

**Categorical information influences conscious perception: An interaction between
object-substitution masking and repetition blindness**

Stephanie C. Goodhew¹, John A. Greenwood², and Mark Edwards¹

¹Research School of Psychology, The Australian National University

²Experimental Psychology, University College London

Word count: (main text): 8,106

Corresponding Author: Stephanie C. Goodhew

Address: Research School of Psychology (Building 39)

The Australian National University, Canberra, 0200

Email: stephanie.goodhew@anu.edu.au

Running head: OSM and RB

Abstract

The visual system is constantly bombarded with dynamic input. In this context, the creation of enduring object representations presents a particular challenge. Here we used object-substitution masking (OSM) as a tool to probe these processes. In particular, we examined the effect of target-like stimulus repetitions on OSM. In visual crowding, the presentation of a physically-identical stimulus to the target reduces crowding and improves target perception, whereas in spatial repetition blindness, the presentation of a stimulus that belongs to the same category (type) as the target impairs perception. Here, across two experiments, we found an interaction between spatial repetition blindness and OSM, such that repeating a same-type stimulus as the target increased masking magnitude relative to presentation of a different-type stimulus. These results are discussed in the context of the formation of object files. Moreover, the fact that the inducer only had to belong to the same ‘type’ as the target in order to exacerbate masking, without necessarily being physically-identical to the target, has important implications for our understanding of OSM per se. That is, our results show the target is processed to a categorical level in OSM despite effective masking, and, strikingly, demonstrate that this category-level content directly influences whether or not the target is perceived, not just performance on another task (as in priming).

Keywords: object-substitution masking; object file; repetition blindness; object individuation; type-token individuation; crowding; object perception; visual masking; perception; consciousness.

As we navigate through the visual world, the brain is confronted with incomplete input due to both internal and external factors. Internal factors include the transitory blindness experienced during eye movements (Burr, Morrone, & Ross, 1994; Ibbotson & Krekelberg, 2011; Irwin & Brockmole, 2004; Matin, 1974), and our attentional capacity, which limits our ability to process more than a fraction of the information available at a given point in time (Broadbent, 1958; Desimone & Duncan, 1995; Kastner & Pinsk, 2004). External factors additionally affect the quality of this input: objects move within a scene, enter and exit scenes, and appear different at different points in time (e.g. due to viewpoint variation). From this impoverished input, the visual system is nonetheless able to create the stable and coherent visual scene that we consciously perceive. In order to achieve this feat, the brain must carve up this noisy input and parse it into discrete object representations. This is particularly difficult with dynamic input, where the incoming information could derive from either multiple distinct objects ('object individuation') or a single object that maintains its identity despite changes in its appearance and/or location. Such inferences occur pre-consciously, but can ultimately determine the contents of conscious perception. Note that such inferences go beyond mere temporal integration, which is strongly determined by physical presentation parameters and sensory limitations (Coltheart, 1980; Di Lollo, 1980). Object individuation is instead about how objects should be consciously represented (e.g., a single object continuing over time), even when the temporal resolution of the system is sufficient to encode the discrete events across time (e.g., sufficient to recognise that two instances of an object occurred, but the inference is about whether this belongs to the same object occurring in different locations over time versus two discrete objects).

Object-substitution masking (OSM) has been used as a paradigm to gauge such inferences about the creation of object representation in the face of dynamic and impoverished input. In OSM, perception of a briefly-presented target surrounded by four-dots is intact when

all elements offset simultaneously, and impaired (‘masked’) when the offset of the four-dots is delayed in time (Di Lollo, Enns, & Rensink, 2000; Enns & Di Lollo, 1997; for a review see Goodhew, Pratt, Dux, & Ferber, 2013). Historically, target-like distractors have been presented in addition to the target (which appeared and disappeared simultaneously with the target). However, recent evidence indicates that the presence of distractors is not necessary to obtain masking (Argyropoulos, Gellatly, Pilling, & Carter, 2013; Filmer, Mattingley, & Dux, 2014, 2015; Pilling, Gellatly, Argyropoulos, & Skarratt, 2014). OSM clearly reflects a failure to register the target as an enduring object for conscious perception. There are, however, two different theoretical accounts as to how this outcome is produced: object substitution versus object-updating. We will now consider each of these in turn, and then review their similarities and difference within the object token framework.

According to the object-substitution account of OSM, masking results when conflict between representations in higher level (extrastriate) areas and sensory input in lower-level areas (e.g., V1) is resolved in favour of the latter (Di Lollo, 2010; Di Lollo et al., 2000; Enns & Di Lollo, 2002). More specifically, information representing the target plus four dots is fed forward from posterior regions into anterior regions, which generates one or more perceptual hypotheses about the identity of the stimulus. However, given the brief presentation of the stimulus, coupled with the increasing size of the receptive fields in more anterior regions, which results in poorer spatial resolution, there is ambiguity about the identity of the stimulus. This triggers re-entrant processing to compare the representation against more high-resolution sensory information in lower visual areas, where a matching process occurs to refine this representation and resolve this ambiguity. Then, if the display remains unchanged throughout this re-entrant loop, the perceptual hypothesis can remain the same, and stable a percept can be achieved. If, however, the display changes (to the mask-alone stimulus), then a mismatch occurs between the descending signals and the incoming sensory input, and thus a new a

perceptual cycle is initiated. Ultimately, masking occurs when the perceptual representation containing the target is discarded in favour of one containing the mask alone (Di Lollo, 2010; Di Lollo et al., 2000; Enns & Di Lollo, 2000). This innovative and insightful model was part of the broader movement that overthrew the historically dominant framework within neuroscience that emphasised the feedforward architecture, without appreciating the now widely-accepted importance of re-entrant processes and their role in visual and cognitive processes (e.g., Lamme, 2000). Furthermore, this re-entrant processing model of OSM has been supported independently by both functional magnetic resonance imaging (fMRI) (Weidner, Shah, & Fink, 2006) and event-related potential (ERP) evidence (Kotsoni, Csibra, Mareschal, & Johnson, 2007).

Since then, however, another account has been proposed, called the *object-updating account*. Unlike object-substitution, this model does not offer a specific neural explanation for how OSM arises, but instead focuses on cognitive-level explanations. That is, it derives from the traditional cognitive-psychological framework which hypothesises the existence of *object files* or *tokens*, which are transitory episodic representations of the spatiotemporal coordinates of an object that are responsible for maintaining coherent object identity over changes in time and space (Kahneman, Treisman, & Gibbs, 1992; Kanwisher & Driver, 1992). That is, according to the object-updating account, the target is initially represented, and then this token undergoes a process of updating, during which information about the mask-alone replaces the representation of the target within the existing object file (Lleras & Moore, 2003; Moore & Lleras, 2005; Pilling & Gellatly, 2010).

If we consider both the object substitution and the object updating theoretical accounts within the object file framework, then where the accounts converge is: an object file is formed for the target (plus four dots) in the first instance. And when masking is effective, there is an object file for the mask alone, but not the target: it is this end-state process – *the failure to form*

an enduring object file containing the target that we are interested in here, and which is common to both the object substitution and object updating accounts. Where the accounts diverge is thus: according to object substitution, the object file for the target (plus four dots) is discarded and a separate one for the mask alone is created, whereas according to object updating, the object file containing the target is updated to reflect the mask alone instead. The key distinction, then, between object-updating and object-substitution then is this: whether there are two separate object tokens (one for the target, one for the trailing mask), which compete for access to consciousness, or whether the interactions between the target and mask happen within one single object file. It is important to note that the object-updating hypothesis in no way undermines or precludes the possibility of re-entrant processing being implicated in masking, it simply specifies different cognitive-level machinations underlying the competition for consciousness.

It is not the purpose of this paper to adjudicate between the object substitution and object-updating theoretical accounts. Instead, we are interested in the process that OSM gauges, which both of these theoretical accounts have in common and agree on: the failure to maintain an enduring object-representation for the target. Much of the evidence in the literature, however, has been interpreted from the perspective of the object-updating framework, and we will review this evidence in the following section. However, proponents of the object-substitution account could equally interpret this as evidence for object substitution. For the present purposes, we will remain agnostic, and instead focus on the functional outcome of interest: the failure to have an enduring object file pertaining to the target, which the following evidence demonstrates that OSM measures.

Lleras and Moore (2003) found evidence that they interpreted as consistent with the object-updating framework, but from a broader perspective clearly illustrate the factors that influence whether the visual system maintains an enduring object file for the target. That is,

when the four-dot mask offsets simultaneously with the target but then subsequently reappears, masking occurs when the interval between the two presentations is conducive to apparent motion, linking the two presentations as belonging to a single object, whereas masking does not occur when the interval is not conducive to apparent motion (Lleras & Moore, 2003). This is consistent with the notion that OSM reflects the updating of object representations over time. Similarly, static manipulations that encourage this distinction between the target and mask can also thwart masking. For example, masking is reduced when the target and four-dot mask objects are different colours (Moore & Lleras, 2005), or luminance polarities (Luiga & Bachmann, 2008). Similarly, masking magnitude scales as a function of the difference in orientation between Gabor targets and masks (Goodhew, Edwards, Boal, & Bell, 2015). These findings have typically been interpreted within the object-updating framework, that is, physical similarity between the target and mask encourages the visual system to update the object file containing the target to reflect the mask. However, equally, the findings could also be interpreted within the object-substitution framework, in that similarity between the target and mask could be conducive to object substitution, encouraging the discarding of the object file containing the target. From a broader perspective, they show that OSM, and therefore the maintenance of an enduring object file for the target, is sensitive to physical similarity parameters.

Masking is also reduced by prior presentation ('preview') of the mask stimulus, even though this preview is not predictive of target location (Neill, Hutchison, & Graves, 2002). More recently, it has been shown that preview of placeholders at the location of target and distractors attenuates masking (Guest, Gellatly, & Pilling, 2012). These findings have also been interpreted within the object-updating framework as suggesting that preview of the non-target stimuli helps protect against alteration or updating of the target object file. Finally, the magnitude of masking is modulated by the temporal resolution of the visual system at the time

that the target and mask are presented. This can be achieved by altering the relative balance of the contribution of magnocellular (vs. parvocellular) neurons, which have superior temporal acuity (C.-M. Chen et al., 2007; Derrington & Lennie, 1984; Livingstone & Hubel, 1988; Maunsell et al., 1999). When the relative magnocellular contribution is increased by having the observer's two hands near the visual stimuli (for a review of hand proximity effects on perception, see Goodhew, Edwards, Ferber, & Pratt, 2015), OSM is reduced (Goodhew, Gozli, Ferber, & Pratt, 2013), whereas when the relative contribution of these cells is reduced via luminance pedestals (for a review of the effects of luminance pedestals on perception, see Pokorny, 2011) masking is increased (Goodhew, Boal, & Edwards, 2014). Collectively, then, this evidence indicates that OSM taps the (failure of) formation of an enduring object representation for a briefly-presented stimulus that is presented close in space and time to other stimuli. This evidence suggests that this formation process is affected by a number of factors, including physical similarity, the opportunity to pre-preview some stimuli as they will appear, and the temporal resolution of the system at the time of encoding.

