



ELSEVIER

Contents lists available at SciVerse ScienceDirect

## Cognition

journal homepage: [www.elsevier.com/locate/COGNIT](http://www.elsevier.com/locate/COGNIT)

# Understanding recovery from object substitution masking<sup>☆</sup>

Stephanie C. Goodhew<sup>\*</sup>, Paul E. Dux, Ottmar V. Lipp, Troy A.W. Visser

School of Psychology, University of Queensland, Australia

## ARTICLE INFO

### Article history:

Received 17 December 2010

Revised 15 November 2011

Accepted 16 November 2011

Available online 10 December 2011

### Keywords:

Object substitution masking

Temporal dynamics

Visual masking

## ABSTRACT

When we look at a scene, we are conscious of only a small fraction of the available visual information at any given point in time. This raises profound questions regarding how information is selected, when awareness occurs, and the nature of the mechanisms underlying these processes. One tool that may be used to probe these issues is object-substitution masking (OSM). In OSM, a sparse, temporally-trailing four dot mask can interfere with target perception, even though the target and mask have different contours and do not spatially overlap (Enns & Di Lollo, 1997). Here, we investigate the mechanisms underlying the recently discovered recovery from OSM observed with prolonged mask exposure (Goodhew, Visser, Lipp, & Dux, 2011). In three experiments, we demonstrate that recovery is unaffected by mask offset, and that prolonged physical exposure of the mask is not necessary for recovery. These findings confirm that recovery is not due to: (a) an offset transient impairing the visibility of other stimuli that are nearby in space and time, or (b) mask adaptation or temporal object-individuation cues resulting from prolonged mask exposure. Instead, our results confirm recovery as a high-level visual-cognitive phenomenon, which is inherently tied to target-processing time. This reveals the prolonged iterative temporal dynamics of conscious object perception.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

Our awareness of the visual world is profoundly limited.<sup>1</sup> When we look at a scene, we are usually able to report on only a small fraction of the available sensory information at any given moment, and the information available for report changes rapidly over time (Broadbent, 1958; Crick & Koch, 1998; Kahneman, 1973; Kastner & Pinsky, 2004). Cognitive scientists have long been intrigued by the limits of awareness, and have intensively investigated the factors

that determine what we become aware of and when awareness arises (Benjamin, 2009; Neisser, 1967). Such questions have broad implications for understanding both normal systems and disorders of visual perception such as spatial neglect (Koch, 2004).

The prolonged time-course of conscious awareness has been demonstrated across a variety of paradigms. For example, in a phenomenon known as the attentional blink (AB), when two targets are presented amongst distractors in a rapid serial visual presentation stream, observers have difficulty identifying the second target when it occurs close in time to the first target. This impairment is typically attributed to the time required to establish a conscious representation of the first item (Chun & Potter, 1995; Raymond, Shapiro, & Arnell, 1992; see also Dux & Marois, 2009; for a review). A similar point is illustrated by perceptual integration tasks. For example, in the missing-dot task, two matrix arrays of dots are presented sequentially which, when combined, have dots occupying all but one of the matrix locations (Di Lollo, 1980). The observer's task

<sup>☆</sup> The authors were supported by an Australian Postgraduate Award (SCG), an Australian Research Council (ARC) Discovery Grant and APD Fellowship (DP0986387, PED), an ARC Discovery Grant and APF Fellowship (DP0770844, OVL), and a UQ Early Career Research Grant (TAWV).

<sup>\*</sup> Corresponding author. Present address: Department of Psychology, University of Toronto, 100 St. George St., Toronto, ON, Canada M5S3G3.

E-mail address: [s.c.goodhew@gmail.com](mailto:s.c.goodhew@gmail.com) (S.C. Goodhew).

<sup>1</sup> We use the terms 'consciousness' and 'awareness' to refer to the reportability of a stimulus on either an identification or detection metric (i.e., access consciousness; Block, 1995).

is to identify the missing dot – a judgment that requires the two arrays to be perceptually integrated. Increasing accuracy on this task occurs well beyond the physical offset of the first array (i.e., with increasing ISI between the two arrays), with results demonstrating an improvement in generating a useable image of the first array continues for over a second and a half after the stimulus is no longer visible (Brockmole, Wang, & Irwin, 2002), and, in turn, that stimulus representations can take several seconds to be consolidated into awareness. During this time, there can be an ongoing ‘competition for consciousness’ amongst stimuli in close spatiotemporal proximity, such that even stimuli that initially fail to be consciously perceived are not irrevocably lost, but instead continue to vie for perceptual processing resources.

Perhaps the most widely used tool for understanding the dynamics of this competition for consciousness is visual masking. Here, the visibility of a target stimulus is altered or impaired by the presentation of another spatiotemporally adjacent stimulus – the mask (Breitmeyer & Ögmen, 2006). In backward masking, the presentation of a spatially-overlapping mask (consisting of a pattern, noise, or light), impairs the visibility of a preceding target presented at the same location. Backward masking is interesting because it appears to work retroactively in time: when presented alone, the target is of sufficient duration and quality to be perceived, but the following mask alters its perceptibility. Even more striking is the phenomenon of metacontrast masking. Here, a non-spatially overlapping mask (e.g., an annulus) impairs the visibility of preceding target (e.g., a disc). Surprisingly, under certain parameters, this impairment is non-monotonic, such that it is virtually absent when the stimulus-onset asynchrony (SOA) between target and mask is short or long, and maximal at intermediate SOAs (Alpern, 1952, 1953; Breitmeyer & Ögmen, 2006). These phenomena demonstrate how the competition for consciousness among stimuli in close spatiotemporal proximity can operate in seemingly counter-intuitive ways.

More recently, a new form of visual masking has garnered considerable research interest: object substitution masking (OSM). In OSM, a target stimulus (e.g., a ring with a gap on the right or left) and a sparse surrounding mask (e.g., four dots adjacent to the corners of the target) appear simultaneously amongst an array of distractor rings spread across space. After a brief interval (e.g., 30 ms), all items offset except the mask, which remains on the display for a variable duration. Under these conditions, identification accuracy declines with increases in both mask duration and number of distractors (Di Lollo, Enns, & Rensink, 2000; Enns & Di Lollo, 1997).

Object substitution is unique in several respects. First, unlike backward masking, the target and mask do not spatially overlap. Second, unlike metacontrast masking, the impairment in target perception occurs in the absence of substantial overlap of the target and mask contours. Third, OSM is only observed when the target is unattended, since the distribution of attention during target presentation is critical to OSM (Di Lollo et al., 2000; Dux, Visser, Goodhew, & Lipp, 2010), whereas attention merely modulates meta-

contrast masking (e.g., Ramachandran & Cobb, 1995; Shellen-Tremblay & Mack, 1999).

