Moore's (2003) paradigm. These authors used a condition akin to that employed in short-ISI trials described above, and assessed the effect of adding another presentation of the mask dots at an outer-most location in between the two successive displays previously used in that condition. The logic here was that this should disrupt the perception that the dots were a re-appearance of the mask that was presented during the target array, and would instead encourage the percept that these were new objects. This, therefore, should allow the visual system to determine that the target and mask were separate object identities, and consequently reduce masking. This is indeed what was found: Masking was reduced with the additional presentation of the dots compared to the standard short-ISI condition (Pilling & Gellatly, 2010).

Moore and Lleras (2005) found that additional manipulations that encourage the visual system to treat the target and mask as a single versus two separate objects also affected masking. Specifically, in Experiment 1, the authors used motion to encourage either common

or separate object representations. In one condition, the mask appeared at a different location to the target, then moved past the location of the target when it was present, and then moved on to a new location after target offset (i.e., separated target and mask objects). In the other condition, after the mask onset at a different location to the target, it then appeared to move in conjunction with the target, rather than slide past its location (i.e., target and mask objects not separate). Masking was greater for the latter compared to the former condition. In Experiment 2, the independence of the target and mask objects was manipulated via coherent versus incoherent motion. The trials began with placeholders: Filled circles (rather than circles with gaps - Cs), with four-dot masks around each of them (the placeholders gave no cue as to the subsequent location of the target). Then, the arrays (both circles and dots) briefly moved. They either did so together (coherent motion), or independently of one another (incoherent motion). After the placeholders stopped moving, the target array was presented. Stronger masking occurred after exposure to motion which implied that the four dots and the circles were a coherent object compared with that after viewing a pattern of motion implying that the circles and dots were separate objects. Again, this suggests that a failure of target and mask object individuation contributes to OSM (Lleras & Moore, 2005).

Finally, in Experiment 3, the target and mask were either coloured identically or appeared in different colours to one another (Lleras & Moore, 2005). This was done on the assumption that when the target and mask were identically coloured, they should be more likely to be treated as belonging to a single object token compared with when they were differently coloured. It was found that masking was weaker when the target and mask were coloured differently. To the extent that featural similarity acts as an object-individuation cue, this is consistent with the object-updating account (Lleras & Moore, 2005). Moreover, it has been shown that rTMS applied to V5/MT+ reduces masking (Hirose et al., 2007). This dovetails with object-updating account, since V5/MT+ processes motion (Born & Bradley,

2005), which inherently involves updating the location of an object over time. Temporarily deactivating this function may reduce object-updating, and this might explain the decrease in masking.

Mask preview findings also provide converging evidence for the crucial role of object individuation computations in OSM. In the mask-preview paradigm, the target display includes four-dot masks around all items (both target and distractors). Since the masks surround all the stimuli, they are not predictive of the location of the target amongst the distractors in the subsequent array and thus do not simply serve as a precue to target location (which is also known to attenuate masking; see Di Lollo et al., 2000). Despite this, the non-predictive preview of the mask objects (and targets) reduces masking (Neill, et al., 2002). It has been hypothesised that this occurs because the preview consolidates the representation of the masks as a separate object identity, protecting against subsequent integration of the target and mask items. In addition to mask preview, prolonged exposure of the target array (in the absence of the cue indicating which item in the array is the target) decreases masking strength (Gellatly, Pilling, Carter, & Guest, 2010). Again, it has been suggested that the pre-exposure of the target array aids in the individuation the target object's representation, thus making it resistant to being confused with that of the mask (Gellatly et al., 2010). These authors have also shown that even when placeholders are used in lieu of the target and mask objects in the preview array, this reduction in masking still occurs, illustrating that effect is not just the result of the previewed stimuli being loaded into visual short-term memory (Guest, et al., 2012).

As noted above, prolonged mask exposure after the target offset can also reduce masking (Goodhew, et al., 2012; Goodhew, et al., 2011a). How would the object-updating framework explain this? As discussed above, recovery of the object representation is obtained when the mask does not offset prior to a response being made, and with a blank response

delay (Goodhew, et al., 2012). This illustrates the importance of time in the recovery effect - with sufficient processing time (untied to any further physical stimulation) the target representation can be accessed. Crucially, both in the conditions where there is a trailing mask and when the screen is blank the visual system has a period of time during which no new stimulation onsets. We suspect that this period of time allows the mask representation to be consolidated in its own right, consequently preventing it from being fused with the target.

Our suggestion is that object-updating reflects a limitation in the temporal resolution of visual encoding. This means that any manipulation which transiently increases the temporal precision of encoding should thwart OSM. One way to do this is to place visual stimuli in near-hand space. It was initially proposed that near-hand space incurs enhanced attentional processing (Reed, Betz, Garza, & Roberts, 2010; Reed, Grubb, & Steele, 2006) or reduced attentional disengagement (Abrams, Davoli, Du, Knapp, & Paull, 2008). More recently, however, there has been evidence to suggest it is the purview of upregulated contribution of the magnocellular channel – which has greater temporal resolution but poorer spatial encoding relative to its parvocellular counterpart (Gozli, West, & Pratt, 2012). Increased temporal resolution implies that the visual system should be especially likely to encode the target and mask as separate objects and it has recently been demonstrated that OSM is indeed reduced at such locations (Goodhew, Gozli, Ferber, & Pratt, in press). This is consistent with the notion that OSM reflects overzealous temporal fusion of two object identities.

It has also been suggested that the object-updating account is distinct from the object substitution account, as the former predicts that the target and mask are fused into a single object representation, whereas the latter that two separate object tokens (target and mask) compete for access to consciousness (Di Lollo, et al., 2000; Guest, Gellatly, & Pilling, 2011). Essentially, it is a question of whether the visual system encodes the target and mask as a

single or two separate events. Recently Guest, Gellatly and Pilling (2011) argued that they had found evidence in favour of object substitution. These authors found that OSM was strongest when the four dots surrounded the target (as per conventional OSM), and was weakened when the dots were instead placed overlapping the critical target feature. These authors argued that this is because OSM reflects spatial competition for consciousness between the whole target and mask object. But this finding is also easily explained by the object-updating framework. A common alignment of the target and mask stimuli (as in the standard OSM condition) would be especially conducive to them being confused as belonging to a single object identity as opposed to when they were spatially offset. Thus, the results of Guest et al. (2011) can be taken as evidence against the notion that the target and mask representations interact at an isolated featural level (e.g., Kahan & Enns, 2010), but they are entirely consistent with the existing evidence for the object-updating account. It remains to be seen, therefore, whether there are findings in OSM that can only be explained by object-substitution, and not object-updating. It is telling, however, that the neural models that fared best in predicting behavioural OSM results (Bridgeman, 2007; Pöder, in press) posit *integration* (as opposed to substitution) of the target and mask stimulus information as the core mechanism underlying OSM.