In the present study, we were interested in how other objects in the visual scene would influence the object file formation processes that lie at the heart of OSM. This question was inspired by findings from crowding, another example of a limitation of visual awareness. Crowding occurs when the perception of a target stimulus, which would be visible if presented in isolation, is impaired by other surrounding items in close spatial proximity (Bouma, 1970; Pelli & Tillman, 2008). Though there are many models of crowding (for a review, see Whitney & Levi, 2011), a popular account characterises it as a pooling process that regularises the noisy representation of position in the periphery by averaging the target and flanker identities (Freeman & Simoncelli, 2011; Greenwood, Bex, & Dakin, 2009; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). Functionally, then, crowding may reflect the blurring and loss of

information across space in the way that OSM reflects the blurring and loss of information across time.

More direct evidence for the interaction between the two processes has been observed, suggesting a potential overlap in the mechanisms underlying the two phenomena. While, as flagged earlier, the presence of simultaneous distractors is not a necessary condition for OSM, and there have been absences of interactions between masking and set-size in OSM reported when the items are widely spaced (Argyropoulos et al., 2013; Filmer et al., 2014), when the distractors are sufficiently close to the target to induce crowding, then an interaction between masking magnitude and the number of distractors is obtained (Camp, Pilling, Argyropoulos, & Gellatly, 2015). This interaction between crowding and OSM suggests that there may be at least some overlap in mechanisms; otherwise purely additive effects would be expected.

We have thus far depicted OSM as a failure of object file continuation, and noted the similarities with visual crowding. In this context, crowding has been found to be strongly modulated by additional objects present in the visual field. Namely, the foveal presentation of an item that matches the identity of the peripheral target reduces crowding, an effect recently referred to as “remote uncrowding” (Geiger & Lettvin, 1986; Sayim, Greenwood, & Cavanagh, 2014). This remote uncrowding effect is specific only to a foveal “inducer” stimulus that is physically identical to the crowded target, and does not generalise to higher-order categorical stimuli (e.g., a foveal letter ‘A’ will uncrowd an uppercase target ‘A’ but not a lowercase ‘a’). That is, in crowding, a simultaneously presented stimulus identical to the target can help to recover the target and thwart the deleterious effects of crowding. Altogether then, OSM and crowding can both be seen as failures of the visual system to individuate unique objects: across space in crowding, across time in OSM. To what extent do these spatial and temporal domains have similar underlying mechanisms? If the reported interaction between crowding and OSM reflects a common underlying process, then we would predict that the remote uncrowding

effects obtained with crowding (i.e., an identical stimulus to the target reduces crowding and thereby improves target perception), should generalise to OSM. This would imply that a physically-identical inducer should reduce OSM. Notably, this prediction is specific to an inducer that is physically identical to the target and not one that belongs to the same category but is visually distinct (e.g., same letter in a different case). We tested these predictions in Experiment 1. To anticipate the results, however, across two experiments we found results in stark contrast to these predictions, which actually highlight an interaction between OSM and spatial repetition blindness.

Experiment 1

To test the impact of target-like stimuli on object file formation inferences, here we employed a standard OSM paradigm in which the target was presented surrounded by four dots which disappeared either simultaneously with the target (simultaneous offset condition) or with a delay of 200 ms after the target offset (delayed mask offset condition). The target was one of four letters (A, S, D, or F), that could appear in either upper or lower case and was presented peripherally either to the left or right of fixation. Participants' task was to identify the target letter. On some trials (20%) no foveal inducer stimulus was present, whereas on most trials an inducer was present. When it was present, it had either the same or different identity to the target (i.e., was the same letter), and was either the same or different case (fully crossed with identity). From the perspective of the remote uncrowding effects discovered in visual crowding, we predicted that a foveal inducer would reduce masking when it was identical (with respect to both identity and case) to the target.

Method

Participants. Twenty-four participants (11 female, 13 male) completed the experiment in exchange for pay (AUD\$15). Their mean age was 22.13 years ($SD = 3.0$). All were right handed except one who reported being ambidextrous.

Stimuli & Apparatus. Stimuli were presented on a gamma-corrected cathode-ray tube (CRT) monitor running at a 75Hz refresh rate. Viewing distance was fixed at 44cm with a chinrest. Stimuli were programmed in Matlab using the Psychophysics Toolbox (Brainard, 1997). The background was set to grey (73 cd/m^2). Target stimuli were black letters (size 18 Courier New font): A, S, D, F, a, s, d, f, presented inside four black dots. Each dot subtended 0.4° of visual angle, and they were arranged in a notional square centred on the target letter such that there was 0.8° between the centre of the letter and the centre of each dot. Stimuli could be presented 5.7° to the left or right of fixation. Inducer stimuli were identical to target stimuli, except that they were presented at fixation and were not surrounded by dots.

Procedure. Each trial began with a white fixation dot presented in the centre of the screen for 560 ms, and observers were instructed to fixate on this throughout the trial. Following this, the target letter was presented inside the four black mask dots, either to the left or right of fixation for 40 ms. If an inducer was presented, it appeared and disappeared simultaneously with the target. On the delayed mask offset trials, the four-dots alone were presented for 200 ms after the offset of the target. Then the fixation dot alone was presented until a response was registered. Participants' task was to indicate the identity of the letter presented (i.e., whether it was the letter A, S, D, or F, irrespective of case), and press the corresponding key on the keyboard (i.e., a four-alternative forced choice, 4AFC). Accuracy rather than speed was emphasised. After the response, the screen was blank (grey) for an inter-trial interval of 1600 ms until the next trial commenced (see Figure 1).

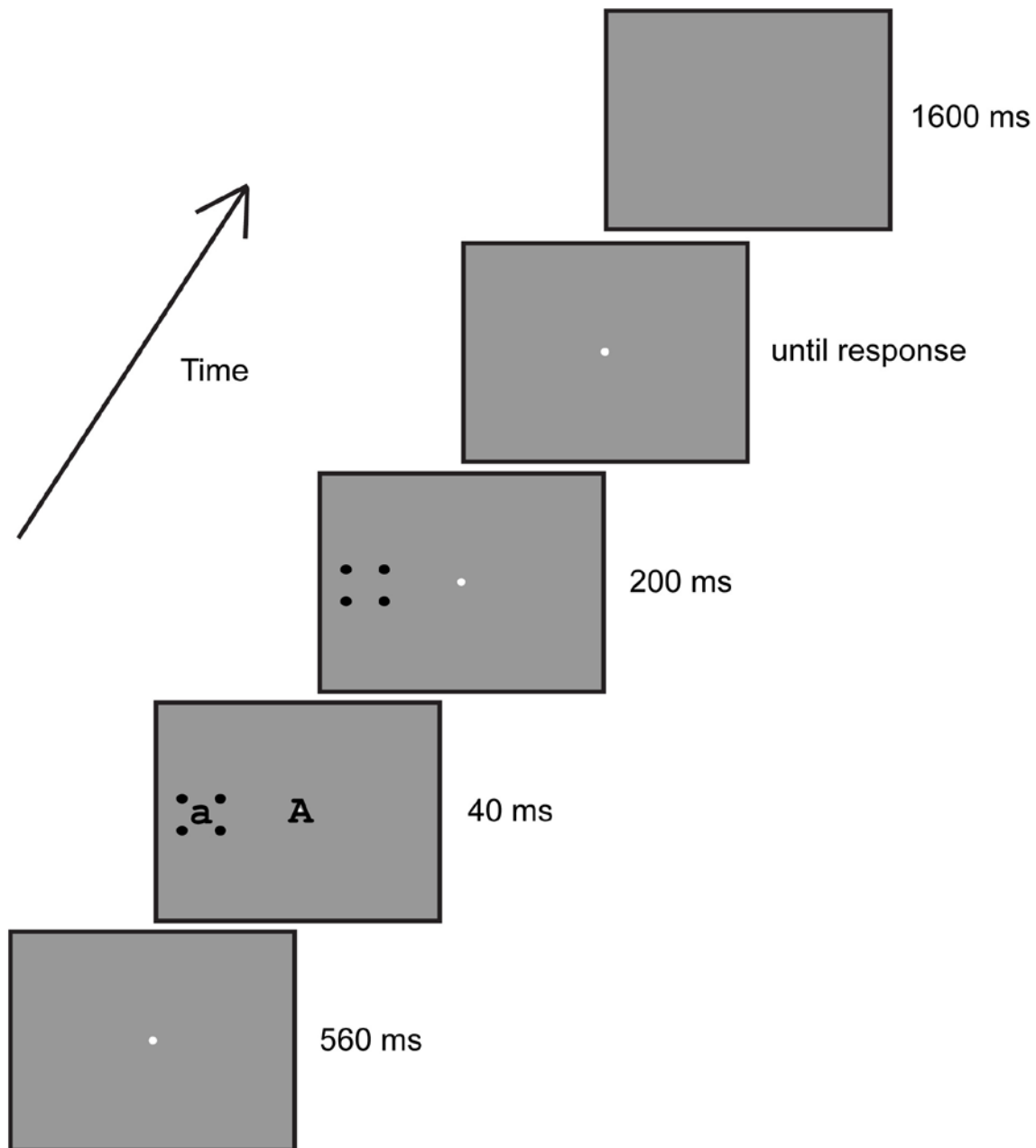


Figure 1. A schematic illustration of a trial in Experiment 1, showing stimuli from a same-identity different-case delayed mask offset trial. On the simultaneous mask offset trials, the trailing four-dot mask would be absent after the target and inducer disappeared.

Each experimental session commenced with a practice block of 12 trials, where the first few trials had slowed presentation times, and the entire practice block presented trial-by-trial feedback on the accuracy of the participants' responses (which was not provided in the experiment proper). Participants were required to reach a minimum of 75% accuracy on the

practice block (with repetition where required) before progressing to the experimental block. Each experimental block consisted of 800 trials, divided equally among the inducer conditions. That is, there were 160 trials for the same-identity same-case inducer condition, 160 trials with a same-identity different-case inducer, 160 trials with a different-identity same-case inducer, 160 trials with a different-identity different-case inducer, and 160 trials for the no inducer control condition. Within each of these conditions, 80 trials were simultaneous target and mask offset trials, and 80 were delayed mask offset trials. The identities of the target and inducer letters were randomly selected from amongst the possible options that met the requirement for that condition (e.g., if the target was the letter 'A' in same-case, different-identity trial, then the inducer had to be one of 'S', 'D', or 'F'). Rest breaks were scheduled every 200 trials, the duration of which was at the discretion of the participant.