To account for the unique properties of OSM, Di Lollo et al. (2000) proposed a framework based on re-entrant neural processing, in which visual analysis begins with a feedforward sweep of information from posterior visual areas (e.g., primary visual cortex, V1) to anterior brain regions (e.g., prefrontal cortex; Hubel & Wiesel, 1959). From this information, anterior regions generate a “perceptual hypothesis” about the identity of the input. Critically, due to the large receptive field size and coarse coding of neurons in anterior regions, this representation is low resolution, and ambiguous. Thus, re-entrant signals must be sent back to posterior brain regions in order to verify the hypothesis against the high-resolution sensory information available in these areas. When this verification process is delayed (by increasing the number of distractors and thus delaying the deployment of attention), and the mask remains on display after target offset, there is a mismatch between the perceptual hypothesis in anterior regions (containing the target and mask), and the incoming sensory input in posterior regions (reflecting the mask alone). Masking occurs when the visual system updates the hypothesis to reflect the incoming sensory evidence about just the mask alone, at the expense of the target and mask representation (Di Lollo et al., 2000; Enns & Di Lollo, 1997).

The re-entrant processing account of OSM (Di Lollo et al., 2000) offers a coherent framework for describing the processes underlying conscious object formation – an important theoretical development. However, although the re-entrant processing account can explain OSM, the phenomenon is by no means itself a confirmation of theory, since it has been argued that existing feedforward-only models can also explain this type of masking (e.g., Francis & Hermens, 2002). It has also been argued that OSM can uniquely probe the formation of object identities (Lleras & Moore, 2003; Moore & Lleras, 2005) – that is, the key properties of a stimulus that persist through time and space (Flombaum & Scholl, 2006; Kahneman, Treisman, & Gibbs, 1992; Scholl, 2002). Thus, it is clear that rigorous testing of the re-entrant and alternative accounts of OSM is an important research endeavour that can shed light onto key processes underlying object recognition.

Some additional research evidence that may implicate feedback mechanisms in OSM comes from the recent discovery of the prolonged time-course of target-mask interactions in the effect. While early reports suggested that masking strength increased monotonically with trailing mask duration (Di Lollo et al., 2000), we found that target identification *improved* with prolonged mask exposure (e.g., 640 ms, Goodhew, Visser, Lipp, & Dux, 2011), yielding a U-shaped masking function across mask durations. Moreover, this non-monotonicity was specific to OSM, as prolonged exposure of a spatially-overlapping backward noise mask produced a monotonic decline in target accuracy.

To explain this finding, we proposed that the recovery from OSM at prolonged mask durations arises from a second cascade of re-entrant processing focused on the target representation that begins after prolonged mask exposure.

That is, we suggest that the lingering target representation can induce a second cascade of reiterative processing once analysis of the mask is complete (i.e., after some minimal duration of mask exposure). This second series of reiterative loops, in turn, allows the target to be identified more accurately relative to intermediate mask durations.

What is the nature of the target representation underlying recovery? The fact that recovery is eliminated with a spatially-overlapping backward noise mask (Goodhew et al., 2011) suggests that the representation must be sufficiently fragile that it is susceptible to low-level interference. Such interference could take the form of masking by integration or interruption (Kaheman, 1968). Integration masking occurs when the target and mask icons are combined, thus substantially impairing target visibility. Interruption masking, in contrast, occurs when the mask interrupts higher-level consolidation of the target representation. The time course of recovery, which peaks over 600 ms after target offset, rules out integration masking, since it is known that such early representations do not survive long enough to underlie recovery (e.g., Di Lollo, 1980; see further discussion in Section 5). Thus, the overlapping noise mask must instead *interrupt* the mechanisms of target consolidation that underlie recovery.

Although our account fits nicely into the re-entrant framework, several alternative explanations for recovery, which do not involve re-entrant processes, are also plausible. One alternative is that recovery arises from increased temporal separation between target and mask offsets. According to the dual-channel sustained-transient model of masking (e.g., Breitmeyer & Ganz, 1976; Singer & Bedworth, 1973), visual processing occurs along two distinct but interacting channels. One is a slow latency, prolonged-response “sustained” channel that encodes information about stimulus form, contrast, and colour; the other is a fast latency, brief-response “transient” channel that encodes motion as well as stimulus onsets and offsets. Masking occurs when transient channel activity arising from the mask suppresses target-related activity in the sustained channel. Critically, this account predicts a U-shaped masking curve, like that obtained by Goodhew et al. (2011), as a function of the duration between target and mask offset (Breitmeyer & Kersey, 1981). This is because at short target-mask stimulus-termination asynchronies (STAs), transient channel activity triggered by the offset of the mask proceeds more rapidly than target-related activity in the sustained channel, rendering the mask ineffective, while at long STAs, masking is avoided because target processing is completed prior to the appearance of the mask. However, at intermediate durations, the fast transient signal triggered by the mask offset overlaps with target processing in the sustained-channel, yielding strong masking (Breitmeyer & Kersey, 1981; Breitmeyer & Ögmen, 2006; see also Macknik & Livingstone, 1998).

A second possible explanation is that recovery from OSM could reflect some type of adaptation to, or temporal cueing from the prolonged mask, which then allows the target to be perceived. For example, Lleras and Moore (2003; see also Moore & Lleras, 2005) demonstrated object-level interference between the target and mask in OSM, and presented evidence that masking occurs when

the target and mask were treated as part of a single object. In contrast, when temporal or featural cues were provided that helped to individuate the target and mask as separate objects, masking was attenuated. From this perspective, prolonged mask exposure may cue the visual system to differentiate between the target and mask, thus releasing the target from masking. Alternatively, it could be that the visual system adapts to the mask after prolonged exposure, and that this is a necessary trigger for renewed target processing, allowing the target to be perceived.

Here, we test these competing accounts of the recovery from OSM. In Experiment 1, we begin by replicating the recovery effect obtained by Goodhew et al. (2011) in order to establish the temporal parameters that lead to masking and recovery in the present paradigm. In Experiment 2, we investigate the role of mask offsets in recovery by using a tone to signal participants' responses, and allowing the mask to remain on display indefinitely. Finally, in Experiment 3, we examine the role of prolonged mask exposure in recovery, by holding mask duration constant and manipulating the duration prior to a response cue, in order to contrast mask adaptation and time-based accounts of the recovery phenomenon.

