RADAR

Research Archive and Digital Asset Repository

OXFORD BROOKES UNIVERSITY

Camp, S, Pilling, M, Argyropoulos, I and Gellatly, A

The role of distractors in Object Substitution Masking

Camp, S, Pilling, M, Argyropoulos, I and Gellatly, A (2015) The role of distractors in Object Substitution Masking. *Journal of Experimental Psychology: Human Perception and Performance*, 41 (4). pp. 940-957.

doi: 10.1037/xhp0000065

This version is available: https://radar.brookes.ac.uk/radar/items/10d8c3d7-976f-4d04-930f-cb98f7156341/1/

Available on RADAR: Semptember 2016

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the postprint version of the journal article. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

WWW.BROOKES.AC.UK/GO/RADAR

Title: The role of distractors in Object Substitution Masking

Sarah Jayne Camp[1], Michael Pilling[1], Ioannis Argyropoulos[2] & Angus Gellatly[1 3]

[1] Oxford Brookes University, Oxford, UK

[2] Heriot-Watt University, Edinburgh, UK

[3] Institute of High Performance Computing, A*STAR, Singapore

In press: (2015) Journal of Experimental Psychology: Human Perception & Performance.

Author contact

Sarah Camp,

Department of Psychology,

Oxford Brookes University,

Gipsy lane,

Oxford, OX3 0BP

Email: sarah.camp-2012@brookes.ac.uk

Author note

Some of these findings were presented at ECVP 2014 (Belgrade, Serbia), and at the British Psychological Society Cognitive Section 2014 annual conference (Nottingham, UK). The experiments in this paper will form part of the first author's doctoral thesis. We would like to thank three anonymous reviewers for their comments on an earlier version of this manuscript.

Abstract

In Object Substitution Masking (OSM) a surrounding mask (typically comprising of four dots) onsets with a target but lingers after offset; under such conditions the ability to perceive the target can be significantly reduced. OSM was originally claimed to occur only when a target was not the focus of attention, for instance when embedded in an array of distractors (Di Lollo, et al., 2000). It was argued that the distractors influenced the time taken for focal attention to reach the target. Some recent work, however, failed to find any such distractor influence: the effect of mask duration being independent of set size when steps were taken to avoid ceiling effects in the smallest set size condition (Argyropoulos et al., 2013; Filmer et al., 2014a). In three experiments we repeatedly found that set size manipulations *can* interact with mask duration (where neither ceiling nor floor effects are evident), the effect of the mask on target perceptibility being amplified according to the number of distractor items. However, a further experiment (Exp. 4) showed that crowding by nearby distractors was actually responsible for this 'set size' effect. When decoupled from crowding, set size alone did not interact with masking, though it did influence overall accuracy. Thus the presence of distractors does influence OSM but not in the way originally assumed by Di Lollo and colleagues in their model (Di Lollo et al., 2000). The crowding × OSM interaction suggests that the two phenomena involve partly overlapping mechanisms.

Abstract word count: 244

Keywords: Object substitution masking, distractors, set size, crowding, mask duration

The role of distractors in Object Substitution Masking

Object substitution masking (OSM) is a form of visual masking in which the mask surrounds a briefly presented target, reducing target perceptibility the longer it trails the target offset (Di Lollo, Enns & Rensink, 2000). OSM has been the focus of a large body of research in the 15 years since the phenomenon was first described (see Goodhew, Pratt, Dux & Ferber, 2013 for a review). OSM, like other forms of masking, is of interest to vision scientists because it reveals something about the microgenesis and spatiotemporal dynamics of perception (Werner, 1957; Bachmann, 2006). OSM is of particular interest because of the seeming object-based nature of the interactions which underlie it: The phenomenon seems to reflect – at least in part – the process by which the visual system resolves whether two stimuli in close spatial and temporal proximity are perceived as the same or different perceptual objects (Lleras & Moore, 2003; Moore & Llearas, 2005; Pilling & Gellatly, 2010; Goodhew, Edwards, Boal & Bell, 2015).