Results & Discussion

Responses were excluded from the analysis if the participant pressed a key other than a designated response key ('invalid responses'; average of 0.17% of trials excluded). All participants scored above chance (25%) in the simplest condition, with a simultaneous offset of the masking dots and no inducer, and therefore all datasets were included in the analysis. Percentages of correct responses (to identify the target letter as A, S, D, or F) were then submitted to a 5 (inducer condition) by 2 (mask offset condition) repeated-measures ANOVA (see Table 1 for accuracy values for each condition). All values reported are with the Greenhouse-Geisser correction for sphericity where appropriate. This revealed a significant main effect of inducer condition, $F(2.75, 63.25) = 10.55, p < .001, \eta_p^2 = .314$, and of mask offset condition, $F(1,23) = 65.89, p < .001, \eta_p^2 = .741$. The interaction between inducer condition and mask offset condition was also significant, $F(3.60, 82.68) = 3.09, p = .024, \eta_p^2 = .118$.

Identity:	Same	Same	Different	Different	Control
Case:	Same	Different	Same	Different	Control
Simultaneous	71.3	71.3	74.7	73.2	76.3
mask offset:					
Delayed	57.9	60.6	66.2	66.3	69.1
mask offset:					
Average:	64.6	66.0	70.5	70.0	72.7

Table 1. Percentage correct accuracy for each of the inducer conditions (relationship between target and inducer identity and case) and for each mask offset condition in Experiment 1. The column on the right is the no-inducer control condition. A graphical illustration of the masking magnitudes calculated from these mean accuracies can be seen in Figure 2.

Closer examination of the percentage-correct accuracy (averaged across mask offset condition) revealed that all of the inducer-present conditions had numerically lower accuracy than the no-inducer condition (see Table 1). When accuracies for the inducer-present trials were submitted to a 2 (identity) $\times$ 2 (case) ANOVA, a main effect of identity was found, $F(1, 23) = 11.77$, $p = .002$, $\eta_p^2 = .339$, such that accuracy was greater when the identity of the target and inducer were *different* (70.1%) compared with when they were the same (65.3%). That is, for example, participants were more likely to be correct on a trial with an *A* target and a *D* inducer, than when both elements were the letter *A*. There was no main effect of case ($F < 1$) nor an interaction between identity and case, $F(1, 23) = 1.35$, $p = .257$, $\eta_p^2 = .055$. In other words, an inducer with the same identity as the target impaired accuracy relative to when the inducer and target were different identities.

To better understand these effects, we submitted the data to a 2 (identity) x 2 (case) x 2 (mask offset condition) repeated measures ANOVA. This revealed a significant main effect of identity on target identification accuracy, $F(1, 23) = 11.77$, $p = .002$, $\eta_p^2 = .339$, and a significant main effect of mask offset condition, $F(1, 23) = 60.77$, $p < .001$, $\eta_p^2 = .725$, but no main effect of case ($F < 1$). There was a significant interaction between identity and mask offset, $F(1, 23) = 6.78$, $p = .016$, $\eta_p^2 = .228$, but no interaction between either identity and case ($F < 1$), or between case and mask offset condition, $F(1, 23) = 1.97$, $p = .174$, $\eta_p^2 = .079$. The three-way interaction among identity, case, and mask offset condition was also not significant ($F < 1$).

Closer examination of the masking magnitudes (the difference in accuracy between simultaneous and delayed mask offset conditions) suggests that masking appeared to scale depending on the relationship between the inducer and the target: masking was greatest when the target and inducer were identical with respect to both case and identity, and weakest when the target and inducer were of different identities and cases (see Figure 2). Indeed, masking was greater when the target and inducer were the same identity (12.0%) compared with when they were different (7.7%). In fact, the masking magnitude when the target and inducer shared the same identity was significantly greater than the no-inducer control condition, $t(23) = 3.15$, $p = .005$, whereas the masking magnitude when the target and inducer were different identities was indistinguishable from the no-inducer control condition ($t < 1$). This means that the visual system was more likely to infer the presence of only a single object at the target location continuing over time (i.e., the mask), and thus fail to consciously represent the target when the inducer and target shared the same identity, whereas their case did not appear to affect this inference.

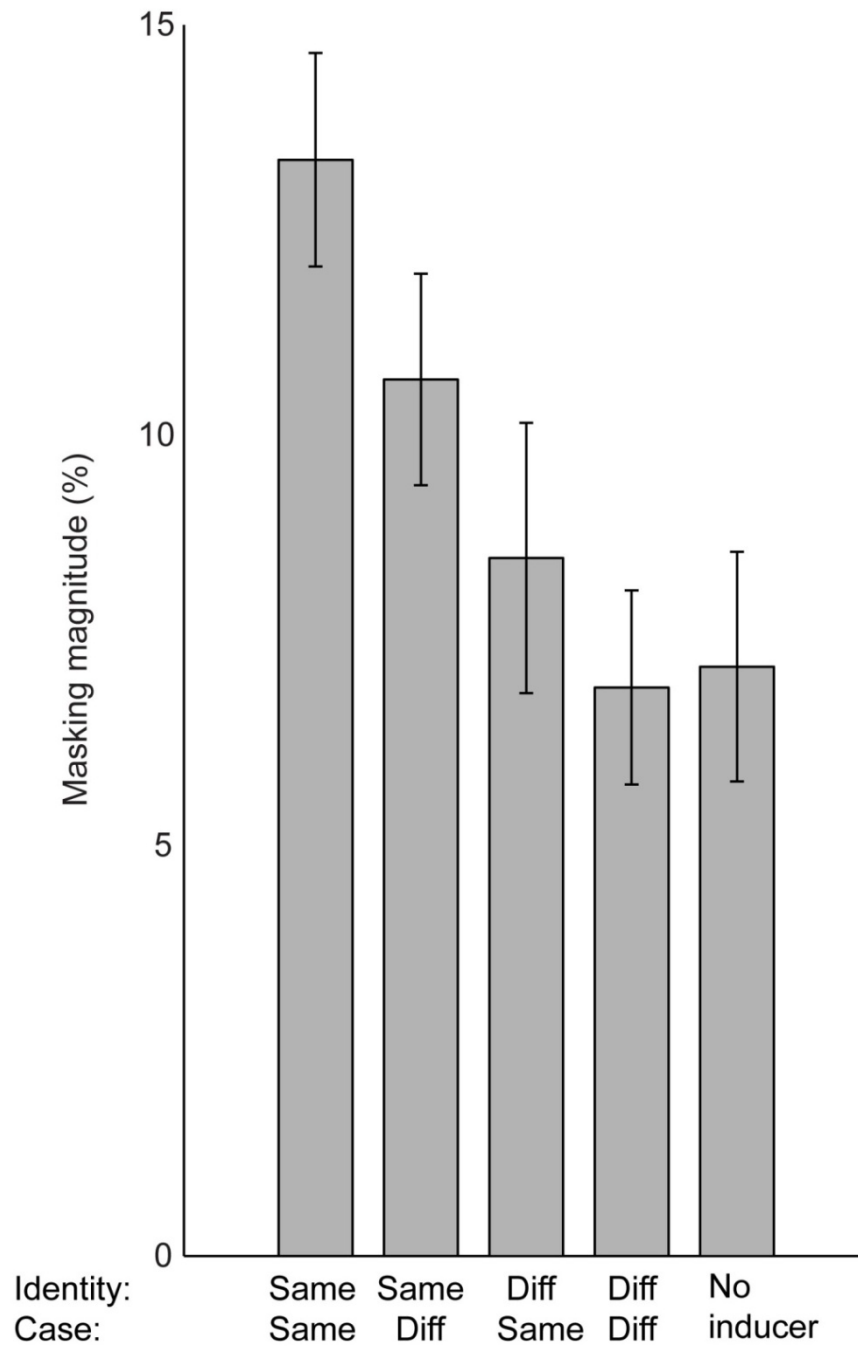


Figure 2. Masking magnitude (the difference in accuracy between simultaneous and delayed mask offset conditions) for each of the inducer condition in Experiment 1. Error bars depict standard error of the mean corrected for within-subjects designs (Cousineau, 2005).

Next, we sought to check that floor and ceiling effects were not constraining the data and thus distorting the pattern of results across conditions. We therefore excluded the data from participants whose performance exceeded 90% or fell below 35% in any condition (7

exclusions), and then repeated the analysis. This revealed that the pattern of results was unchanged. That is, there was still a significant main effect of identity, $F(1, 16) = 8.04$, $p = .012$, $\eta_p^2 = .335$, a main effect of mask offset condition, $F(1,16) = 35.26$, $p < .001$, $\eta_p^2 = .688$, and no main effect of case ($F < 1$). Most critically, there was still a significant interaction between identity and mask offset condition, $F(1, 16) = 4.66$, $p = .046$, $\eta_p^2 = .266$. There was neither an interaction between identity and case ($F < 1$) nor between case and mask offset condition ($F < 1$). This suggests that the results are robust and not reflective of floor or ceiling effects.

These results are in striking contrast to the predictions from visual crowding, both in terms of the qualitative direction of the effect, and the level of stimulus processing at which the effect of the inducer manifested. Specifically, in crowding, a target-like stimulus can decrease crowding, but only when the inducer and target are physically identical. In contrast, here in OSM, a target-like stimulus increased masking, and this was true when the inducer and target shared the same identity, even if they were physically different (i.e., different cases). Object repetitions clearly have markedly different effects on crowding versus OSM, suggesting at least some separation in their mechanisms.