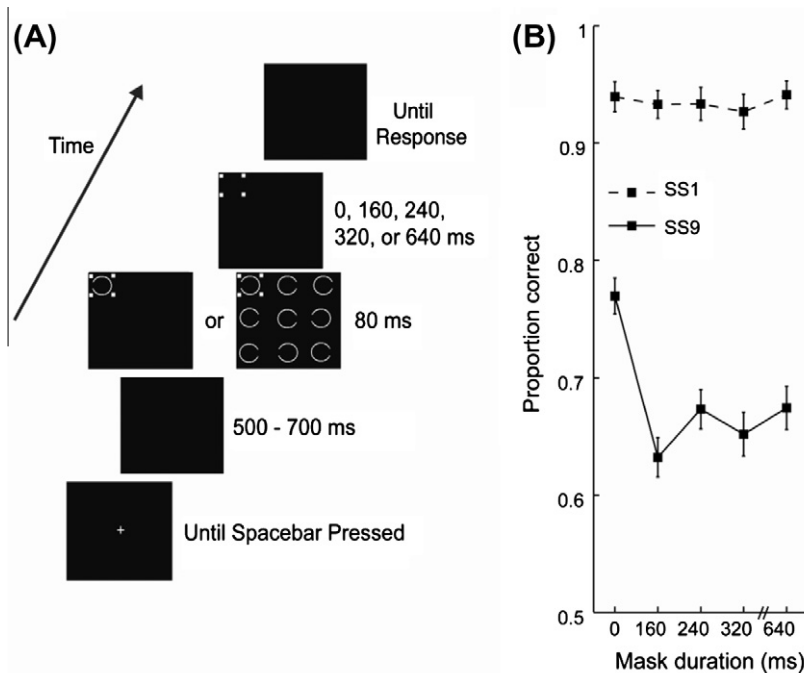
## 2. Experiment 1

The purpose of Experiment 1 was to replicate the OSM recovery effect obtained by Goodhew et al. (2011) using a slightly modified paradigm. This replication can then be used as a baseline against which to compare the results of subsequent experiments.

### 2.1. Method

Stimuli were presented on a Pentium computer, connected to a 19" CRT monitor (Acer AC716) running at a refresh rate of 100 Hz, using the Presentation software package (Neurobehavioral systems, 2009, Albany, CA). Targets were white rings with a gap to the left or right (Landolt-Cs;  $\sim 1.99^\circ$ , luminance 5.07 cd/m<sup>2</sup>, gap size  $\sim 0.53^\circ$ ), and the mask was four white dots appearing on the corners of a notional square ( $\sim 2.13 \times 2.13^\circ$ ) surrounding the target. The edge of the dots was separated from the edge of the target by  $\sim 0.23^\circ$ .

Each trial began with a white, centrally-presented, fixation cross. Observers initiated the search display by pressing the spacebar. After a random delay of 500–700 ms the target appeared either alone (set-size one), or with eight accompanying distractor Landolt-Cs (set-size nine; Fig. 1A), in a  $3 \times 3$  matrix of possible locations ( $\sim 1.53^\circ$  separation horizontally and by  $\sim 1.68^\circ$  vertically). The location of the target was randomly determined for each trial with its position signalled by the mask. This search array, consisting of the target, mask, and distractors (if any) was presented for 80 ms and was followed by the mask alone for 0, 160, 240, 320, or 640 ms (Fig. 1A). Observers indicated the identity of the target (whether the Landolt-C surrounded by the mask had its gap missing on the left or the right) by pressing the left or right arrow key on the keyboard respectively (unsped response). There were 50 trials



**Fig. 1.** (A) Schematic illustration of a trial in Experiment 1. (B) Mean proportion correct as a function of set size and trailing mask duration. Error bars represent standard errors of the mean. SS1 = set size one, SS9 = set size nine.

per condition (set-size  $\times$  mask duration, randomly-intermixed), yielding 500 trials in total.

The University of Queensland's School of Psychology Ethics Committee approved the experimental protocol. For this and subsequent experiments, observers were mostly recruited from a pool of first-year undergraduate psychology students who participated in exchange for course credit. This pool consisted predominately of young female adults (e.g., 18–22 years old). For this experiment, data were collected from 41 observers who gave informed consent prior to participation. One observer's data were excluded from the analysis, because they performed near chance (<.55 proportion correct) at set-size nine, 0 ms mask duration).

## 2.2. Results and discussion

Trials on which an observer made an invalid response (one that was not either of the possible response options) were excluded from analysis. Typically these invalid responses arose from errant presses of the spacebar used to initiate each trial. This occurred on an average of 1% of trials. Fig. 1B plots mean proportion correct on the target identification task as a function of set size and mask duration. Examination of the figure shows that performance was near ceiling across the mask durations at set-size one, as is typical in OSM studies. In contrast, performance varied substantially across the mask durations at set-size nine. Consistent with this, a repeated measures ANOVA indicated significant main effects of set-size,  $F(1,39) = 325.31$ ,  $p < .001$ ,  $\eta_p^2 = .893$ , and mask duration,  $F(4,156) = 25.63$ ,  $p < .001$ ,  $\eta_p^2 = .397$ . The interaction be-

tween mask duration and set-size was also significant,  $F(4,156) = 18.22$ ,  $p < .001$ ,  $\eta_p^2 = .318$ . To determine whether OSM and recovery had occurred, we conducted follow-up comparisons at set-size nine. To protect against making a Type I error, we adjusted our alpha-level to correct for the resulting four comparisons using the Bonferroni procedure (masking: 0 vs. 160 ms; recovery: 160 vs. 640 ms, 240 vs. 640 ms; 320 vs. 640 ms) which yielded a critical alpha level of .0125. These follow-up comparisons revealed significant masking, as indexed by a decrease in target accuracy from the 0 to the 160 ms mask duration,  $t(39) = 8.60$ ,  $p < .001$ ,  $\eta_p^2 = .655$ , followed by significant recovery, as indexed by a significant improvement in target accuracy from the 160 to 640 ms mask duration condition,  $t(39) = 2.83$ ,  $p = .007$ ,  $\eta_p^2 = .171$ . Performance at 240 ms and 640 ms did not significantly differ ( $t < 1$ ), nor did performance at 320 and 640 ms,  $t(39) = 1.84$ ,  $p = .074$ ,  $\eta_p^2 = .080$ .