OSM and object correspondence

It has long been noted that object representations play a key role in OSM (e.g., Lleras & Moore, 2003; Moore & Lleras, 2005). But, is there a link between OSM and one of the fundamental challenges of vision in everyday life? That is, humans make several eye movements a second (Henderson & Hollingworth, 1998), and objects themselves move in the environment. This means the brain receives visual information that is constantly being interrupted and in a state of flux. Yet, despite this, our subjective sense of the world is one of stability and continuity. This demonstrates that the brain transforms the dynamic and

impoverished input into stable recognisable object representations. It would be a confusing world if the visual system treated all changes in stimulation as new objects. This would mean, for example, that making a saccade to another object in between two fixations would obliterate any recognition of persisting identity.

While maintaining object identities is important, the brain cannot indiscriminately compute continuing object identity. In some instances, objects do change: An object at one location may be replaced by another, and it is equally important to appreciate when a new object appears on the scene as to maintain object identities when they persist. Thus, making an inference of correspondence versus non-correspondence under conditions of ambiguous input is a key challenge for the visual system. Some paradigms exist for approaching this question, which are discussed below, but it is plausible that OSM cuts to the core of this process as this phenomenon involves presenting observers with two instantiations of a visual stimulus – the first of which contains the target object, as well as surrounding visual information – followed by the second which just contains the surrounding visual information. This means the brain is confronted with the need to draw an object correspondence (or non-correspondence) inference, and, we suggest, on masked trials this ambiguity is resolved in favour of a judgement that just a single object appeared throughout the presentation and consequently the target is suppressed from visual awareness. That is, the observer has no sense of having seen the target object at all, and when forced to respond, cannot accurately identify it. According to this perspective, when the target and mask are temporally fused, it is not a flaw of the system, but rather reflects the operation of a mechanism that allows us to maintain coherent object identity representations through time and across change.

Why, according to this perspective, does spatial attention play a role in OSM? This could be explained in that it serves to lower the resolution of the forms being encoded to such an extent that the target plus mask representation is more likely to be perceptually confused

with the mask-alone representation. It is already known that attention increases the spatial resolution of visual encoding (Yeshurun & Carrasco, 1998, 1999, 2008). Thus, in OSM, the fact that spatial attention cannot be readily focussed on the target would decrease any featural dissimilarity between the target and mask. This, in turn, would increase the likelihood that they would be deemed to be two instantiations of a single object.

There is another aspect of attention in OSM – that attention to the mask increases masking (Neill, et al., 2002) – which has been argued to be necessary for OSM to be observed (Tata & Giaschi, 2004). However, it is not necessary for the mask to be a singleton and ‘pop-out’ in the target array, since OSM is obtained even when four-dots surround all target locations and the target is instead signalled by appearing in a different colour to the distractors (e.g., Lleras & Moore, 2003). From the object correspondence perspective, attention to the mask likely serves two important purposes. Firstly, it enhances the strength of the mask’s stimulation, increasing the likelihood that the target instantiation, with relatively less strength, will be subject to integration rather than be treated as a discrete object. When the mask surrounding the target is not a singleton, the longer duration of the mask relative to the target likely serves the same purpose. Secondly, when integration occurs, attention to the mask ensures that it is the stimulus that dominates in the integration. Otherwise, if the representation of the mask were equally weak as the target, then integration could occur in favour of the target. That is, the target plus mask and subsequent trailing mask could be treated as a single instance of the target plus mask object rather than the mask alone stimulus. If such integration occurred, it would of course not be evident as an impairment in target visibility, but instead an impairment in mask visibility. Attention to the mask may preclude this possibility, and instead OSM reflects the integration of the target and mask into a single object representation that largely reflects the mask alone stimulus.

Relationship between OSM and other visual-cognitive phenomena

Although OSM has unique features that are not shared with other paradigms, the basic mechanisms that produce OSM may overlap with those underlying other visual-cognitive phenomena, especially other ‘failures of awareness’ (Kim & Blake, 2005).

Metacontrast masking

There are obvious superficial differences between metacontrast masking and OSM. For example, mask duration is held constant in metacontrast masking with only SOA being varied, whereas OSM conventionally involves the common onset of the target and mask (i.e., a SOA of equals zero). It has been demonstrated, however, that when mask duration is varied, metacontrast masking magnitude increases over the durations 0-160ms (Di Lollo, von Muhlenen, Enns, & Bridgeman, 2004), and over these durations OSM shows the same pattern. Furthermore, inducing a delay between the stimulus and the response can reduce metacontrast masking when the SOA between the target and mask is zero (Lachter & Durgin, 1999), and common-onset OSM shows a similar improvement with response delay (Goodhew et al., 2012). But, ultimately, as described previously, there is empirical evidence that OSM and metacontrast masking are indeed dissociable, with metacontrast masking having an earlier locus of suppression (Chakravarthi & Cavanagh, 2009).

When it comes to the role of attention, however, we might think of the distinction between OSM and metacontrast as one of degree, rather than kind. Whereas OSM is critically dependent on the dispersal of attention during target exposure, metacontrast masking is not, although it is modulated by attention. That is, metacontrast masking is obtained even with a centrally presented and attended target; in fact, this is the arrangement in which metacontrast masking is traditionally used (Alpern, 1952, 1953; Breitmeyer & Ögmen, 2006). However, the magnitude of metacontrast masking increases when spatial attention is dispersed during the presentation of the target (e.g., number of distractors in the target array) and attenuated

when rapid attention to the target is facilitated in such arrays (Shelley-Tremblay & Mack, 1999; Tata, 2002). While the label ‘metacontrast masking’ was not used, this interaction between attention and metacontrast masking was first noticed decades ago. Averbach and Coriell (1961) used a partial-report paradigm (similar to Sperling, 1960) in which a 2 x 4 array of letters was presented briefly, and after a variable delay, a visual cue was used to signal the target. When the visual cue was a circle that appeared around the location that had contained the target letter, these authors found an unexpected impairment in the visibility of the target letter that conformed to a U-shaped function across target-cue SOA (Averbach & Coriell, 1961). It is noteworthy that when attention was spread over the target array, a circle could serve as a mask for letters, even though these have somewhat different contours.