In fact, the reduced identification accuracy when the inducer and target shared the same identity strongly resembles a less well-known failure of visual awareness called *spatial repetition blindness*. RB is the impaired perception of a repeated item compared to a non-repeated item under otherwise identical presentation parameters (Kanwisher, 1987; Kanwisher & Potter, 1989, 1990). Typically RB has been demonstrated using rapid serial visual presentation (RSVP) streams, in which items are represented at the rate of about 10 images per second, and participants are likely to fail to perceive or report the repeated items. RB, however, is not limited to sequential presentation of the items, because it is also found when repeated items are presented simultaneously in different spatial locations, known as ‘spatial RB’ (Harris,

Wong, & Andrews, 2015; Kanwisher, 1991; Kanwisher & Potter, 1989; Luo & Caramazza, 1996). RB has been documented with a range of stimuli, including letters, words, and pictures (Bavelier, 1994; Egeth & Santee, 1981; Kanwisher, 1991; Luo & Caramazza, 1996; Marohn & Hochhaus, 1988). RB is not limited to physically-identical stimuli, but also occurs for stimuli that belong to the same category but are visually dissimilar (e.g., ‘A’ and ‘a’, or the word ‘cat’ and a picture of a cat) (Bavelier, 1994; Egeth & Santee, 1981; Marohn & Hochhaus, 1988). It is a strikingly robust phenomenon, occurring even when the omission of the repeated item violates overarching structures like the semantic sense of a sentence (Kanwisher & Potter, 1989, 1990).

The dominant explanation for RB is a failure to individuate *tokens* (i.e., object files) for the same *type* (Chun & Cavanagh, 1997; Goldfarb & Treisman, 2011; Kanwisher, 1987, 1991; Kanwisher, Driver, & Machado, 1995; Kanwisher & Potter, 1989). That is, when the system registers multiple activations of the same type (e.g., a given word) in close succession, it attributes this to a single object, and thus fails to recognise the unique instances (tokens) that led to this activation. For example, the system recognises that the word ‘cat’ was presented, but fails to individuate the two separate presentations. While other explanations for RB have been suggested, such as type-node refractory period (Luo & Caramazza, 1996), and response-level and memory-based explanations (Fagot & Pashler, 1995), these have since been refuted (Chun & Cavanagh, 1997; Kanwisher, Kim, & Wickens, 1996), and thus the type-token individuation failure remains the prevailing explanation.

One of the key demonstrations for the role of object files in RB, and thus support for the explanation attributing the failure to individuate tokens of the same type, is that RB interacts with apparent motion. Since apparent motion is the percept that links two spatiotemporal events to the same object token (Anstis, 1980), apparent motion is also a failure to individuate multiple tokens. Unlike RB, however, it does not depend on the tokens belonging to the same type, as

apparent motion can even occur when the two instances are featurally distinct, although it is enhanced by featural similarity (Hein & Moore, 2012). Chun and Cavanagh (1997) showed that repeated items occurring within the same stream of apparent motion led to increased RB compared with items occurring in different streams. In other words, when the physical parameters encouraged the inference of a single token continuing across time, this exacerbated the failure to form to separate tokens when they belonged to the same type.

Here in Experiment 1, we found that accuracy was reduced when the inducer and target shared the same identity. This in fact replicates the basic phenomenon of spatial repetition blindness. Most strikingly, however, we also found that masking magnitude and spatial repetition blindness *interacted*. That is, masking magnitude, rather than just accuracy per se, was affected by the relationship between the inducer and the target. This suggests that OSM and RB share some common underlying mechanisms, or at least have distinct mechanisms that share some common properties.

One might argue that a key difference between the experimental parameters used here and those employed in standard RB is that in the latter, participants are required to report two items, whereas here they only had to report one (the target, not the inducer). However, spatial RB is not dependent on these dual-item report requirements, and is robustly observed even when participants only have to report a single item (e.g., Luo & Caramazza, 1995). The number of items to report, therefore, does not distinguish the present experiment from RB.

If we consider this relation in terms of the process of the object-updating principles that are commonly used to explain the presence of masking (Goodhew, Pratt, et al., 2013; Lleras & Moore, 2003; Moore & Lleras, 2005; Pilling & Gellatly, 2010), OSM occurs when the visual system fails to create separate object files for the target and mask, and ultimately the initial representation of the target is updated to reflect the mask. According to this object-updating

framework, therefore, the present results indicate that perceiving a stimulus that shares the same identity as the target makes it even more likely that the visual system will favour an *updating* solution when confronted with the presentation parameters of OSM. If we consider the present findings in relation to the object substitution account, according to which the initial representation of the target is discarded in favour of the mask representation (Di Lollo, 2010; Di Lollo et al., 2000), then the present results imply that perceiving a stimulus that shares the same identity as the target takes makes it even more likely that the visual system will favour a *substitution* solution when confronted with dynamic input. Recasting object-updating and object-updating in terms of their broader, functionally-isomorphic mechanisms (i.e., the failure to maintain an enduring object token that contains the target), the present results tell us that perceiving a stimulus that shares the same identity as the target decreases the probability that the visual system will maintain an enduring representation for the target when confronted with the dynamic input that constitute the presentation parameters of OSM.

Before we can interpret this result further, however, we need to establish whether this effect of simultaneous repetition of target identity on masking magnitude reflected a true change in perceptual sensitivity to the target stimulus, or merely a shift in response criterion. There are two main motivations for this. Firstly, previous research has shown that RB entails a true change in perceptual sensitivity, rather than just a change in response criterion (Kanwisher et al., 1996). Thus, if the results are related to mechanisms implicating RB, they should be related to changes in perceptual sensitivity. Secondly, it is possible that if participants adopted a response criterion to make a response indicating that the target was an identity different to that of the inducer on that trial, that this would have systematically reduced their accuracy on the same-identity trials relative to the different-identity trials. In order test this, in Experiment 2 we modified the task such that participants were making a two-alternative forced choice

(2AFC) on each trial and applied a signal detection theory (SDT) framework (Macmillan & Creelman, 2005) to analyse the results.

Experiment 2

The results of Experiment 1 revealed an interaction between spatial RB and OSM, such that the simultaneous repetition of a stimulus that shared the same identity as the target in another spatial location increased masking magnitude. But did this reflect a true change in perceptual sensitivity, or a strategic response bias that participants adopted? Here we used a 2AFC and applied the SDT framework to disentangle changes in perceptual sensitivity (d') from response criterion (c) in this paradigm.

We employed both 2AFC identification and detection tasks. In the identification task, a letter was always presented, and participants judged whether the given target letter 'A' was present or not. When it was not, another non-target letter ('B') was present. In the detection task, participants were detecting whether a letter stimulus was present at all, regardless of its identity. When it was not, on half of trials, no letter target was present and only the four-dot mask was shown. We did this because evidence suggests that within-category identification (e.g., daffodil versus daisy, letter 'A' versus letter 'B') is much slower and involves deeper processing compared to either detection (noticing that there was an object present) or categorisation (knowing that a flower was present, that a letter was present) (Grill-Spector & Kanwisher, 2005). Performance on the detection task, therefore, will tell us whether processing the identity of the target is necessary to obtain the interaction between OSM and spatial repetition blindness.

Method

Participants. Twenty-seven (17 female, 10 male) participants completed the experiment in exchange for pay (AUD\$15). Their mean age was 20.6 years ($SD = 3.2$). Four reported being left-handed, two reported being ambidextrous, and the remainder reporting being right handed.

Stimuli & Apparatus. Stimuli and apparatus were identical to Experiment 1, with the following exceptions. The inducer and target letters were now the letter ‘A’ or ‘B’ (both uppercase and lowercase, fully crossed). For the identification task, participants were told that their target was the letter ‘A’, and were instructed to press the ‘Z’ key if the target was present, and ‘/?’ if it was absent. This target was present on 50% of trials, and when it was absent, the letter ‘B’ was present instead. For the detection task, a target letter was either present or absent, and participants were instructed to press the ‘Z’ key if the target was present, and ‘/?’ if it was absent. A target was present on 50% of trials, and when it was, its identity was equally likely to be ‘A’ or ‘B’.

Procedure. There were two experimental blocks (identification versus detection), each consisting of 400 trials, divided equally among the inducer conditions. That is, there were 80 trials for same-identity same-case inducer condition, 80 trials for same-identity different-case inducer, 80 trials for different-identity same-case inducer, 80 trials for different-identity different-case inducer, and 80 trials for the no inducer control condition. Within each of these conditions, 40 trials were simultaneous target and mask offset trials, and 40 were delayed mask offset trials. The procedure was identical to Experiment 1 in all other respects, except that for the detection task, target exposure duration was reduced to 26 ms and its contrast was reduced from black (100% contrast) to grey (20% contrast).

Results & Discussion

As per Experiment 1, invalid responses were excluded from analysis. One participant made invalid responses on almost 25% of trials, and was therefore excluded from further analysis. The average percentage of invalid responses excluded from the remaining participants was very low ($<0.5\%$).

Within the SDT framework, a ‘hit’ constitutes the correct identification of the target on trials when it was present. Since in the identification task the target was the letter A (and therefore ‘B’ was the nontarget), when ‘A’ was present inside the four dots and participants selected the ‘present’ response, this was classified as a hit. A ‘false alarm’ reflects the incorrect response of target presence when the target is in fact absent. Here, therefore, when the letter ‘B’ was presented inside the four dots but participants responded that the target was present, this was classified as a false alarm. For the detection task, the definitions of hit, miss, false alarm, and correct rejection map onto the standard definitions of these labels. We then calculated measures of sensitivity (d') and criterion (c) (Macmillan & Creelman, 2005), with d' as $z(\text{hit-rate}) - z(\text{false-alarms})$, and the criterion as $-(z(\text{hit-rate}) + z(\text{false-alarms}))/2$. The d' and c values for the identification and detection conditions can be seen in Tables 2 and 3, respectively.

As in Experiment 1, we first examined each participant’s performance in the simplest condition: the no-inducer, simultaneous mask offset condition. We reasoned that if participants were unable to reach a sufficient level of performance here, then their performance would be subject to floor effects and thus insensitive to variation due to condition. Two participants demonstrated poor performance ($d' < 1$) in the identification task, and were therefore excluded from further analysis. Five participants showed poor performance in the detection task and were similarly excluded. Sensitivity (d') to detect the identity/presence of the target letter for the remaining datasets was analysed (see Table 2 and 4 for mean values) separately for each of the tasks (identification and detection).

Identification task. For the identification task, a 5 (inducer condition) x 2 (mask offset condition) repeated-measures ANOVA revealed a significant main effect of inducer condition on sensitivity to the target, $F(2.26, 51.87) = 12.51, p < .001, \eta_p^2 = .352$, a significant main effect of mask offset condition, $F(1, 23) = 22.67, p < .001, \eta_p^2 = .496$, as well as a significant interaction between inducer condition and mask offset condition, $F(3.36, 77.33) = 3.08, p = .028, \eta_p^2 = .118$.

Identity:	Same	Same	Different	Different	Control
Case:	Same	Different	Same	Different	Control
Simultaneous	1.94	1.91	2.54	2.34	2.44
mask offset:					
Delayed	1.24	1.10	2.17	2.23	1.85
mask offset:					

Table 2. Sensitivity (d') for each of the inducer conditions and for each mask offset condition for the identification task in Experiment 2. A graphical illustration of the masking magnitudes calculated from these values can be seen in Figure 3.