Although these results closely replicate those obtained by Goodhew et al. (2011), it is notable that peak masking here occurred at a shorter duration (160 ms vs. 320 ms in earlier work). This raises the question of why earlier studies, which used mask durations greater than 160 ms (e.g., Di Lollo et al., 2000), failed to obtain recovery. A close examination of the data suggests that individual variability in masking functions is the most likely explanation. For example, in Experiment 1A of Goodhew et al. (2011), three trained observers completed approximately 200 trials for each data point (2304 trials overall). Nevertheless, despite this extensive measurement, two observers showed maximal masking at 320 ms, whereas the other showed it at 160 ms. Similarly, in the present experiment, although

the majority of participants (47.5%) showed peak masking at 160 ms, a substantial portion (20%) reached maximal masking at 320 ms. It should also be noted that the majority of OSM studies have compared a single delayed mask offset duration with a simultaneous mask duration baseline to assess masking. Without a means of measuring peak masking in such paradigms, it is impossible to assess whether recovery was obtained.

Having confirmed the temporal parameters of masking and recovery with our particular OSM paradigm, in Experiment 2, we investigated the role of mask offsets in the recovery after prolonged mask exposure (e.g., Goodhew et al., 2011). Specifically, we tested whether recovery arises from increased temporal separation between target and mask offsets as mask duration increases or whether the offset of the mask plays no role in the recovery effect (see Breitmeyer & Kersey, 1981; Macknik & Livingstone, 1998).

### 3. Experiment 2

In Experiment 2, we replicated the method of the first experiment, but used a tone as a signal to respond at a designated duration after the target, while leaving the mask on the display until a response was made. If recovery were to occur under these conditions, it would rule out temporal separation of offsets as a viable explanation of the effect as the mask never offset before the participants responded.

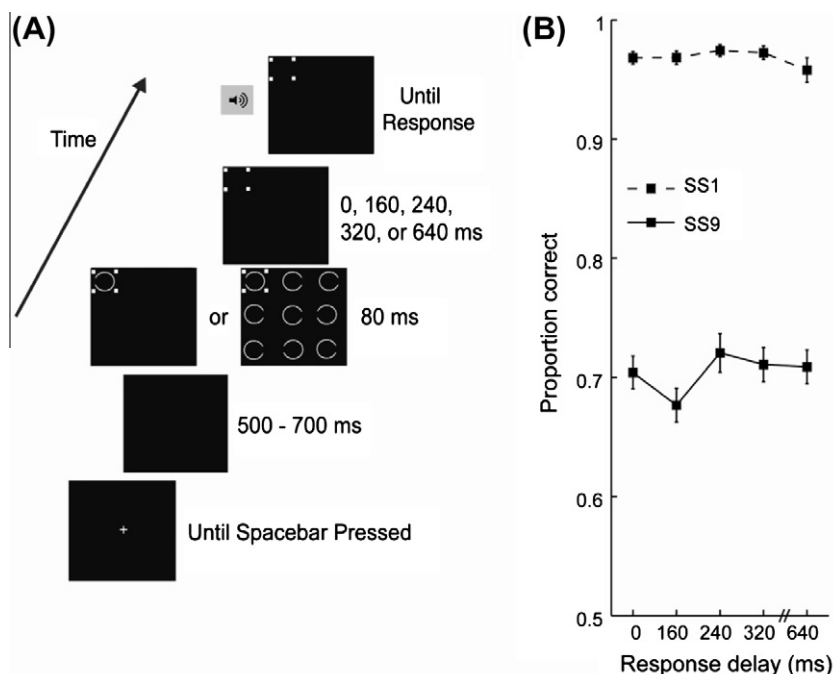
#### 3.1. Method

The stimuli and procedure replicated Experiment 1 with the following exceptions. After a designated duration (0,

160, 240, 320, or 640 ms), a 50 ms tone (787.5 Hz) sounded via headphones and observers were instructed to make the target identification response after this tone. The mask display terminated after the response was made. Thus, for this experiment, 'response delay' refers to the duration of the mask prior to the auditory cue to respond. These durations, in turn, correspond to the mask durations used in Experiment 1. That is, in both Experiments 1 and 2 we systematically varied the duration between target offset and the cue to respond regarding target identity. However, whereas in Experiment 1, this cue was mask offset, in Experiment 2, an auditory signal served as the response cue, and the mask remained visible until the response was complete. This allowed us to investigate the time-course of masking in the absence of a mask offset. There were 50 trials per condition (set-size  $\times$  response delay, randomly-intermixed), yielding 500 trials. We tested 43 observers who gave informed consent to participate. Data from three observers were excluded from analysis due to near-chance accuracy ( $<.55$ ) at set-size nine, 0 ms mask duration.

#### 3.2. Results and discussion

Trials on which an observer made an invalid response were not analysed (this occurred on an average of 5.5% of trials). Fig. 2B plots target identification accuracy as a function of set-size and response delay. From this it can be seen that at set-size one, there is little or no masking and performance is near ceiling. In contrast, at set-size nine, performance varies across the response delays. A repeated measures ANOVA demonstrated significant main effects of set-size,  $F(1, 39) = 695.58$ ,  $p < .001$ ,  $\eta_p^2 = .947$ , and mask



**Fig. 2.** (A) Schematic illustration of a trial in Experiment 2. (B) Mean proportion correct as a function of set size and trailing mask duration. Error bars represent standard errors of the mean. SS1 = set size one, SS9 = set size nine.

duration,  $F(1,4) = 2.48$ ,  $p = .046$ ,  $\eta_p^2 = .060$ . However, the interaction of these factors was not significant,  $F(4,156) = 1.89$ ,  $p = .116$ ,  $\eta_p^2 = .046$ .

Although we did not obtain a significant interaction, we conducted follow-up comparisons at set-size nine to test for masking and recovery. Our rationale was twofold. First, it seems likely that the non-significant set-size  $\times$  mask interaction arose because strong masking occurred at all response delays across at set-size nine. This contrasts sharply with Experiment 1, where little masking was obtained in the 0 ms mask duration condition. The source of this difference lies in the fact that the mask remained on the display in Experiment 2 until a response was made, thus leading to masking even at the 0 ms response delay. Second, because Experiment 1 allowed us to identify the temporal parameters of the masking and recovery functions in our task, we felt comfortable proceeding with planned comparisons here in Experiment 2. Finally, since only one comparison was planned to probe for recovery, we did not need to correct for multiple comparisons.