Why, then, does attention affect metacontrast masking? And why does the requirement for similarity of the contours of the target and mask appear to become more flexible under conditions of distributed spatial attention? We believe it is for the same reason that four dots can mask almost any stimulus in OSM when focussed attention on the target is precluded: Under these conditions the representation of the target is of sufficiently low spatiotemporal resolution that it can be confused for a prior instantiation of the trailing mask, rather than an object in its own right. When metacontrast masking occurs centrally, the greater spatiotemporal resolution of central vision demands that the target and mask have similar contours and appear in very close spatial proximity in order for them to become perceptually confusable. Once we move into the periphery or conditions under which attention is not focussed on the target, however, this facilitates the perceptual decision that the target and mask reflect two instantiations of a single object representation, and thus the requirement for precise physical similarity between the target and mask is relaxed.

Backward masking

Similarities between OSM and metacontrast masking have often been noted; given both phenomena appear to reflect a low-resolution target representation being integrated with that of the mask. It also appears, however, that backward pattern masking, where the mask spatially overlaps the target, may at least partially tap similar mechanisms (Di Lollo et al., 2000; first proposed this link). For example, it has been found that backward masking is differentially exacerbated when the perceptual quality of the target is already degraded, such as with object occlusion (Wyatte, Curran, & O'Reilly, 2012). Further evidence for object-updating in backward masking is that when the overlapping mask consists of systematically alternating black and white squares, (and thus its form is more easily discernible when the mask is overlaid), it is a less effective mask than one that has exactly the same number of black and white squares but has them arranged in a random, unsystematic way (Coltheart & Arthur, 1972). This is evidence that the target and mask are integrated into a single percept, and the mask that provides better camouflage when they are merged is more effective at obscuring target visibility. If the purpose of the mask was merely to terminate further processing of the target, then the nature of the mask itself should be irrelevant. This suggests that integration of the target and mask representations plays a role in backward masking.

Change blindness

Attentional limitations that impoverish the quality of perceptual representations and render them vulnerable to integration is apparent in many other visual-cognitive phenomena. For example, change blindness refers to the finding that observers often fail to see large changes in a scene (Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1998). Specifically, when two versions of a scene are presented in alternation - an exact version and a version with a significant change to an object in the scene - with an irrelevant image (e.g., grey screen), observers have difficulty in detecting the change/difference (Rensink et al., 1997). However, the irrelevant image interleaved between the two versions of the scene is not

necessary to produce change blindness as the effect can be observed with the scene being continuously present, if visual noise stimuli ('mudsplashes') are added to the scene concurrently with the change, even though they do not obscure the target (O'Regan, Rensink, & Clark, 1999). Change detection may be related to OSM because when attention and processing resources are occupied elsewhere, the critical event (the change to the scene in change blindness, or the target in OSM) is encoded with poor spatiotemporal resolution, making it missed entirely, or susceptible to integration. In this sense, we support the notion that attention represents a necessary, but not a sufficient condition, for visual awareness (Cohen, Cavanagh, Chun, & Nakayama, 2012), rather than the alternative that spatial attention and perceptual consciousness reflect distinct operations (Koch & Tsuchiya, 2007, 2012).

Object correspondence through occlusion

If OSM does indeed reflect the operation of mechanisms involved in integrating (versus segregating) object identities, then there should be converging evidence between OSM and other traditional studies of object correspondence. On closer examination of the literature, such convergence is apparent. Many laboratory object correspondence tasks involve the presentation of two suprathreshold events, before and after occlusion. For example, a green circle travels across the screen (event one), disappears behind an occluder, and then reemerges on the other side (event two). The extent to which observers perceive the emerging object as the same object that disappeared behind the occluder can be gauged with subjective report of perceived object identity persistence (Burke, 1952), indirectly via an objective change detection task (Hollingworth & Franconeri, 2009), or the end location of saccades (Richard, Luck, & Hollingworth, 2008). Alternatively, some studies have used apparent motion, but such tasks usually employ similar dependent measures, such as report on the nature or direction of motion (Green & Odom, 1986; Hein & Moore, 2012).

What properties of the events need to match in order for a persisting object identity (rather than two discrete objects) to be perceived? Both an object's surface features and its spatiotemporal history make independent contributions to object correspondence across occlusion by another object (Hollingworth & Franconeri, 2009), saccades (Richard, et al., 2008), and apparent motion events (Hein & Moore, 2012). It is interesting, therefore, that results in the OSM paradigm converge on the same conclusion: While the mask's spatiotemporal trajectory influences masking (e.g., Lleras & Moore, 2003), so does its features. For example, masking strength is modulated by whether the target and mask are the same colour (Moore & Lleras, 2005) or the same luminance (Luiga & Bachmann, 2008), such that masking is greater when the target and mask had common features compared to when they had conflicting features. This convergence between traditional measures of object correspondence and OSM suggests that they may tap common operations.

Visual Short-Term Memory

Object-updating implies that on at least a proportion of trials, the first event (i.e., the target) and the second event are stored in order for a comparison to take place. What memory store mediates this process? Visual short-term memory (VSTM) is a store in which about three to four bound objects are stored and explicitly reportable (Luck & Vogel, 1997). However, masking is affected by mechanisms other than loading information into VSTM (Guest, et al., 2012), and ERP evidence suggests that targets which fail to be reported in OSM are not encoded into VSTM (Prime, Pluchino, Eimer, Dell'Acqua, & Jolicoeur, 2011). Thus, VSTM does not appear to be the most likely candidate.