As can be seen in Figure 3, in the identification task, masking was greater when the inducer and target shared a common identity than when they did not, akin to the pattern observed in Experiment 1. In order to fully understand this interaction, we submitted the inducer-present trials to a 2 (identity) x 2 (case) x 2 (mask offset condition) ANOVA, which revealed a significant main effect of inducer identity on sensitivity to the target, $F(1, 23) = 19.21, p < .001, \eta_p^2 = .455$, such that sensitivity was lower when the target and inducer were the same

identity (1.55), compared with when they were different (2.32). This confirms the presence of spatial RB. There was also a main effect of mask offset condition, $F(1, 23) = 17.91$, $p < .001$, $\eta_p^2 = .438$, such that sensitivity was reduced in the delayed mask offset condition (1.68) relative to the simultaneous mask offset condition (2.19). This confirms the presence of OSM. Crucially, there was also an interaction between inducer identity and mask offset condition, $F(1, 23) = 7.44$, $p = .012$, $\eta_p^2 = .244$. None of the other main effects or interactions were significant ($ps \geq .176$ & $\eta_p^2 \leq .078$). The interaction between inducer identity and mask offset condition reflects the fact that masking magnitude was greater when the inducer and target shared the same identity (0.76) compared with when they were different identities (0.24). This means that when the inducer and target belonged to the same type, the visual system was less likely to maintain an enduring object token for the target. In other words, there was an interaction between spatial RB and OSM. This is consistent with the results of Experiment 1, and demonstrates that the same effect occurs even when pure perceptual sensitivity, uncontaminated by response bias, is the dependent measure.

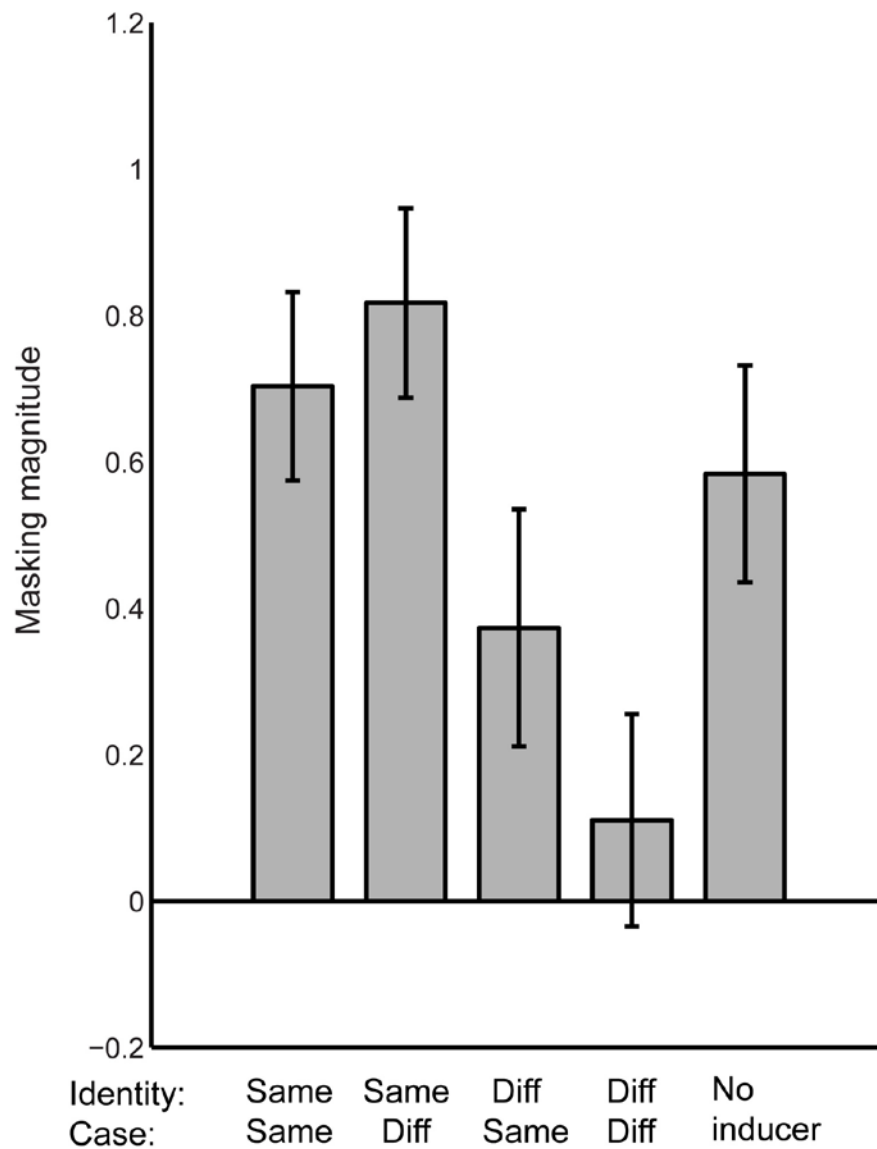


Figure 3. Masking magnitude (difference in sensitivity (d') between simultaneous and delayed mask offset conditions) for each of the inducer condition for the identification block in Experiment 2. Error bars represent standard error of the mean corrected for within-subjects designs (Cousineau, 2005).

One interesting aspect of the results that does differ from Experiment 1 is that rather than the same identity inducer increasing masking relative to the no-inducer baseline, here this was not the case, and while the same-identity condition showed numerically greater masking (0.76) than the no-inducer baseline (0.58), and the different-identity condition showed (0.24) showed

less than baseline, neither of these comparisons were significant ($ps = .358$ and $.145$ respectively). Instead, it was the substantial difference in masking magnitude between the same-identity and different-identity inducer conditions that was reliable. The conclusion is unchanged, however, because in spatial RB, the appropriate comparison is between the repeated and non-repeated item (rather than between a repeated item and no item). The equivalent comparison here is between the same-identity and different-identity inducers (rather than the same-identity inducer compared with no inducer), and this comparison yielded a highly reliable impact of identity on masking magnitude. Altogether then, the results demonstrate an interaction between spatial RB and OSM on perceptual sensitivity to the target.

One possibility is that is that here, despite the fact that the main effect and interaction implicating case were not significant, perceptual similarity did in fact impact participants' perception of the target. This could have been obscured by the fact that the *different identity* condition averages across two different target-inducer stimulus pairs: "A" and "B" versus "a" and "b". The latter are more perceptually similar than the former, despite both representing different-identity pairs. To assess whether this 'same versus different case' definition of perceptual similarity had an impact on the results, we divided all of the trials according whether the target and inducer pair were upper or lower case on that trial, and submitted the data to a 2 (perceptual similarity) x 2 (identity) x 2 (case) x 2 (mask offset condition) repeated measures ANOVA. This revealed a significant main effect of perceptual similarity, $F(1, 23) = 4.66$, $p = .042$, $\eta_p^2 = .169$, such that target detection accuracy was significantly higher when the target and inducer were both uppercase (2.42) compared with when they were lowercase (1.97). (Note that here we will not repeat reporting of main effects/interactions that are redundant with previous analyses of this data, e.g. main effect of identity). Perceptual similarity also interacted with mask offset condition, $F(1, 23) = 5.29$, $p = .031$, $\eta_p^2 = .187$, such that average masking magnitude was greater with uppercase (lower similarity) inducer-target pairs (0.91) compared

with lowercase (higher similarity) inducer-target pairs (0.33). That is, the more perceptual similar to the inducer to the target, the weaker the masking magnitude. Crucially, however, however, there was neither a three-way interaction among perceptual similarity, identity, and mask offset condition ($F < 1$), nor a four-way among perceptual similarity, identity, case, and mask offset condition, $F(1, 23) = 1.43$, $p = .174$, $\eta_p^2 = .079$, meaning that the (still significant, of course) two-way interaction between identity and mask offset condition was unchanged by considering perceptual similarity as a factor. This analysis suggests that while perceptual similarity between target and the inducer appears to be an important factor in its own right for both target perception and masking, this does not alter the relationship between categorical identity and both target perception and masking magnitude. In other words, the category of the letter is still an important dimension and impacts on perception and masking magnitude, even when perceptual similarity is accounted for.

Next, we once again sought to establish that this pattern of results was not a product of floor or ceiling effects. To this end, we excluded datasets where performance in any condition fell below $d' = 0$, or where performance in any condition exceeded $d' = 5$. (While chance level performance is 0, and thus we could have adopted a more stringent exclusion criterion, this would have led to further exclusions and we were already losing a substantial chunk of the data and thus power. Indeed, critically, if we do attempt to apply more stringent criteria, such as $d' = 0.5$ and 4.5 as minimum and maximum cut-offs, then we lose all effects except that of mask offset condition, including identity, and thus cannot make any inferences without the presence of spatial RB). This led to four exclusions, which revealed an interesting result, whereby the main effect of identity was still significant, $F(1, 19) = 14.39$, $p = .001$, $\eta_p^2 = .431$, as was the main effect of mask offset condition, $F(1, 19) = 16.94$, $p = .001$, $\eta_p^2 = .471$, whereas there was no main effect of case ($F < 1$). There was neither an interaction between identity and case ($F < 1$) nor between case and mask offset condition ($F < 1$). The interaction between identity and mask

duration was now of borderline significance, $F(1, 19) = 4.37$, $p = .050$, $\eta_p^2 = .187$. There was now, however, a significant three-way interaction among identity, case, and mask offset condition, $F(1, 19) = 4.87$, $p = .040$, $\eta_p^2 = .204$. That is, for the first time, the case of the letter appeared to impact participants' target detection performance, in addition to its identity. The average masking magnitudes followed the same patterns as those shown in Figure 3, simply appearing more pronounced (e.g., greater advantage for the different-case relative to the same-case). Altogether then, this suggests that case may also play an interactive role with identity in determining masking magnitude. Most crucially, however, identity still did have an impact on masking, even when here more steps were taken to even further mitigate the likelihood of floor or ceiling effects.

The response criterion values for each condition in the identification task are shown in Table 3. These were submitted to a 5 (inducer condition) by 2 (mask offset condition) repeated-measures ANOVA, which revealed a main effect of mask offset condition, $F(1,23) = 4.34$, $p = .049$, $\eta_p^2 = .159$, but no effect of either inducer condition ($p = .119$, $\eta_p^2 = .079$) or the interaction between inducer condition and mask offset condition ($p = .609$, $\eta_p^2 = .028$). Furthermore, the inducer-present trials were submitted to a 2 (target identity) x 2 (case) x 2 (mask duration) repeated-measures ANOVA. None of the main effects or interactions from this analysis were significant ($ps \geq .070$ & $\eta_p^2s \leq .136$). Thus, the impact of spatial repetition of the same type as the target impacts sensitivity to the target, rather than response criterion.