We did not use the 0–160 ms mask duration comparison to index masking in this experiment, because, as discussed above, the 0 ms mask duration does not represent a true ‘absence of trailing mask condition’, as it did in Experiment 1. Instead, the presence of masking can be inferred from the main effect of mask duration. A planned comparison revealed significant recovery, with performance superior at the 640 compared to the 160 ms response delay,  $t(39) = 2.24$ ,  $p = .031$ ,  $\eta_p^2 = .114$ . In sum, there was recovery in the absence of a mask offset, suggesting offsets are not necessary for the effect to occur. The present experiment also rules out the mask-offset induced transient-on-sustained inhibition account as an explanation for masking and recovery in OSM (Breitmeyer & Kersey, 1981; Breitmeyer & Ögmen, 2006). To recap, this theory suggests that the offset of the mask triggers transient channel activity, which inhibits the prior sustained channel response to the target, leading to masking. Therefore, it would predict that in the absence of mask offset, and with target and mask onset constant (common-onset), masking should not vary across our response delay durations – a pattern of results that was not observed.

In order to assess whether the recovery obtained without the mask offset was qualitatively different to that obtained with the mask offset, we examined set-size nine performance at comparable mask and target delays between Experiments 1 and 2. Note that this comparison did not include the 0 ms delay durations as these yielded very different presentation conditions and target performance levels for the reasons noted above. This analysis revealed a significant main effect of mask duration/response delay,  $F(3,234) = 6.70$ ,  $p < .001$ ,  $\eta_p^2 = .079$ , but no interaction with experiment ( $F < 1$ ). This demonstrates that the pattern of recovery observed did not differ statistically between experiments with or without a mask offset, further confirming that an interaction between target and mask offsets cannot account for recovery from OSM.

In sum, Experiment 2 demonstrated that recovery from OSM does not arise from increasing temporal separation between the offsets of target and mask (see Breitmeyer & Kersey, 1981; Macknik & Livingstone, 1998). However, an

outstanding question remains: what is the role of the prolonged mask in recovery? For example, it may be that with prolonged mask exposure, the visual system adapts to the mask, allowing the target to gain relative competitive strength and thus to be perceived. Alternatively, it might be that prolonged mask exposure acts as a cue to the visual system to individuate the target and mask as separate object representations (see Lleras & Moore, 2003; Moore & Lleras, 2005), thus releasing the target from the suppressive effect of the mask. A final possibility is that prolonged physical exposure of the mask is not instrumental, and that recovery is instead an inherently time-based phenomenon. Experiment 3 was designed to investigate these possibilities.

#### 4. Experiment 3

Experiment 3 was identical to Experiment 2, except that mask duration was held constant (160 ms) and an auditory response cue was presented at a variable interval following target offset.

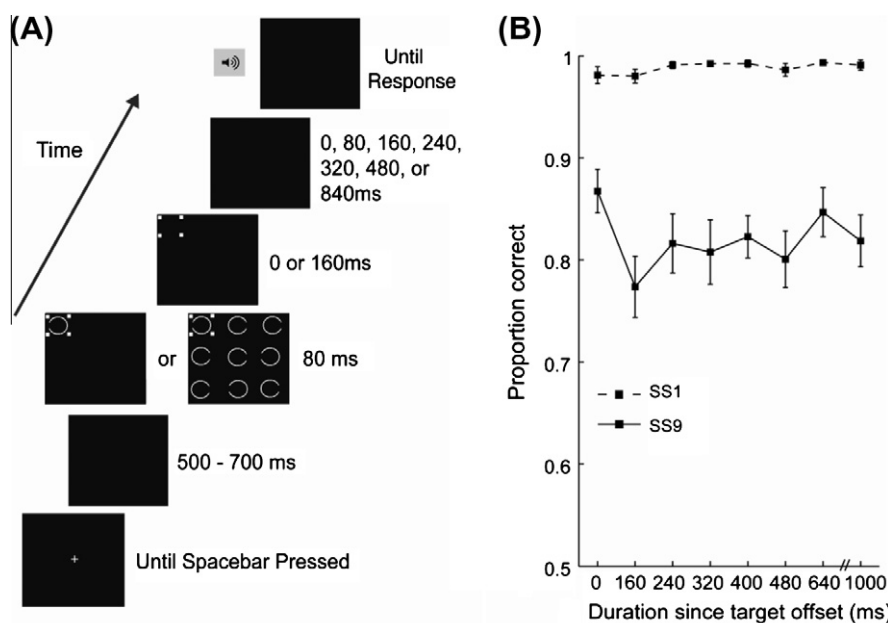
##### 4.1. Method

The search arrays were identical to those used in Experiment 1. In the simultaneous offset condition, the mask offset simultaneously with the target array, and a tone prompted observers to respond immediately (0 ms) after mask offset. For all other conditions, the mask always trailed for 160 ms, the screen was then blank for an interval of 0, 80, 160, 240, 320, 480, or 840 ms, after which an auditory response cue was given. This yielded durations since target offset (DSTOs; mask duration + ISI) of 160, 240, 320, 400, 560, 640, and 1000 ms. There were 60 trials per condition (2 set-sizes  $\times$  8 DSTOs = 16 conditions), producing 960 trials in total. This experiment was programmed in Matlab using the Psychophysics Toolbox extension (Brainard, 1997). We tested 12 observers who gave informed consent prior to participation. Data from one subject were excluded due to near-chance (<.55) performance at set-size nine, 0 ms DSTO.

##### 4.2. Results and discussion

Trials for which an invalid response was registered were excluded from the analysis. Invalid responses could result from a response being made before the completion of the response delay or a spacebar response to initiate the next trial, being registered as the target identification response. This occurred on an average of 4.26% of trials. Fig. 3 shows target identification accuracy plotted as a function of set-size and response delay duration. Examination of this figure shows a pattern of results mirroring that of the first two experiments – masking followed by recovery.

A repeated measures ANOVA revealed a significant main effect of set-size,  $F(1,10) = 55.55$ ,  $p < .001$ ,  $\eta_p^2 = .847$ , and a significant main effect of mask duration,  $F(7,70) = 4.05$ ,  $p = .001$ ,  $\eta_p^2 = .288$ . The interaction between these variables was also significant,  $F(7,70) = 4.25$ ,  $p = .001$ ,



**Fig. 3.** (A) Schematic illustration of a trial in Experiment 3. (B) Mean proportion correct as a function of set size and trailing mask duration. Error bars represent standard errors of the mean. SS1 = set size one, SS9 = set size nine.