An alternate store, visible persistence, is a briefly-lingering low-level stimulus representation that is extinguished within approximately 100-200ms of stimulus offset (Coltheart, 1980; Di Lollo, 1980; Hogben & Di Lollo, 1974). However, manipulations that

occur beyond this time window still influence OSM (Goodhew, et al., 2012; Goodhew, et al., 2011a; Lleras & Moore, 2003). Informational persistence (also known as iconic memory) is preserved for longer durations, and lingering informational persistence is typically gauged with cue after target offset, which is the first point at which the to-be-reported target is differentiated from distractors (Sperling, 1960). In OSM, the location of the target is cued from the time of exposure. Informational persistence, furthermore, is purported to be an abstract representation, and yet we know that improvement with prolonged mask exposure is extinguished when a backward pattern mask overlaps the spatial location of the target, suggesting that even at long durations it is a visual, rather than informational (Goodhew, et al., 2011a). Consistent with this, since OSM appears to fundamentally reflect object integration mechanisms, it would make most sense for the target representation to be a mid-level representation, like that of the ‘object token’, a marker for an identity whose location can be updated over time, and whose features are irrelevant (Kahneman, Treisman, & Gibbs, 1992). Unlike the representation suggested by Kahneman et al., however, the representation here appears to retain featural information (Luiga & Bachmann, 2008; Moore & Lleras, 2005).

It is worth noting that Brockmole and colleagues (Brockmole, Irwin, & Wang, 2003; Brockmole & Wang, 2003; Brockmole, Wang, & Irwin, 2002) found that performance on a temporal integration task follows a curvilinear function much like OSM. These studies used the missing-dot paradigm, in which observers were presented with two successive matrices of dots (e.g., 4 x 4 grid) separated by variable intervals. If the matrices were overlaid, all of the locations were filled except one. Observers’ task was to identify this location (i.e., find the missing dot). It was found that missing-dot localisation accuracy was maximal at an inter-stimulus interval (ISI) of 0 ms between the to-be-integrated images, declined across the first 100ms or so, and then steadily increased over the course of at least another 1000ms.

Brockmole et al. suggested that this was evidence that the visual system is capable of integrating a currently-maintained visual image with perceptual information, and that the prolonged timescale over which this occurs is due to sluggish mechanisms for generating a consolidated first image.

More recent work, however, has shown that consolidation of an image in such missing-dot tasks takes about 200ms (Jiang, 2004), and that after 500ms the visual system maintains separate representations of each array in VSTM (Jiang & Kumar, 2004). This is somewhat dissimilar to Brockmole's timecourse, but still not inconsistent with OSM. Recovery from OSM peaks by 640ms after target exposure (Goodhew et al., 2012), so by this point there may be separate representations for the target and mask available, which may even be stored in VSTM. But of greater interest is the nature of the target representation before this, in the form that allows for object integration. Recently it has been suggested that the most efficient way to perform the missing-dot task is to memorise the location in which there are *not* dots and compare this with the second array, meaning that the task is not tapping true perceptual integration (Hollingworth, Hyun, & Zhang, 2005; Jiang, Kumar, & Vickery, 2005), casting doubt on the ability of the results from this paradigm to inform our understanding of OSM. Thus, we have not yet identified the precise nature of the memory representation in OSM, but it has several important characteristics – visual in form, contains both featural and spatiotemporal information, relatively long in duration, and, as we shall see in the next section, capable of activating semantic-level recognition of the target.

The fate of masked targets in OSM

Focussing on the mechanism that suppresses the target from, or catapults it into, visual awareness is vital for understanding OSM and, indeed, vision in general. Equally important, however, is assessing *unconscious* vision – that is, the level of processing of the

masked target. This final section will review the evidence regarding the fate of masked targets in OSM.

How do salient stimuli fare as targets in OSM? One example of such a stimulus category is the human face. Processing of faces is associated with increased negativity over occipito-temporal electrode sites about 130-200 ms after stimulus presentation (Bentin, Allison, Puce, Perez, & McCarthy, 1996). This component of the event-related potential (ERP) is known as the N170 and it is thought to reliably predict the conscious perception of a stimulus as a face. The delayed offset of a four-dot mask has been found to obliterate the N170 (Reiss & Hoffman, 2007), suggesting that OSM impairs early levels of visual analysis. Examination of Reiss and Hoffman's (2007) data presented in their paper, however, reveals that while there was no difference in the amplitude of the waveforms triggered by masked face versus house stimuli in the couple of hundred milliseconds after the target, a systematic and sustained difference did appear to emerge several hundred milliseconds later - beyond the time window of analysis for the N170. This suggests that there could be an electrophysiological correlate of faces that are processed (to the level that they are differentiated from other objects, such as houses) and yet not consciously perceived due to masking. However, a limitation of this study is that the authors did not separately compare on a trial-by-trial basis when participants were aware of the target and when they were not. Since masking was not complete (target identification accuracy 70.4% where chance is 50%), this means that on a considerable portion of trials in the 'masking' condition, participants would have actually been aware of the target. This means that the ERP difference could be driven by target awareness on those trials. As it stands, this work is not definitive on the fate of unconscious stimuli in OSM.

While it remains unclear whether faces are perceived during OSM, features of objects clearly are processed. Chen and Treisman (2009) conducted a series of experiments

investigating the level of implicit processing in OSM. These authors found that when the target and mask stimuli were featurally identical (e.g., left-pointing double-headed arrows, [$\leftarrow\rightarrow$]), response times to identify the mask were facilitated compared with when they were different (e.g., target left-pointing and mask right-pointing arrows), even for targets that were missed on a target detection task. This demonstrates that there is processing of the basic physical features of the target, even in the absence of explicit perceptual awareness (Chen & Treisman, 2009). They attempted to assess the level of implicit semantic perception in subsequent experiments using isolated letter stimuli and defining compatibility in a somewhat arbitrary manner (vowels versus consonants) that did not reflect the way we typically categorise or process letters. Typically, we see them as parts of words, and rapidly extract the meaning of those words. Thus, what Chen and Treisman essentially showed is that there is not implicit processing of the vowel/consonant distinction in OSM. But when word stimuli are used, there is evidence for implicit semantic perception of suppressed targets.

In order to capture the richness of abstract meaning, Goodhew et al. (2011b; Experiment 2) used target words (PINK, BLUE, HOUR, MAIL, coloured grey), and masked them with four coloured dots (all either coloured pink or blue on a given trial), with (####) as distractors. Targets were present on 50% of trials, and the mask had either a simultaneous or delayed (200 ms) offset. Participants' task was to 1) make a speeded mask colour judgement (pink versus blue), and 2) make an unspeeded target detection task. Compatibility was defined as the relationship between the semantic meaning of the target word and the colour of the mask. That is, when the target word was present, on some trials the semantic meaning of the target matched the colour of the four dots (e.g., 'PINK' inside pink dots; compatible), on others, the semantic meaning of the target mismatched the colour of the four dots (e.g., 'PINK' inside blue dots; incompatible), and on other still there were non-colour words were neutral with respect to the colour of the four dot mask. Crucially, the only variable that

differed across the compatible, neutral and incompatible trials was the semantic relationship between the target and the coloured dots – featurally, the conditions were equivalent (four black letters inside coloured dots).