Identity:	Same	Same	Different	Different	Control
------------------	------	------	-----------	-----------	---------

Case:	Same	Different	Same	Different	Control
Simultaneous	.23	.14	.16	.12	.29
mask offset:					
Delayed	.11	-.03	-.02	.11	.13
mask offset:					

Table 3. Criterion (c) for each of the inducer conditions and for each mask offset condition for the identification task in Experiment 2. Note that a higher positive value indicates a more conservative criterion with respect to target presence, and a value of zero indicates an absence of a systematic response bias.

Detection task. The first three participants who completed this task were run on a version where target exposure time was 40 ms and the target was black. These participants showed performance levels indicative of ceiling effects, including d 's all > 2 and has high as 4.6 even in delayed offset conditions. We therefore then reduced target exposure time and target contrast, and the following results exclude those three datasets from the analysis. Sensitivity (d') to detect the presence of the target letter were firstly analysed via a 5 (inducer condition) $\times$ 2 (mask offset condition) repeated measures ANOVA, which indicated a significant main effect of inducer condition, $F(2.15, 36.51) = 3.43$, $p = .040$, $\eta_p^2 = .168$, and a main effect of mask offset condition, $F(1, 17) = 13.22$, $p = .002$, $\eta_p^2 = .437$, but no interaction ($F < 1$). Secondly, a 2 (identity) $\times$ 2 (case) $\times$ 2 (mask offset condition) repeated measures ANOVA on the inducer-present trials showed a significant main effect of mask offset condition, $F(1, 17) = 12.30$, $p = .003$, $\eta_p^2 = .420$, indicating OSM, whereas none of the other main effects or interactions approached significance ($F_s < 1$, $p_s \geq .373$, & $\eta_p^2 \leq .047$). This indicates that when the task is merely to detect target presence and no identity-level processing is required, the interaction between spatial RB and OSM is eliminated.

Identity:	Same	Same	Different	Different	Control
Case:	Same	Different	Same	Different	Control
Simultaneous	2.08	2.02	2.18	2.04	1.75
mask offset:					
Delayed	1.62	1.63	1.64	1.70	1.20
mask offset:					

Table 4. Sensitivity (d') for each of the inducer conditions and for each mask offset condition for the detection task in Experiment 2. A graphical illustration of the masking magnitudes calculated from these values can be seen in Figure 4.

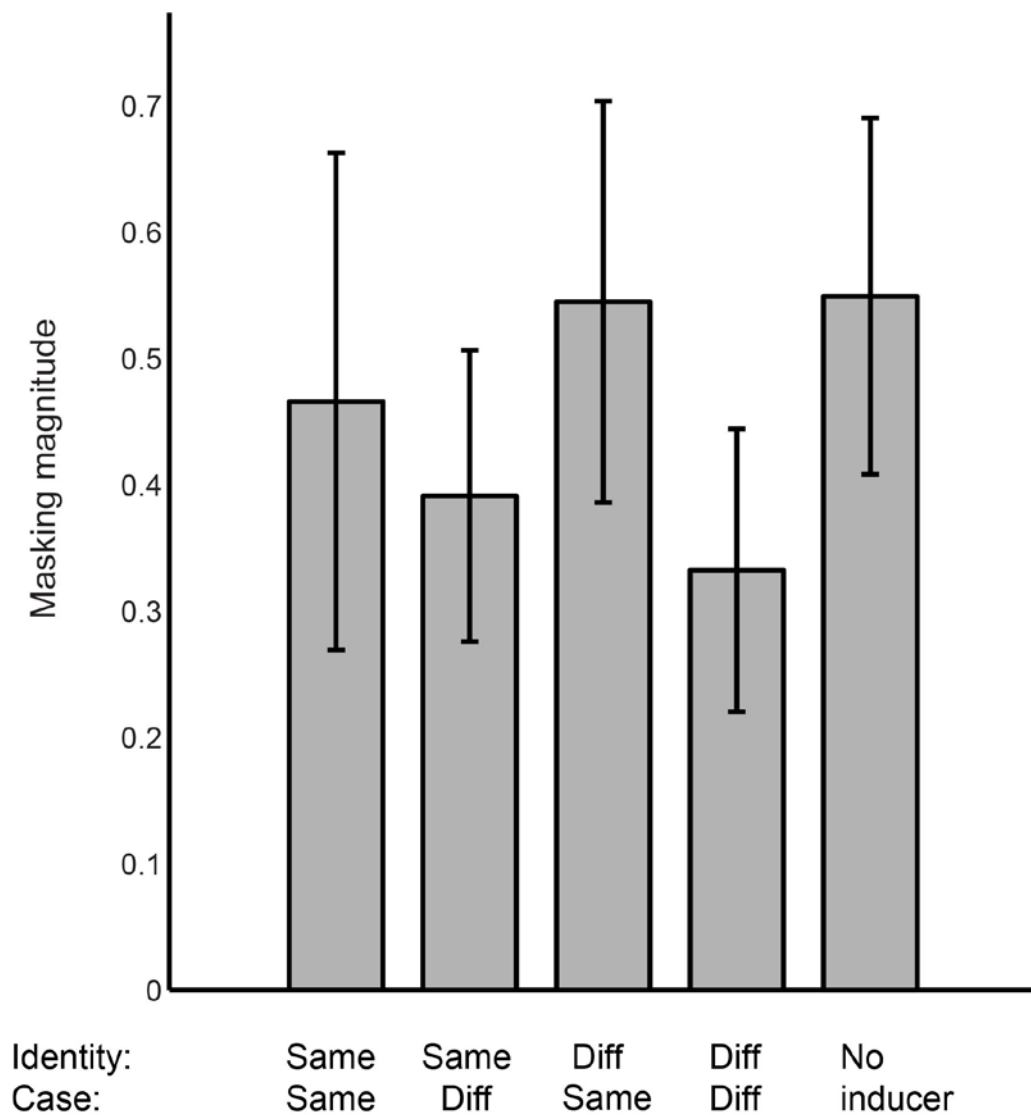


Figure 4. Masking magnitude (difference in sensitivity (d') between simultaneous and delayed mask offset conditions) for each of the inducer condition for the detection block in Experiment 2. Error bars represent standard error of the mean corrected for within-subjects designs (Cousineau, 2005).

We also submitted the criterion values from the detection task to a 5 (inducer condition) by 2 (mask offset condition) repeated-measures ANOVA, which showed a non-significant main effect of inducer condition, $F(1.4, 23.55) = 3.16$, $p = .076$, $\eta_p^2 = .157$, a main effect of mask

offset condition, $F(1, 17) = 28.67$, $p < .001$, $\eta_p^2 = .628$, and a significant interaction between inducer condition and mask offset condition, $F(2.70, 45.96) = 3.91$, $p = .017$, $\eta_p^2 = .187$. In order to better understand this interaction, we then submitted the inducer-present trials to a 2 (identity) $\times$ 2 (case) $\times$ 2 (mask offset condition) repeated-measures ANOVA, which revealed again a significant main effect of mask offset condition, $F(1, 17) = 21.87$, $p < .001$, $\eta_p^2 = .563$, such that criterion was more conservative on the delayed mask offset trials relative to the simultaneous mask offset trials. No other main effects or interactions were significant ($ps \geq .127$, η_p^2 s $\leq .131$). Thus, while it appears the presence of the inducer may have impacted the effect of mask offset condition on criterion, this was not dependent on any systematic relationship between the inducer and the target.

Identity:	Same	Same	Different	Different	Control
Case:	Same	Different	Same	Different	Control
Simultaneous	.67	.79	.75	.73	.86
mask offset:					
Delayed	1.18	1.12	1.08	1.05	1.54
mask offset:					

Table 5. Criterion (c) for each of the inducer conditions and for each mask offset condition for the detection task in Experiment 2. Note that a higher positive value indicates a more conservative criterion with respect to target presence, and a value of zero indicates an absence of a systematic response bias.

General Discussion

Here, across two experiments, we found that repetition of a stimulus that shared the same identity as the target increased OSM magnitude relative to the presentation of a stimulus that did not share the same identity as the target. While perceptual similarity appeared to also play a role in target perception and masking in Experiment 2, crucially, this was always orthogonal to the impact of identity. Spatial repetition blindness has been demonstrated previously, such that the repetition of a stimulus of the same type impairs perception (Harris et al., 2015; Kanwisher, 1991; Luo & Caramazza, 1996). Here, the conditions that created spatial RB exacerbated masking, demonstrating an interaction between RB and OSM. This is consistent with the fact that these two failures of visual awareness are theorised to have similar underlying mechanisms.

Gellatly, Pilling, Cole, and Skarratt (2006) were the first to note the conceptual similarity between RB and OSM. Their analysis focussed on stimulus similarity effects, specifically the finding in previous research that RB is not modulated by repetition along unattended feature dimensions, in conjunction with their empirical finding that similarity between the target and mask along unreported dimensions did not modulate OSM. Gellatly et al. did not, however, demonstrate a direct interaction between repetition blindness and OSM, and moreover, other studies have shown that similarity along unreported dimension can modulate masking (Goodhew, Edwards, Boal, et al., 2015; Luiga & Bachmann, 2008; Moore & Lleras, 2005). Thus, up until now, the link between RB and OSM has not been compelling. The present results, however, show a direct interaction and thus support the notion that both OSM and RB reflect a failure to maintain enduring object tokens for two distinct objects (a process we dub a failure of *object individuation*), thus resulting in the perception of just a single object. This does not imply that they are identical phenomena, but what they have in common is they demonstrate the brain's use of a simple but powerful heuristic in the face of dynamic and

interrupted visual input: the probability that two featurally or category-level related events reflects two distinct objects is lower than the probability for two events that are not related.

It is notable that the effect of target-like repetitions on OSM qualitatively diverged from the effect of this manipulation in visual crowding. In crowding, a target-like stimulus *releases* the target from crowding, but only if physically identical to the target (Geiger & Lettvin, 1986; Sayim et al., 2014). One possible interpretation of this difference is that the locus of visual crowding is at a lower level in the visual system, such that the image-level representation of the target is degraded. In OSM, in contrast, the image-level representation appears to remain largely intact, even to the extent that it can activate semantic representations when masking is effective (Goodhew, Visser, Lipp, & Dux, 2011b). Instead, in OSM it is the subsequent inferences about how to interpret this information that lead to the suppression of visual awareness of the target (Goodhew, Pratt, et al., 2013). As a result, the presentation of a stimulus identical to the target can boost the target signal and thus overcome visual crowding, whereas in OSM this same element will create confusion about the possible source of category-level activations and thus increase masking.