$\eta_p^2 = .298$ . Again, masking peaked at 160 ms (0 ms DSTO vs. 160 ms DSTO,  $t(10) = 4.60$ ,  $p = .001$ ,  $\eta_p^2 = .679$ ), followed by recovery that was at its maximum at the 640 ms DSTO (160 ms DSTO vs. 640 ms DSTO,  $t(10) = 2.81$ ,  $p = .019$ ,  $\eta_p^2 = .441$ ).

The observed recovery effect in the absence of prolonged mask exposure converges with the results of Experiment 2 to demonstrate that offsets are not driving the recovery effect (since here target-mask offset was constant for all but the first DSTO, and yet recovery was observed). In addition, the present findings show that prolonged exposure of the mask is not a necessary condition for recovery. This last point indicates that recovery is mediated neither by adaptation to a prolonged mask, nor because prolonged exposure of the mask serves as cue to individuate the mask as a separate object, thus releasing the target from masking (see Lleras & Moore, 2003; Moore & Lleras, 2005).

A unique aspect of the results here is the drop in accuracy between the 640 and 1000 ms DSTO conditions. Examination of the data suggests this drop does not reflect a perceptual phenomenon, but rather the observers' failure to comply with task instructions. While the overall proportion of invalid responses was modest, a disproportionate number of these errors (which typically reflected pressing a response key prior to the tone) occurred at the longest DSTO. For example, at the set-size nine, 1000 ms DSTO, invalid responses occurred on 19.24% of trials, more than double the 8.03% of trials on which it occurred at the 640 ms DSTO. For these reasons, we excluded the 1000 ms DSTO from the analysis, and any subsequent analyses were performed on the remaining DSTOs.

It is also worthwhile to discuss the relationship between our results and those obtained by Lachter and Dur-

gin (1999) using metacontrast masking. In their experiment, observers were first presented with a target disc and an annulus mask on either side of a central fixation. Following a variable duration blank screen, a second frame was then presented which consisted of an annulus mask at both locations. Observers were required to report the location of the target following a short (<480 ms) or long response delay (>730 ms). The critical result was that masking decreased monotonically with increasing target-mask SOA at the short response delay, but varied in a U-shaped fashion with SOA at the long response delay. The authors attributed this to the differential latencies of two stages of processing: a fast contour processing stage, followed by a prolonged surface "filling in" stage. In the short response delay condition, observers would only have access to contour information, which would make it difficult to differentiate the target + mask display on one side of fixation from the mask + mask display on the other. What differentiated the two displays was the surface information – which was only available later – hence conferring an advantage to the long response delay condition (Lachter & Durgin, 1999).

In OSM, mask duration, not SOA, is manipulated. Thus, we cannot directly compare the results of Lachter and Durgin (1999) to ours, except at their 0 ms target-mask SOA (i.e., simultaneous onset) condition which corresponds to the parameters of our own Experiment 3. Here, our results correspond nicely with those of Lachter and Durgin (1999) with reduced masking at long (480 ms) relative to short (0 ms) response delays. Despite the similar pattern of results, however, we suspect the mechanisms underlying our effect are quite different from those proposed by Lachter and Durgin (1999). This follows from the fact that we obtained OSM in the absence of substantial adjacent tar-

get-mask contours, using a target without an appreciable surface area, and employing a task that did not require judgements about target surfaces.

## 5. General discussion

In this study we employed OSM to explore the dynamics of the competition for consciousness among stimuli that appear in close spatiotemporal proximity. Specifically, we tested several possible explanations for the OSM recovery effect observed by Goodhew et al. (2011). In Experiment 1 we replicated the recovery effect and established its temporal parameters. Experiment 2 examined the role of mask offsets in producing recovery. Eliminating the mask offset yielded recovery that was statistically indistinguishable from that observed when the mask did offset, thus eliminating offsets as a necessary factor for recovery. The fact that there was significant masking in this condition also eliminates a dual-channel account (Breitmeyer & Ganz, 1976) for OSM more generally (see above). Finally, in Experiment 3, we showed that recovery occurred in the absence of prolonged mask exposure, when observers simply delayed their responses, demonstrating that the recovery is neither due to adaptation nor temporal cues derived from the prolonged exposure of the mask.

Taken together, these results demonstrate that recovery is an inherently time-based phenomenon in which properties of the mask itself play only a minimal role. Elsewhere, Goodhew et al. (2011) have argued that the importance of time in recovery from OSM stems from the need to re-establish iterative re-entrant processing on a lingering target representation in order to consolidate this information and render it explicitly reportable. When a target has been successfully masked, and thus has its representation degraded or suppressed, this “second cascade” of processing is required to recover the target information. However, before discussing this hypothesis further, it is important to consider other explanations for our findings that do not draw on re-entrant processing.

As discussed above, when OSM was first demonstrated, Di Lollo et al. (2000) firmly grounded their explanation in terms of re-entrant processing. However, the necessity of re-entry mechanisms to account for OSM has been questioned. For example, Francis and Hermens (2002) argued that existing metacontrast masking models (including Bridgeman, 1971, 1978; Francis, 1997; Weisstein, 1968) could account for the pattern of masking reported by Di Lollo et al. (2000). But this was only true if the simulations operationalized attention indirectly by varying the mask intensity parameter. As noted by Di Lollo, Enns, and Rensink (2002), the validity of this change is open to criticism on the grounds that it is post hoc and may not accurately capture the construct of attention.

Another possible alternative explanation arises from the work of Francis (2000) who considered computational mechanisms which could account for U-shaped masking as a function of target-mask SOA under conditions of metacontrast. This paper did an excellent job of simulating the mechanisms underlying metacontrast, including – efficient masking (where the onset of the mask is optimally effec-

tive at suppressing target perception when it occurs later rather than earlier, due to the time required for consolidation of the target in the network), mask blocking (where the strength of the activity representing the target in the network prevents a weak mask from being effective shortly after the target, but the mask becomes increasingly effective as the target activity decays), and target blocking (where the effect of the mask is so strong that it thwarts target processing, but again, not immediately after the target). However, it is unlikely that this framework could encompass recovery in OSM because it assumes a temporal asynchrony between the target and mask onsets in order to explain masking, a feature that is notably absent in OSM.

More recently, Bridgeman (2007) modified his recurrent lateral inhibition model of metacontrast, which is a single-layer network that employs distributed stimulus representations (e.g., Bridgeman, 1971) to account for OSM. According to this model, each neuron inhibits and in turn is inhibited by neighbouring neurons at recurrent 30 ms intervals. The duration of these iterations is based on the latency of reciprocal lateral inhibition in the cat lateral geniculate nucleus cells (Singer & Creutzfeldt, 1970). A key feature of this model is that it explains metacontrast by analysing collective distributed activities within the network, rather than by the response patterns of single cells.