Target detection was impaired for delayed relative to simultaneous offset trials. On the delayed mask offset trials, when the target was correctly detected, response times were significantly faster on the compatible compared to the incompatible trials, with neutral words yielding intermediate response times. When the target was missed, the opposite pattern was observed – response times were significantly faster for incompatible compared to compatible trials. This ‘negative compatibility effect’ (NCE) (Bennett, Lleras, Oriet, & Enns, 2007; Eimer & Schlaghecken, 1998) is striking, and demonstrates different processing in conscious and unconscious vision (see Goodhew et al., 2011b, for discussion of this effect). But the most important finding was that a systematic relationship between the meaning of the target and response times to the mask, indicating that the target word was processed to the level of semantics, despite failing to be detected. This demonstrates that there is implicit semantic perception in OSM (Goodhew, et al., 2011b).

The conclusion that there is implicit semantic perception may seem at odds with the findings that OSM abolishes the N400 to word targets (Reiss & Hoffman, 2006), an event-related potential used as an electrophysiological signature of semantic processing (Kutas & Hillyard, 1980). However, if we examine Reiss and Hoffman’s (2006) methodology, we find that these results are also not definitive regarding the level of processing in OSM. These authors had participants read a context word followed by a semantically related or unrelated target word surrounded by dots. The participants then identified the target from multiple display options. The delayed mask offset impaired target identification performance and while the differential amplitude of the N400 to semantically-related versus unrelated words was present for the consciously perceived targets, it was abolished in the delayed offset

condition. While these results point toward OSM impairing semantic processing, there are a couple of limitations of this study. Firstly, the authors did not compare perceptually aware versus unaware trials, instead grouping all delayed mask offset trials together for the analysis. This makes it possible that any effect for unaware trials was obscured, especially if it differed qualitatively from that for aware trials. Secondly, there was, in fact, some behavioural evidence for implicit semantic perception – responses were significantly more accurate for compatible trials (context and target word semantically related) compared to incompatible trials, but this is difficult to interpret, since as the authors acknowledged, it could be attributed to a guessing bias. Thus, this study, as it stands, does not conclusively reveal the fate of the masked target in OSM.

To summarise, the evidence indicates that (a) successfully masked targets in OSM are nonetheless implicitly processed at both the basic physical feature and higher-level abstract semantic level, and (b) there is a qualitative difference in the effect of the perceived versus missed targets at the semantic level. This is not because the missed targets failed to be processed - this would result in an absence of priming on the miss trials. Instead, there was strong priming, but it differed qualitatively in nature from the pattern when the target was visible.

Conclusion

OSM has been an exciting development in visual masking techniques that are used for understanding the mechanisms of conscious visual perception. It appears to tap properties of the visual system that are common to multiple paradigms, including metacontrast masking, backward masking, change blindness, change detection, and object correspondence through occlusion, to name but a few. That is, when focal attention to the target is prevented, this stimulus is encoded with sufficiently low spatiotemporal resolution that it is perceptually

confused for a prior instantiation of the mask. Ultimately, we conclude that OSM reflects a perceptual decision to a fundamental problem in visual cognition: whether stimulation in close spatiotemporal proximity reflects a persisting object identity or discrete object representations. Masking occurs when the visual system decides in favour of the latter, thereby suppressing target visibility. Finally, OSM selectively impairs awareness while leaving high-level processing of the target intact, making it an ideal tool for the study of unconscious perception.

Acknowledgements

This work was supported by an Ontario Government Fellowship awarded to S.C.G, NSERC grants awarded to S.F. and J.P., an Australian Research Council Discovery Grant and Fellowship awarded to PED (DP0986387) and an Early Researcher Award and CIHR grant awarded to S.F.

References

- Abrams, R. A., Davoli, C., Du, F., Knapp, W. H., & Paull, D. (2008). Altered vision near the hands. *Cognition*, *107*(3), 1035-1047. doi: 10.1016/j.cognition.2007.09.006
- Alpern, M. (1952). Metacontrast: Historical introduction. *American Journal of Optometry*, *29*, 631-646.
- Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America*, *43*, 648-657. doi: 10.1364/JOSA.43.000648
- Anstis, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London B*, *290*(1038), 153-167. doi: 10.1098/rstb.1980.0088
- Argyropoulos, I., Gellatly, A., Pilling, M., & Carter, W. (in press). Set size and mask duration do not interact in object-substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*. doi: 10.1037/a0030240
- Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, *40*, 309-328.
- Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in object recognition. *Journal of Cognitive Neuroscience*, *15*(4), 600-609. doi: 10.1162/089892903321662976
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., . . . Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, *103*(2), 449-454. doi: 10.1073/pnas.0507062103
- Baxt, N. (1871). Über die Zeit, welche nötig ist, damit ein Gesichtseindruck zum Bewusstsein kommt und über die Grösse (Extension) der bewussten Wahrnehmung bei einem Gesichtseindrücke von gegebener Dauer *Pfluegers Arch Gesamte Physiol Menschen Tiere*, *4*, 325-336.