Our interpretation that visual crowding should occur at a lower level than OSM is consistent with prior work reported by Chakravarthi and Cavanagh (2009). In a crowding paradigm, Chakravarthi and Cavanagh masked the flankers of a crowded object with three different types of masks: a noise mask, a metacontrast mask, or a four-dot (OSM) mask. They reasoned that if a given mask were to suppress the flankers, then it should reduce crowding and thus improve target perception. This uncrowding effect would then suggest that the form of masking capable of doing this had occurred at an earlier locus in the system than visual crowding. They found that both noise and metacontrast masks applied to the flankers reduced crowding, whereas object-substitution masks did not. They concluded that the suppression characteristic of object substitution occurs higher up in system than visual crowding (Chakravarthi & Cavanagh,

2009). Others have however found that crowding and OSM do appear to interact in some circumstances (Camp et al., 2015). We suggest that while crowding and OSM have different loci and reflect dissociable underlying mechanisms, crowding can impact on the object-file formation mechanisms underlying OSM. That is, if crowding weakens the quality of the target representation, then this likely feeds into the object-formation and consolidation process, encouraging the inference that the target does not warrant a distinct object file from the mask. Altogether then, we suggest that visual crowding reflects an earlier locus where the representation of target quality is degraded, whereas OSM reflects more about inferences of assignment of object files to visual events.

Our results also have important implications for our understanding of OSM in its own right. That is, a key finding to emerge from the present study is that categorical information about the target directly impacts object file consolidation, and therefore whether or not the target object is consciously perceived. There is a growing body of evidence to indicate advanced, high-level processing of the target, even when OSM is effective. It has been shown in OSM that the semantic content (i.e., meaning) of a target word influences response efficiency to another task, even though masking is effective such that participants cannot accurately identify whether the target is a word or a random string of letters, or even detect its presence (Goodhew et al., 2011b). While not direct evidence for semantic processing, there is also other evidence for processing of the target despite effective masking, such as the finding that targets masked via OSM still influence motor actions (Binsted, Brownwell, Vorontsova, Heath, & Saucier, 2007), and with prolonged mask exposure, the target can be ‘recovered’ (i.e., target identification improves relative to intermediate mask durations), suggesting that effective masking does not necessitate a total loss of target-related information (Goodhew, Dux, Lipp, & Visser, 2012; Goodhew, Visser, Lipp, & Dux, 2011a). In other words, there appears to be high-level processing, including semantic processing despite effective OSM. But here, rather

than category-level processing of the target merely affecting response efficiency to a secondary task, it actively impacted whether the target was perceived. That is, the categorical relationship between the inducer and target (e.g., ‘a’ and ‘A’) was a strong determinant of masking. In order for this *relationship* to have an impact, it necessitates that the target itself was processed to a categorical level, even though masking was effective. This is compelling evidence that OSM allows for high-level processing, and only interferes with late-stage processes involved in conscious perception of the target.

It should be noted that in the literature that are documented instances of failures to find unambiguous evidence of semantic and category-level processing in OSM (Z. Chen & Treisman, 2009; Reiss & Hoffman, 2006, 2007). There are, however, a number of methodological pitfalls in these studies, and these notwithstanding, the absence of evidence for implicit semantic and categorical perception is not complete. For example, in Chen and Treisman’s (2009) behavioural priming paradigm, the key category boundaries were between *vowels* versus *consonants* – as opposed a more naturalistic categorisation such as the letter category to which they belong. That is, intuitively, humans much more easily and readily classify the grapheme that constitutes the second letter of the alphabet as *b* than as belonging the category of *consonant*. Indeed, children typically learn the individual letter of the alphabet (*a, b, c...*) well before they learn the vowel versus consonant distinction (*vowel, consonant, consonant*). Despite choosing to categorise stimuli along somewhat arbitrary lines, that is, not carving language at its joints, Chen and Treisman still found some behavioural evidence for implicit semantic perception in OSM, which they dismissed as perceptual in nature.

In a similar vein, Reiss and Hoffman (2006) examined the N400 event-related potential (ERP) component, widely considered to be the electrophysiological signature of semantic processing (Kutas & Hillyard, 1980), in response to both masked and unmasked targets. Measuring this ERP component requires comparison of *congruent* and *incongruent* semantic

conditions. Strikingly, Reiss and Hoffman found reduced OSM magnitude on the congruent compared with the incongruent trials – which could constitute direct evidence for semantic processing of the masked target. However, the way that the study was designed did not allow this behavioural effect to be disentangled from a guessing bias confound, due to the presentation of a context word at the beginning of the trial (for a full explanation, see review paper, Goodhew, Pratt, et al. 2013).

Furthermore, Reiss and Hoffman (2007) examined the N170 ERP component, the purportedly face-specific component in response to pictorial face versus house targets amongst house distractor images. These authors found no ERP evidence that the categorical distinction between faces and houses, which was registered when the target was unmasked, was present when it was masked (the authors collapsed across congruent and incongruent trials to calculate accuracy, and thus the behavioural data provide no insight into the question of implicit categorical perception in OSM). There are two possible interpretations of the ERP data. One is that the methodology and analysis obscured a true effect of implicit categorical perception in OSM. Notably, masking was far from complete (70% accuracy on the delayed mask offset trials), meaning that the ‘*masked*’ trials included a substantial mix of trials where the target was perceived in addition to those where it was not. Evidence from visual masking, both with OSM (Goodhew et al., 2011b), and with other more traditional forms of masking, on both behavioural and electrophysiological measures (Eimer & Schlaghecken, 1998; Schlaghecken & Eimer, 2000), indicate that ‘*aware*’ versus ‘*unaware*’ processing of content can differ *qualitatively* differ. Notably for Reiss and Hoffman (2007), this has been found to be true with face stimuli (Bennett, Lleras, Oriet, & Enns, 2007; Kiss & Eimer, 2008). This is important because if qualitatively different effects of equivalent magnitude are averaged over, it could yield an apparent null result. For instance, a priming effect of a 20 ms advantage to the congruent condition for *aware* trials coupled with a 20 ms advantage for the *incongruent*

condition on the unaware trials (assuming equi-probable trial types) would dilute to 0ms priming when the aware and unaware trials are averaged over. It would have been preferable had Reiss and Hoffman (2007) examined the ERP data separately for masked trials on which participants correctly identified the target and those on which they had not. This would have allowed for the assessment of such possibilities. As it stands, the possibility remains that the treatment of the data obscured an effect that would have revealed implicit categorical perception in OSM.

The second interpretation of Reiss and Hoffman's (2007) as reported is that perhaps the null result is meaningful and reflective of the true state of affairs, and not a product of the methodological choices in the study. Given the clear evidence for implicit semantic perception in OSM in previous research with word stimuli (Goodhew et al., 2011b), and implicit categorical perception with individual letters here, why would there not be implicit processing of category boundaries such as that between faces versus houses? We can only speculate at this point. However, one seemingly likely possibility is the marked differences in *perceptual load* between pictorial faces, especially as deployed in Reiss and Hoffman (2007), versus simple linguistic stimuli, especially in their most basic form as isolated letters as used here. It is well-established in other contexts that perceptual load modulates the magnitude of processing of task-relevant content, such that there is greater processing under conditions of low load, which is attenuated or eliminated under conditions of high load (Lavie, 1995, 2005; Lavie, Beck, & Konstantinou, 2014). The target displays in Reiss and Hoffman (2007) consisted of four large complex real-world images and eight-dot masks in order to obscure the visibility of the target. This is considerably higher perceptual load than the displays used here, where either one letter was presented with four dots around it alone, or with one other letter a considerable distance away. Perceptual load is a likely candidate to be an important moderating factor in whether evidence for implicit categorical perception is obtained.

For the present purposes, we can be agnostic about whether object-updating or object substitution underlies OSM. This is because both of these mechanisms converge on a perceptual inference to devalue the representation of the target (i.e., not maintain an enduring representation of it) and consciously represent only the mask, due to the close spatial and temporal proximity of the target and mask, and that was the functional process that we were interested in studying here. For our purposes, it is not important whether this inference is generated via hypothetical *substitution* or *updating* processes. One might argue that the semantic-level effects found here lend greater credence to the object-substitution framework underlying OSM, since object substitution is often couched in terms of re-entrant processing mechanisms (Di Lollo et al., 2000; Dux, Visser, Goodhew, & Lipp, 2010; Kotsoni et al., 2007; Weidner et al., 2006), and re-entrant processing would permit semantic-level interactions to occur. However, while object-updating theory is more typically focussed on cognitive-level explanations, this process could equally be achieved via re-entrant processing mechanisms, even if not in the precise way specified by the re-entrant processing explanation of object substitution. The present results, therefore, do not, and need not, adjudicate between these different explanations of OSM.

Finally, we would like to add a caveat about our use of the term ‘failure’ in the context of visual awareness throughout this manuscript (e.g., failure to form enduring representation of the target in the context of object-updating framework). We use this term because it is a convenient way to refer to experimental paradigms in which target perception suffers, such as masking. And indeed, the circumstances in which OSM is typically measured push the visual system towards an inference that is objectively ‘wrong’, resulting in a failure to perceive a stimulus that was physically present. However, this does not necessarily imply that OSM represents a *dysfunction* in vision. If, for example, when weighing up whether or not to update an object file in the presence of dynamic input, the system *always* decided in favour of

updating, then we would be unable to keep track of continuing objects as they move and change and disappear behind other objects (Burke, 1952; Hollingworth & Franconeri, 2009). So in this sense, OSM reflects the functioning of an adaptive mechanism, rather than a system struggling to encode rapidly presented objects. This makes OSM a fundamentally useful window into these preconscious visual-cognitive inferences that occur seamlessly, allowing us to have a stable and coherent percept of the world around us filled with recognisable objects.

In conclusion, here we discovered an interaction between spatial repetition blindness and OSM. That is, the repetition of a target-like stimulus elsewhere in the visual scene exacerbates masking relative to a stimulus unlike the target. The repeated item need not be identical to the target, but instead only needs to belong to the same category (or ‘type’). This supports a common basis to OSM and RB, namely the failure to consciously perceive two objects close in space and time. This result contrasts with that found in visual crowding, where presentation of a stimulus identical to the target reduces crowding. The results are also convergent evidence for implicit post-featural processing of the target despite effective masking, but also take this a step further, demonstrating that the category-level content of the target can influence masking, and therefore the process of object file consolidation.