In modifying this framework to account for OSM, focused attention was simulated by varying the number of iterations over which the neural net activity is collected. This change not only maps nicely onto the putative role of attention suggested by Di Lollo et al. (2000), but also seems reasonable on the grounds that an attended stimulus will be responded to faster than an unattended stimulus. Using this model, Bridgeman (2007) was able to reproduce masking functions that resembled the psychophysical data from Di Lollo et al. (2000), and, critically, yielded U-shaped masking functions under some conditions, similar to those seen here.

In the present context, it is obviously interesting that Bridgeman's simulation produced a U-shaped function with OSM; however as it stands, this model cannot account for the conditions under which recovery is observed. First, according to Bridgeman's model, recovery occurs with fewer iterations (strongly attended), but not with many iterations (less attended). Of course, it is possible that the magnitude of recovery varies as a function of distributed attention (i.e., set-size), but we cannot be certain based on the present experiments because they each had only one set-size at which recovery could be detected (set-size 1 data constrained by ceiling effects). However, it is clear that a strongly attended target cannot be a necessary condition for recovery, since we found the effect under conditions of distributed attention (at set-size 9). Also, recovery varies with response delay (or duration since target offset; DSTO), and it is unclear how Bridgeman's model could account for these findings. Perhaps DSTO could vary with the ‘number of iterations’ parameters, but then the model would seem to give the opposite result. That is, we should see recovery with more rather than fewer iterations (prolonged relative to shorter DSTOs).



Taken together, the analyses above suggest that alternative theories not based on iterative processing fail to account for recovery. However, it remains to be shown that our notion of a second cascade is a viable option. In particular, if a second series of iterative loops gives rise to the prolonged time course of recovery, shown here and elsewhere (Goodhew et al., 2011), a representation of the target not only must survive initial competition with the trailing mask, but also persist for over half a second. Below we consider the issue of what sort of representation would be necessary in order for it to persist for a sufficient duration to drive some form of recovery based on a second re-entrant cascade.

Conventional accounts of persistence suggest it can take two forms: visible and informational (Coltheart, 1980). Visible persistence is thought to arise from a briefly lingering low-level stimulus representation. This follows from repeated demonstrations that the strength and duration of visible persistence is closely tied to the physical properties (e.g., duration, luminance) of the inducing stimulus. Most importantly for the present purposes, this form of persistence seems to be reliably extinguished within about 100–200 ms after stimulus offset (Coltheart, 1980; Di Lollo, 1980; Hogben & Di Lollo, 1974). Thus, visible persistence cannot underlie recovery.

By contrast, informational persistence is not locked to the physical properties of the stimulus, and lasts for at least several hundred milliseconds. Informational persistence can be equated with iconic memory, such as that measured in classic partial-report tasks (Coltheart, 1980; Loftus & Irwin, 1998). Here, observers are flashed an array of stimuli (e.g., letters), and a cue indicating which line to report is presented after the offset of the array. This method reveals that observers do indeed have access to rich and spatially-accurately representations of stimuli after they offset, because they can report any line of the array when cued, but that this representation dissipates with time. In addition, the duration of informational persistence is largely unrelated to the physical properties of the array (Coltheart, 1980; Sperling, 1960). Thus, on the face of it, the properties of informational persistence closely resemble those required for recovery. However, there are aspects of recovery which are difficult to reconcile with the notion of informational persistence. Most critically, informational persistence has been hypothesised to reflect a high-level, abstract representation, and yet the representation in OSM is sufficiently fragile that it is susceptible to a spatially-overlapping backward mask – a feature that is inconsistent with an abstract representation.

In light of the evidence that representations of recovered targets are neither low-level and sensory in nature, nor entirely abstract, we propose that these stimuli are stored as weak (unconscious) representations in visual short-term memory (VSTM). It is known that stimuli in VSTM are vulnerable to interference from backward masking (Makovski, Sussman, & Jiang, 2008; Pashler, 1988; Phillips, 1974; Sligte, Scholte, & Lamme, 2008; Sligte, Wokke, Tesselaaar, Scholte, & Lamme, 2011). Thus, if targets are weakly represented in VSTM, this would explain why recovery is eliminated when a spatially-overlapping mask is used in place of the trailing four-dot mask. Namely, a

trailing noise mask produces more potent interference: its greater size, signal strength, and spatial overlap allow it to obliterate the target representation in VSTM, leaving no substance for recovery. In contrast, the trailing four-dot mask produces interference initially, but does not affect the target representation in VSTM. Thus, with sufficient time for consolidation of the mask, target information can become available again, permitting recovery to occur.

The present findings (and recovery effects in general) not only have implications for the nature of the representations underlying OSM, they also provide further grounds for distinguishing OSM from other forms of masking more generally. The relative insensitivity of recovery to mask duration is particularly striking when considered in light of repeated demonstrations of the critical role of mask stimulus factors, such as its temporal duration, offset, or relation to the target (e.g., Breitmeyer & Ögmen, 2006) in other forms of masking. Instead, in OSM, once a mask was of sufficient duration (160 ms) to impair target visibility, it then did not matter whether it offset immediately, was presented for a further 480 ms, or disappeared following response. Recovery occurred irrespective of such changes. This demonstrates that the recovery from OSM cannot merely be another demonstration of already-established masking phenomena such as the impact of offsets on stimulus perception (see Macknik & Livingstone, 1998), or the rescuing effect of object individuation cues (see Lleras & Moore, 2003; Moore & Lleras, 2005). Rather, the evidence is compelling that OSM reflects prolonged high-level competition that occurs between mental representations of stimuli, which occurs even in the absence of new visual inputs.