- Baxt, N. (1982). On the time necessary for a visual impression to come into consciousness. *Psychological Research*, 44(1), 1-12. doi: 10.1007/BF00308551
- Bennett, J. D., Lleras, A., Oriet, C., & Enns, J. T. (2007). A negative compatibility effect in priming of emotional faces. *Psychonomic Bulletin & Review*, 14(5), 908-912. doi: 10.3758/BF03194120
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8(6), 551-565. doi: 10.1162/jocn.1996.8.6.551
- Binsted, G., Brownwell, K., Vorontsova, Z., Heath, M., & Saucier, D. (2007). Visuomotor system uses target features unavailable to conscious awareness. *Proceedings of the National Academy of Sciences*, 104(31), 12669-12672. doi: 10.1073/pnas.0702307104
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 15(2), 205-206. doi: 10.1017/S0140525X00038188
- Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, 15(12), 567-575. doi: 10.1016/j.tics.2011.11.001
- Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience*, 28, 157-189. doi: 10.1146/annurev.neuro.26.041002.131052
- Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, 83(1), 1-36. doi: 10.1037/0033-295X.83.1.1
- Breitmeyer, B. G., & Kersey, M. (1981). Backward masking by pattern stimulus onset. *Journal of Experimental Psychology: Human Perception and Performance*, 7(5), 972-977. doi: 10.1037/0096-1523.7.5.972

- Breitmeyer, B. G., & Ögmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, 62(8), 1572-1595. doi: 10.3758/BF03212157
- Breitmeyer, B. G., & Ögmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (2nd ed.): Oxford University Press.
- Bridgeman, B. (1971). Metacontrast and lateral inhibition. *Psychological Review*, 78, 528-539. doi: 10.1037/h0031782
- Bridgeman, B. (1977). A correlational model applied to metacontrast: Reply to Weisstein, Ozog, and Szoc. *Bulletin of the Psychonomic Society*, 10, 85-88.
- Bridgeman, B. (2007). Common-onset masking simulated with a distributed-code model. *Advances in Cognitive Psychology*, 3(1-2), 33-40. doi: 10.2478/v10053-008-0012-5
- Brockmole, J. R., Irwin, D. E., & Wang, R. F. (2003). The locus of spatial attention during the temporal integration of visual memories and visual percepts. *Psychonomic Bulletin & Review*, 10(4), 510-515. doi: 10.3758/BF03196562
- Brockmole, J. R., & Wang, R. F. (2003). Integrating visual images and visual percepts across time and space. *Visual Cognition*, 10(7), 853-873. doi: 10.1080/13506280344000121
- Brockmole, J. R., Wang, R. F., & Irwin, D. E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 315-334. doi: 10.1037//0096-1523.28.2.315
- Bullier, J. (2001). Feedback connections and conscious vision. *Trends in Cognitive Sciences*, 5(9), 369-370. doi: 10.1016/S1364-6613%2800%2901730-7
- Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, 4, 121-138. doi: 10.1080/17470215208416611
- Chakravarthi, R., & Cavanagh, P. (2009). Recovery of a crowded object by masking the flankers: Determining the locus of feature integration. *Journal of Vision*, 9(10), 1-9.

- Chen, Z., & Treisman, A. (2009). Implicit perception and level of processing in object-substitution masking. *Psychological Science*, 20(5), 560-567. doi: 10.1111/j.1467-9280.2009.02328.x
- Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences*, 16(8), 411-417. doi: 10.1016/j.tics.2012.06.013
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, 27(3), 183-228. doi: 10.3758/BF03204258
- Coltheart, M., & Arthur, B. (1972). Evidence for an integration theory of visual masking. *Quarterly Journal of Experimental Psychology*, 24(3), 262-269. doi: 10.1080/14640747208400279
- Cudeiro, J., & Sillito, A. M. (2006). Looking back: Corticothalamic feedback and early visual processing. *Trends in Neurosciences*, 29(6), 298-306. doi: 10.1016/j.tins.2006.05.002
- Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in the lateral geniculate nucleus of the macaque. *Journal of Physiology*, 357, 219-240.
- Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, 109(1), 75-97. doi: 10.1037/0096-3445.109.1.75
- Di Lollo, V. (2010). Iterative reentrant processing: A conceptual framework for perception and cognition (the blinding problem? No worries, mate). In V. Coltheart (Ed.), *Tutorials in Visual Cognition*. New York: Psychology Press.
- Di Lollo, V. (2012). The feature-binding problem is an ill-posed problem. *Trends in Cognitive Sciences*, 16(6), 317-321. doi: 10.1016/j.tics.2012.04.007

- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, 129(4), 481-507. doi: 10.1037/0096-3445.129.4.481
- Di Lollo, V., von Muhlenen, A., Enns, J. T., & Bridgeman, B. (2004). Decoupling stimulus duration from brightness in metacontrast masking: Data and models. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 733-745. doi: 10.1037/0096-1523.30.4.733
- Dux, P. E., Visser, T. A. W., Goodhew, S. C., & Lipp, O. V. (2010). Delayed re-entrant processing impairs visual awareness: An object substitution masking study. *Psychological Science*, 21(9), 1242-1247. doi: 10.1177/0956797610379866
- Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, 24(6), 1737-1747. doi: 10.1037/0096-1523.24.6.1737
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, 8(2), 135-139. doi: 10.1111/j.1467-9280.1997.tb00696.x
- Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, 4(9), 345-352. doi: 10.1016/S1364-6613%2800%2901520-5
- Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience*, 19(9), 1488-1497. doi: 10.1162/jocn.2007.19.9.1488
- Francis, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review*, 104(3), 572-594. doi: 10.1037/0033-295X.104.3.572

- Francis, G., & Hermens, F. (2002). Comment on "Competition for consciousness among visual events: The psychophysics of reentrant visual processes" (Di Lollo, Enns, & Rensink, 2000). *Journal of Experimental Psychology: General*, 131(4), 590-593. doi: 10.1037/0096-3445.131.4.590
- Gellatly, A., Pilling, M., Carter, W., & Guest, D. (2010). How does target duration affect object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1267-1279. doi: 10.1037/a0018733
- Goodale, M. A. (2008). Action without perception in human vision. *Cognitive Neuropsychology*, 25(7-8), 891-919. doi: 10.1080/02643290801961984
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25. doi: 10.1016/0166-2236(92)90034-8
- Goodale, M. A., & Westwood, D. A. (2004). An evolving view of duplex vision: Separate but interacting cortical pathways for perception and action. *Current Opinion in Neurobiology*, 14(2), 203-211. doi: 10.1016/j.conb.2004.03.002
- Goodhew, S. C., Dux, P. E., Lipp, O. V., & Visser, T. A. W. (2012). Understanding recovery from object substitution masking. *Cognition*, 122(3), 405-415. doi: 10.1016/j.cognition.2011.11.010
- Goodhew, S. C., Gozli, D. G., Ferber, S., & Pratt, J. (in press). Reduced temporal fusion in near-hand space. *Psychological Science*.
- Goodhew, S. C., Visser, T. A. W., Lipp, O. V., & Dux, P. E. (2011a). Competing for consciousness: Prolonged mask exposure reduces object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2), 588-596. doi: 10.1037/a0018740