Acknowledgements

This research was supported by an Australian Research Council (ARC) Discovery Early Career Research Award (DE140101734) awarded to S.C.G., and an ARC Discovery Project (DP110104553) grant awarded to M.E, and a UK Medical Research Council (MRC) Career Development Award to J.A.G. The authors thank Reuben Rideaux for assistance with the data collection. Correspondence regarding this study should be addressed to Stephanie Goodhew (stephanie.goodhew@anu.edu.au), Research School of Psychology, The Australian National University.

References

- 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. (2013). Set size and mask duration do not interact in object-substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, 39(3), 646-661. doi: 10.1037/a0030240
- Bavelier, D. (1994). Repetition blindness between visually different items: the case of pictures and words. *Cognition*, 51(3), 199-236. doi: 10.1016/0010-0277(94)90054-X
- 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
- 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
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226(5241), 177-178.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433-436. doi: 10.1163/156856897X00357
- Broadbent, D. E. (1958). *Perception and communication*. Elmsford, New York: Pergamon Press.
- Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, 4(3), 121-138. doi: 10.1080/17470215208416611
- Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature*, 371(6497), 511-513. doi: 10.1038/371511a0

- Camp, S. J., Pilling, M., Argyropoulos, I., & Gellatly, A. (2015). The Role of Distractors in Object Substitution Masking. *Journal of Experimental Psychology: Human Perception and Performance*. doi: 10.1037/xhp0000065
- 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. doi: 10.1167/9.10.4
- Chen, C.-M., Lakatos, P., Shah, A. S., Mehta, A. D., Givre, S. J., Javitt, D. C., & Schroeder, C. E. (2007). Functional Anatomy and Interaction of Fast and Slow Visual Pathways in Macaque Monkeys. *Cerebral Cortex*, 17(7), 1561-1569. doi: 10.1093/cercor/bhl067
- 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
- Chun, M. M., & Cavanagh, P. (1997). Seeing two as one: Linking apparent motion and repetition blindness. *Psychological Science*, 8(2), 74-79. doi: 10.1111/j.1467-9280.1997.tb00686.x
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, 27(3), 183-228. doi: 10.3758/BF03204258
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology*, 1(1), 42-45.
- 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.

- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222. doi: 10.1146/annurev.ne.18.030195.001205
- 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* (pp. 9-42). New York: Psychology Press.
- 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
- 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
- Egeth, H. E., & Santee, J. L. (1981). Conceptual and perceptual components of interletter inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, 7(3), 506-517. doi: 10.1037/0096-1523.7.3.506
- 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
- Enns, J. T., & Di Lollo, V. (2002). What competition? *Trends in Cognitive Sciences*, 6, 118.
- Fagot, C., & Pashler, H. (1995). Repetition blindness: Perception or memory failure? *Journal of Experimental Psychology: Human Perception and Performance*, 21(2), 275-292. doi: 10.1037/0096-1523.21.2.275
- Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2014). Size (mostly) doesn't matter: The role of set size in object substitution masking. *Attention, Perception, & Psychophysics*. doi: 10.3758/s13414-014-0692-5
- Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2015). Object substitution masking for an attended and foveated target. *Journal of Experimental Psychology: Human Perception and Performance*. doi: 10.1037/xhp0000024
- Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, 14(9), 1195-1201. doi: 10.1038/nn.2889
- Geiger, G., & Lettvin, J. Y. (1986). Enhancing the perception of form in peripheral vision. *Perception*, 15(2), 119-130. doi: 10.1068/p150119
- Gellatly, A., Pilling, M., Cole, G., & Skarratt, P. (2006). What is being masked in object substitution masking? *Journal of Experimental Psychology: Human Perception and Performance*, 32(6), 1422-1435. doi: 10.1037/0096-1523.32.6.1422
- Goldfarb, L., & Treisman, A. (2011). Repetition blindness: The survival of the grouped. *Psychonomic Bulletin & Review*, 18(6), 1042-1049. doi: 10.3758/s13423-011-0135-4
- Goodhew, S. C., Boal, H. L., & Edwards, M. (2014). A magnocellular contribution to conscious perception via temporal object segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, 40(3), 948-959. doi: 10.1037/a0035769

- 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., Edwards, M., Boal, H. L., & Bell, J. (2015). Two objects or one? Similarity rather than complexity determines objecthood when resolving dynamic input. *Journal of Experimental Psychology: Human Perception and Performance*, 41(1), 102-110. doi: 10.1037/xhp0000022
- Goodhew, S. C., Edwards, M., Ferber, S., & Pratt, J. (2015). Altered visual perception near the hands: A critical review of attentional and neurophysiological models. *Neuroscience & Biobehavioral Reviews*, 55, 223-233. doi: 10.1016/j.neubiorev.2015.05.006
- Goodhew, S. C., Gozli, D. G., Ferber, S., & Pratt, J. (2013). Reduced temporal fusion in near-hand space. *Psychological Science*, 24(6), 891-900. doi: 10.1177/0956797612463402
- Goodhew, S. C., Pratt, J., Dux, P. E., & Ferber, S. (2013). Substituting objects from consciousness: A review of object substitution masking. *Psychonomic Bulletin & Review*, 20(5), 859-877. doi: 10.3758/s13423-013-0400-9
- 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

- Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional Averaging Explains Crowding with Letter-like Stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31), 13130-13135. doi: 10.2307/40484661
- Grill-Spector, K., & Kanwisher, N. (2005). Visual Recognition: As Soon as You Know It Is There, You Know What It Is. *Psychological Science*, 16(2), 152-160. doi: 10.1111/j.0956-7976.2005.00796.x
- 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
- Harris, I. M., Wong, C., & Andrews, S. (2015). Visual field asymmetries in object individuation. *Consciousness and Cognition*, 37, 194-206. doi: 10.1016/j.concog.2015.09.004
- 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
- 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
- Ibbotson, M., & Krekelberg, B. (2011). Visual Perception and Saccadic Eye Movements. *Current Opinion in Neurobiology*, 21(4), 553-558. doi: 10.1016/j.conb.2011.05.012
- Irwin, D. E., & Brockmole, J. R. (2004). Suppressing Where but Not What: The Effect of Saccades on Dorsal- and Ventral-Stream Visual Processing. *Psychological Science*, 15(7), 467-473. doi: 10.1111/j.0956-7976.2004.00703.x

- 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%2892%2990007-O
- Kanwisher, N. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, 27(2), 117-143. doi: 10.1016/0010-0277%2887%2990016-3
- Kanwisher, N. (1991). Repetition blindness and illusory conjunctions: Errors in binding visual types with visual tokens. *Journal of Experimental Psychology: Human Perception and Performance*, 17(2), 404-421. doi: 10.1037/0096-1523.17.2.404
- Kanwisher, N., & Driver, J. (1992). Objects, attributes, and visual attention: Which, what and where. *Current Directions in Psychological Science*, 1(1), 26-31. doi: 10.1111/1467-8721.ep10767835
- Kanwisher, N., Driver, J., & Machado, L. (1995). Spatial repetition blindness is modulated by selective attention to color or shape. *Cognitive Psychology*, 29, 303-337.
- Kanwisher, N., Kim, J. W., & Wickens, T. D. (1996). Signal detection analyses of repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1249-1260.
- Kanwisher, N., & Potter, M. C. (1989). Repetition blindness: The effects of stimulus modality and spatial displacement. *Memory & Cognition*, 17(2), 117-124. doi: 10.3758/BF03197061
- Kanwisher, N., & Potter, M. C. (1990). Repetition blindness: Levels of processing. *Journal of Experimental Psychology: Human Perception and Performance*, 16(1), 30-47. doi: 10.1037/0096-1523.16.1.30
- Kastner, S., & Pinsk, M. A. (2004). Visual attention as a multilevel selection process. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 483-500. doi: 10.3758/CABN.4.4.483

- Kiss, M., & Eimer, M. (2008). ERPs reveal subliminal processing of fearful faces. *Psychophysiology*, 45(2), 318-326.
- 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
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205. doi: 10.1126/science.7350657
- Lamme, V. A. F. (2000). Neural mechanisms of visual awareness: A linking proposition. *Brain and Mind*, 1, 385-406.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3), 451-468. doi: 10.1037/0096-1523.21.3.451
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, 9(2), 75-82. doi: 10.1016/j.tics.2004.12.004
- Lavie, N., Beck, D. M., & Konstantinou, N. (2014). Blinded by the load: attention, awareness and the role of perceptual load. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1641). doi: 10.1098/rstb.2013.0205
- Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240(4853), 740-749. doi: 10.1126/science.3283936
- 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

- 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
- Luo, C. R., & Caramazza, A. (1995). Repetition blindness under minimum memory load: Effects of spatial and temporal proximity and the encoding effectiveness of the first item. *Perception & Psychophysics*, 57(7), 1053-1064. doi: 10.3758/BF03205464
- Luo, C. R., & Caramazza, A. (1996). Temporal and spatial repetition blindness: Effects of presentation mode and repetition lag on the perception of repeated items. *Journal of Experimental Psychology: Human Perception and Performance*, 22(1), 95-113. doi: 10.1037/0096-1523.22.1.95
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide*: Lawrence Erlbaum Associates.
- Marohn, K. M., & Hochhaus, L. (1988). Different-case repetition still leads to perceptual blindness. *Bulletin of the Psychonomic Society*, 26(1), 29-31.
- Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin*, 81(12), 899-917. doi: 10.1037/h0037368
- Maunsell, J. R., Ghose, G. M., J.A., A., McAdams, C. J., Boudreau, C. E., & Noerager, B. D. (1999). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Visual Neuroscience*, 16, 1-14.
- 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
- 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

- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4(7), 739-744. doi: 10.1038/89532
- 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
- Pilling, M., Gellatly, A., Argyropoulos, Y., & Skarratt, P. (2014). Exogenous spatial precuing reliably modulates object processing but not object substitution masking. *Attention, Perception, & Psychophysics*, 76(6), 1560-1576. doi: 10.3758/s13414-014-0661-z
- Pokorny, J. (2011). Review: Steady and pulsed pedestals, the how and why of post-receptoral pathway separation. *Journal of Vision*, 11(5), 1-23. doi: 10.1167/11.5.7
- 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
- Sayim, B., Greenwood, J. A., & Cavanagh, P. (2014). Foveal target repetitions reduce crowding. *Journal of Vision*, 14(6), 1-12. doi: 10.1167/14.6.4
- Schlaghecken, F., & Eimer, M. (2000). A central-peripheral asymmetry in masked priming. *Perception & Psychophysics*, 62, 1367-1382.

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

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, 15(4), 160-168. doi: 10.1016/j.tics.2011.02.005