This point may be further illustrated by comparing recovery from OSM with a seemingly-related phenomenon in metacontrast masking known as “target disinhibition”. Here, a second mask presented adjacent to the initial masking stimulus leads to improved target perception (e.g., Breitmeyer, Rudd, & Dunn, 1981; Dember & Purcell, 1967; Robinson, 1966). Although both forms of recovery possess obvious behavioural similarities, there is strong evidence that target disinhibition occurs at a much earlier locus. First, disinhibition is produced by the onset of a new stimulus (exogenously triggered), whereas, as noted above, recovery from OSM can occur in the absence of a new object onset, and is unaffected by mask offset and mask duration. Second, whereas disinhibition appears to depend on the “mask being masked”, recovery in OSM depends on the passage of time and the presence of an intact target representation (see Goodhew et al., 2011). Third, target disinhibition and recovery from OSM operate over disparate timescales: disinhibition typically occurs with very short (tens of milliseconds) intervals separating the two masks and at short mask durations. Recovery from OSM, on the other hand, takes over half a second after target offset to emerge reliably. Finally, whereas all the evidence suggests that OSM taps re-entrant processes such as those proposed by Di Lollo et al. (2000; see Dux et al. (2010) for confirmatory evidence of the role of re-entrant processes in OSM), target disinhibition does not occur with interocular presentation, implicating a pre-cortical locus for the effect (Robinson, 1968).

## 5.1. Conclusion

The recovery from OSM reported by Goodhew et al. (2011) is invariant to the offset of the mask, and even the prolonged physical exposure of this trailing stimulus. These findings rule out explanations of recovery based on the temporal separation of target and mask offsets, mask adaptation, and temporal object-individuation cues due to prolonged mask exposure. Instead, it solidifies recovery as a high-level visual-cognitive phenomenon, which is inherently tied to target-processing time. This reveals the especially prolonged iterative temporal dynamics of conscious object perception.

## References

- Alpern, M. (1952). Metacontrast: Historical review. *American Journal of Optometry*, 29, 631–646.
- Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America*, 43, 648–657.
- Benjamin, L. T. J. (2009). *A history of psychology: Original sources and contemporary research*. Oxford: Blackwell Publishing.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 15, 205–206.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- 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, 1–36.
- Breitmeyer, B. G., & Kersey, M. (1981). Backward masking by pattern stimulus onset. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 972–977.
- Breitmeyer, B. G., & Ögmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (2nd ed.). Oxford University Press.
- Breitmeyer, B. G., Rudd, M., & Dunn, K. (1981). Metacontrast investigations of sustained-transient channel inhibitory interactions. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 770–779.
- Bridgeman, B. (1971). Metacontrast and lateral inhibition. *Psychological Review*, 78, 528–539.
- Bridgeman, B. (1978). Distributed sensory code applied to simulations of iconic storage and metacontrast. *Bulletin of Mathematical Biology*, 40, 605–623.
- Bridgeman, B. (2007). Common-onset masking simulated with a distributed-code model. *Advances in Cognitive Psychology*, 3, 33–40.
- Broadbent, D. E. (1958). *Perception and communication*. Elmsford, New York: Pergamon Press.
- 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, 315–334.
- Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 109–127.
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception and Psychophysics*, 27, 183–228.
- Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, 8, 97–107.
- Dember, W. N., & Purcell, D. G. (1967). Recovery of masked visual targets by inhibition of the masking stimulus. *Science*, 157, 1335–1336.
- Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, 109, 75–97.
- 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, 481–507.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2002). Object substitution without reentry? *Journal of Experimental Psychology: General*, 131, 594–596.
- Dux, P. E., & Marois, R. (2009). How humans search for targets through time: A review of data and theory from the attentional blink. *Attention, Perception, and Psychophysics*.
- 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, 1242–1247.
- Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science*, 8, 135–139.
- Flombaum, J. K., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 840–853.
- Francis, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review*, 104, 572–594.
- Francis, G. (2000). Quantitative theories of metacontrast masking. *Psychological Review*, 107, 768–785.
- 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, 590–593.
- Goodhew, S. C., Visser, T. A. W., Lipp, O. V., & Dux, P. E. (2011). Competing for consciousness: Prolonged mask exposure reduces object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 588–596.
- Hogben, J. H., & Di Lollo, V. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vision Research*, 14, 1059–1069.
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. *The Journal of Physiology*, 148, 574–591.
- Kaheman, D. (1968). Method, findings, and theory in studies of visual masking. *Psychological Bulletin*, 70, 404–425.
- Kahneman, D. (1973). *Attention and effort*. New Jersey: Prentice-Hall.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219.
- Kastner, S., & Pinsk, M. A. (2004). Visual attention as a multilevel selection process. *Cognitive, Affective, and Behavioral Neuroscience*, 4, 483–500.
- Koch, C. (2004). *The quest for consciousness: A neurobiological approach*. Englewood: Roberts and Company Publishers.
- Lachter, J., & Durgin, F. H. (1999). Metacontrast masking functions: A question of speed? *Journal of Experimental Psychology: Human Perception and Performance*, 25, 936–947.
- 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, 106–120.
- Loftus, G. R., & Irwin, D. E. (1998). On the relations among different measures of informational and visible persistence. *Cognitive Psychology*, 35, 135–199.
- Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience*, 1, 144–149.
- Makovski, T., Sussman, R., & Jiang, Y. H. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 369–380.
- Moore, C. M., & Lleras, A. (2005). On the role of object representations in substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1171–1180.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Pashler, H. (1988). Familiarity and the detection of change in visual displays. *Perception and Psychophysics*, 44, 369–378.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics*, 16, 283–290.
- Ramachandran, V. S., & Cobb, S. (1995). Visual attention modulates metacontrast masking. *Nature*, 373, 66–68.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849–860.
- Robinson, D. N. (1966). Disinhibition of visually masked stimuli. *Science*, 154, 157–158.
- Robinson, D. N. (1968). Visual disinhibition with binocular and interocular presentations. *Journal of the Optical Society of America*, 58, 254–257.
- Scholl, B. (2002). *Objects and attention*. Cambridge, MA: MIT Press.
- Shelley-Tremblay, J., & Mack, A. (1999). Metacontrast masking and attention. *Psychological Science*, 10, 508–515.
- Singer, W., & Bedworth, N. (1973). Inhibitory interaction between X and Y units in the cat lateral geniculate nucleus. *Brain Research*, 49, 291–307.

- 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*, 311–330.
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores. *PLoS ONE*, *3*, e1699.
- Sligte, I. G., Wokke, M. E., Tesselaar, J. P., Scholte, H. S., & Lamme, V. A. F. (2011). Magnetic stimulation of the dorsolateral prefrontal cortex dissociates fragile visual short-term memory from visual working memory. *Neuropsychologia*, *49*, 1578–1588.
- Sperling, G. (1960). The information available in brief presentations. *Psychological Monographs: General and Applied*, *74*, 1–29.
- Weisstein, N. (1968). A Rashevsky–Landahl neural net: Simulation of metacontrast. *Psychological Review*, *75*, 494–521.