- Goodhew, S. C., Visser, T. A. W., Lipp, O. V., & Dux, P. E. (2011b). Implicit semantic perception in object substitution masking. *Cognition*, *118*(1), 133-137. doi: 10.1016/j.cognition.2010.10.013
- Gozli, D. G., West, G. L., & Pratt, J. (2012). Hand position alters vision by biasing processing through different visual pathways. *Cognition*, *124*(2), 244-250. doi: 10.1016/j.cognition.2012.04.008
- Green, M., & Odom, J. V. (1986). Correspondence matching in apparent motion: Evidence for three-dimensional spatial representation. *Science*, *233*(4771), 1427-1429. doi: 10.1126/science.3749887
- Guest, D., Gellatly, A., & Pilling, M. (2011). The effect of spatial competition between object-level representations of target and mask on object substitution masking. *Attention, Perception, and Psychophysics*, *73*(8), 2528-2541. doi: 10.3758/s13414-011-0196-5
- Guest, D., Gellatly, A., & Pilling, M. (2012). Reduced OSM for long duration targets: Individuation or items loaded into VSTM? *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1541-1553. doi: 10.1037/a0027031
- Heath, M., Neely, K. A., Yakimishyn, J., & Binsted, G. (2008). Visuomotor memory is independent of conscious awareness of target features. *Experimental Brain Research*, *188*(4), 517-527. doi: 10.1007/s00221-008-1385-x
- Hein, E., & Moore, C. M. (2012). Spatio-temporal priority revisited: The role of feature identity and similarity for object correspondence in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 975-988. doi: 10.1037/a0028197

- Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 269-293). Oxford: Elsevier Science Ltd.
- Hirose, N., Kihara, K., Mima, T., Ueki, Y., Fukuyama, H., & Osaka, N. (2007). Recovery from object substitution masking induced by transient suppression of visual motion processing: A repetitive transcranial magnetic stimulation study. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6), 1495-1503. doi: 10.1037/0096-1523.33.6.1495
- Hirose, N., & Osaka, N. (2009). Object substitution masking induced by illusory masks: Evidence for higher object-level locus of interference. *Journal of Experimental Psychology: Human Perception and Performance*, 35(4), 931-938. doi: 10.1037/a0012734
- Hogben, J. H., & Di Lollo, V. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vision Research*, 14(11), 1059-1069. doi: 10.1016/0042-6989(74)90202-8
- Hollingworth, A., & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues. *Cognition*, 113(2), 150-166. doi: 10.1016/j.cognition.2009.08.004
- Hollingworth, A., Hyun, J. S., & Zhang, W. (2005). The role of visual short-term memory in empty cell localization. *Perception, and Psychophysics*, 67(8), 1332-1343. doi: 10.3758/BF03193638
- Hubel, D. H. (1963). The visual cortex of the brain. *Scientific American*, 209(5), 54-63. doi: 10.1038/scientificamerican1163-54
- James, T. W., Culham, J., Humphreys, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An

- fMRI study. *Brain: A Journal of Neurology*, 126(11), 2463-2475. doi: 10.1093/brain/awg248
- Jiang, Y., & Kumar, A. (2004). Visual short-term memory for two sequential arrays: One integrated representation or two separate representations? *Psychonomic Bulletin & Review*, 11(3), 495-500. doi: 10.3758/BF03196601
- Jiang, Y., Kumar, A., & Vickery, T. J. (2005). Integrating sequential arrays in visual short-term memory. *Experimental Psychology*, 52(1), 39-46. doi: 10.1027/1618-3169.52.1.39
- Kahan, T. A., & Enns, J. T. (2010). Object trimming: When masking dots alter rather than replace target representations. *Journal of Experimental Psychology: Human Perception and Performance*, 36(1), 88-102. doi: 10.1037/a0016466
- Kahan, T. A., & Lichtman, A. S. (2006). Looking at object-substitution masking in depth and motion: Toward a two-object theory of object substitution. *Perception & Psychophysics*, 68(3), 437-446. doi: 10.3758/BF03193688
- Kahneman, D. (1967). An onset-onset law for one case of apparent motion and metacontrast. *Perception & Psychophysics*, 2(12-A), 577-584. doi: 10.3758/BF03210272
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175-219. doi: 10.1016/0010-0285(92)90007-O
- Kim, C. Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible 'invisible'. *Trends in Cognitive Sciences*, 9(8), 381-388. doi: 10.1016/j.tics.2005.06.012
- Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences*, 11(1), 16-22. doi: 10.1016/j.tics.2006.10.012
- Koch, C., & Tsuchiya, N. (2012). Attention and consciousness: Related yet different. *Trends in Cognitive Sciences*, 16(2), 103-105. doi: 10.1016/j.tics.2011.11.012

- Kotsoni, E., Csibra, G., Mareschal, D., & Johnson, M. H. (2007). Electrophysiological correlates of common-onset visual masking. *Neuropsychologia*, 45(10), 2285-2293. doi: 10.1016/j.neuropsychologia.2007.02.023
- Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, 14(7), 301-307. doi: 10.1016/j.tics.2010.04.006
- Kouider, S., Sackur, J., & de Gardelle, V. (2012). Do we still need phenomenal consciousness? Comment on Block. *Trends in Cognitive Sciences*, 16(3), 140-141. doi: 10.1016/j.tics.2012.01.003
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205. doi: 10.1126/science.7350657
- Lachter, J., & Durgin, F. H. (1999). Metaccontrast masking functions: A question of speed? *Journal of Experimental Psychology: Human Perception & Performance*, 25(4), 936-947. doi: 10.1037/0096-1523.25.4.936
- Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23(11), 571-579. doi: 10.1016/S0166-2236(00)01657-X
- Lamme, V. A. F., Zipser, K., & Spekreijse, H. (2002). Masking interrupts figure-ground signals in V1. *Journal of Cognitive Neuroscience*, 14(7), 1044-1053. doi: 10.1162/089892902320474490
- Lleras, A., & Moore, C. M. (2003). When the target becomes the mask: Using apparent motion to isolate the object-level component of object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 29(1), 106-120. doi: 10.1037/0096-1523.29.1.106

- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279-281. doi: 10.1038/36846
- Luiga, I., & Bachmann, T. (2008). Luminance processing in object substitution masking. *Vision Research*, 48(7), 937-945. doi: 10.1016/j.visres.2008.01.001
- Moore, C. M., & Lleras, A. (2005). On the role of object representations in substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6), 1171-1180. doi: 10.1037/0096-1523.31.6.1171
- Mumford, D. (1991). On the computational architecture of the neocortex: I. The role of the thalamo-cortical loop. *Biological Cybernetics*, 65(2), 135-145. doi: 10.1007/BF00202389
- Neill, W. T., Hutchison, K. A., & Graves, D. F. (2002). Masking by object substitution: Dissociation of masking and cueing effects. *Journal of Experimental Psychology: Human Perception and Performance*, 28(3), 682-694. doi: 10.1037/0096-1523.28.3.682
- O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature*, 398(6722), 34-34. doi: 10.1038/17953
- Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, 292(5516), 510-512. doi: 10.1126/science.1057099
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 11(10), 1129-1135. doi: 10.1038/nn.2187
- Pilling, M., & Gellatly, A. (2010). Object substitution masking and the object updating hypothesis. *Psychonomic Bulletin & Review*, 17(5), 737-742. doi: 10.3758/PBR.17.5.737

- Pöder, E. (in press). Attentional gating models of object substitution. *Journal of Experimental Psychology: General*. doi: 10.1037/a0030575
- Prime, D. J., Pluchino, P., Eimer, M., Dell'Acqua, R., & Jolicoeur, P. (2011). Object-substitution masking modulates spatial attention deployment and the encoding of information in visual short-term memory. *Psychophysiology*, 48(5), 687-696. doi: 10.1111/j.1469-8986.2010.01133.x
- Reed, C., Betz, R., Garza, J. P., & Roberts, R. J. (2010). Grab it! Biased attention in functional hand and tool space. *Attention, Perception, and Psychophysics*, 72(1), 236-245. doi: 10.3758/APP.72.1.236
- Reed, C., Grubb, J., & Steele, C. (2006). Hands up: Attentional prioritization of space near the hand. *Journal of Experimental Psychology: Human Perception and Performance*, 32(1), 166-177. doi: 10.1037/0096-1523.32.1.166
- Reiss, J. E., & Hoffman, J. E. (2006). Object substitution masking interferes with semantic processing: Evidence from event-related potentials. *Psychological Science*, 17(12), 1015-1020. doi: 10.1111/j.1467-9280.2006.01820.x
- Reiss, J. E., & Hoffman, J. E. (2007). Disruption of early face recognition processes by object substitution masking. *Visual Cognition*, 15(7), 789-798. doi: 10.1080/13506280701307035
- Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8(5), 368-373. doi: 10.1111/j.1467-9280.1997.tb00427.x
- Richard, A. M., Luck, S. J., & Hollingworth, A. (2008). Establishing object correspondence across eye movements: Flexible use of spatiotemporal and surface information. *Cognition*, 109(1), 66-88. doi: 10.1016/j.cognition.2008.07.004

- Shapley, R. (1990). Visual sensitivity and parallel retinocortical channels. *Annual Review of Psychology*, 41, 635-658. doi: 10.1146/annurev.ps.41.020190.003223
- Shelley-Tremblay, J., & Mack, A. (1999). Metacontrast masking and attention. *Psychological Science*, 10(6), 508-515. doi: 10.1111/1467-9280.00197
- Sherman, S. M., Wilson, J. R., Kaas, J. H., & Webb, S. V. (1976). X- and Y-Cells in the dorsal lateral geniculate nucleus of the Owl Monkey (*Aotus trivirgatus*). *Science*, 192(4238), 475-477. doi: 10.1126/science.816006
- Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5(4), 644-649. doi: 10.3758/BF03208840
- Singer, W., & Creutzfeldt, O. D. (1970). Reciprocal lateral inhibition of on- and off-center neurones in the lateral geniculate body of the cat. *Experimental Brain Research*, 10(3), 311-330. doi: 10.1007/BF00235054
- Sperling, G. (1960). The information available in brief presentations. *Psychological Monographs: General and Applied*, 74(11), 1-29. doi: 10.1037/h0093759
- Tapia, E., & Breitmeyer, B. G. (2011). Visual consciousness revisited: Magnocellular contributions to conscious and nonconscious vision. *Psychological Science*, 22(7), 934-942. doi: 10.1177/0956797611413471
- Tata, M. S. (2002). Attend to it now or lose it forever: Selective attention, metacontrast masking, and object substitution. *Perception & Psychophysics*, 64(7), 1028-1038. doi: 10.3758/BF03194754
- Tata, M. S., & Giaschi, D. E. (2004). Warning: Attending to a mask may be hazardous to your perception. *Psychonomic Bulletin & Review*, 11(2), 262-268. doi: 10.3758/BF03196568

- Weidner, R., Shah, N. J., & Fink, G. R. (2006). The neural basis of perceptual hypothesis generation and testing. *Journal of Cognitive Neuroscience*, 18(2), 258-266. doi: 10.1162/jocn.2006.18.2.258
- Weisstein, N. (1968). A Rashevsky-Landahl neural net: Simulation of metacontrast. *Psychological Review*, 75(6), 494-521.
- Werner, H. (1935). Studies on counter: I. Qualitative analyses. *American Journal of Psychology*, 47, 40-64. doi: 10.2307/1416707
- Wunderlich, K., Schneider, K. A., & Kastner, S. (2005). Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nature Neuroscience*, 8(11), 1595-1602. doi: 10.1038/nn1554
- Wyatte, D., Curran, T., & O'Reilly, R. (2012). The limits of feedforward vision: Recurrent processing promotes robust object recognition when objects are degraded. *Journal of Cognitive Neuroscience*, 24(11), 2248-2261. doi: 10.1162/jocn_a_00282
- Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, 396(6706), 72-75. doi: 10.1038/23936
- Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, 39(2), 293-306. doi: 10.1016/S0042-6989(98)00114-X
- Yeshurun, Y., & Carrasco, M. (2008). The effects of transient attention on spatial resolution and the size of the attentional cue. *Perception & Psychophysics*, 70(1), 104-113. doi: 10.3758/PP.70.1.104
- Zeki, S. (1993). *A vision of the brain*. Oxford, England: Blackwell.
- Zeki, S. (2001). Localization and globalization in conscious vision. *Annual Review of Neuroscience*, 24, 57-86. doi: 10.1146/annurev.neuro.24.1.57