OSM is clearly differentiated from other forms of backward masking such as *noise masking* (Kinsbourne & Warrington, 1962) and *pattern masking* (Turvey, 1973) by virtue of the OSM mask lacking spatial overlap with the target. OSM seems more akin to the phenomenon of *metacontrast masking* (MM, Alpern, 1953), in having a surrounding but non-spatially overlapping mask. However despite this superficial resemblance OSM and MM are distinct phenomena: A striking aspect of OSM, is that the mask can be very sparse, while in MM the mask completely surrounds the target and often the inner surface of the mask tightly follow the outer contours of the target, in OSM just four surrounding dots are sufficient for the effect to occur (Di Lollo et al., 2000; Enns, 2004)[i]. As a consequence, OSM is often referred to in the literature as *four-dot masking* (Di Lollo et al, 2000; Enns & Di Lollo, 2000, Dell'Acqua, Pascali, Jolicoeur & Sessa, 2003; Vroomen & Keetels, 2009).

Furthermore OSM has been found to be largely insensitive to the amount of contour in the surrounding mask. Evidence for this can be seen in a study by Enns (2004), who directly compared a four-dot mask with the ring mask typical of MM studies (along with several other types of mask).

SOA between target and mask onset was varied and target and mask duration fixed at 30 ms. With short target-mask SOAs (up to 50 ms) the type of mask was found to influence the masking observed: in particular the presence of the MM mask led to reduced target perceptibility while the four-dot mask had no effect; however at longer SOAs (>=150 ms) masking was the same irrespective of the type of mask given. Enns argued based on these results that there are two distinct masking processes: *object formation masking* and *object substitution masking*: object formation masking is deemed an early visual process involving target-mask interactions such as lateral inhibition, while object substitution masking is a later process which is largely independent of the physical characteristics of the mask. Consistent with this fact OSM also largely fails to show the characteristic decline in strength exhibited in MM as the spatial separation between the target and surrounding mask elements is increased (Di Lollo et al. 2000; Guest, Gellatly & Pilling, 2011; *cf*. Lefton, 1973). Finally, unlike MM, in OSM the mask's time of onset with respect to the target is not critical. OSM is typically demonstrated with a target and mask which onset simultaneously; however there is a temporal window – to date not fully investigated – within which OSM can occur when the target onset precedes that of the mask. In all cases is the extent to which the mask lingers on screen following target offset which is critical in determining masking (Enns & Di Lollo, 1997; Di Lollo et al., 2000; Enns, 2004; *cf*. Jannati, Spalek & Di Lollo, 2013).

The fact that OSM is mostly insensitive to mask contour combined with the lack of importance of relative mask and target onset times seems to rule out conventional explanations of masking, i.e. those forged in terms of local inhibitory interactions or other feed-forward processes (Kahneman, 1968; Coltheart, 1975; Breitmeyer & Ganz, 1976; Breitmeyer & Öğmen, 2000; *cf*. Frances & Hermens, 2002; Di Lollo, Enns & Rensink 2002; Bridgeman, 2006; Pöder, 2013; Di Lollo, 2014; Pöder, 2014). Instead, it is suggested OSM emerges from the process of re-entrant exchanges between different levels of the visual system (Di Lollo et al. 2000; Di Lollo et al., 2002; Jannati, et al., 2013; Di Lollo, 2014; *cf*. Frances & Hermens, 2002; Pöder, 2013, 2014). Neurophysiological evidence has shown that different levels within the visual system are connected by both ascending and descending pathways (Felleman & van Essen, 1991), an arrangement that allows iterative communication between lower level visual areas representing stimulus-bound information and higher level areas responsible for the generation of perceptual hypotheses (Mumford, 1992; Lamme, Supèr & Spekreijse, 1998). The *Object Substitution Theory of Masking* (hereafter OSTM) purports that OSM occurs as a consequence of such iterative communications which occur during the normal course of perception (Di Lollo et al. 2000; Jannati, et al., 2013). As an object appears in view the retinotopic pattern of the image cast by the stimulus is encoded in V1. This input level representation is fed forward to higher extrastriate visual areas where cells have broader spatial tuning but are sensitive to more abstract stimulus properties, such as shape. Extrastriate cells form one or more *perceptual hypotheses* about the newly presented stimulus which are then fed back via re-entrant pathways to check against the representation based on current input. Representation at the input level does not, by itself, result in awareness. For awareness to occur the current input representation has to be successfully matched with the (delayed) re-entrant signal.

In the OSTM masking is viewed as a consequence of the inherent sluggishness of the re-entrant architecture in responding to rapid changes in visual input. Consider what happens according to the model in the OSM paradigm. The presentation of a target on screen surrounded by a mask generates a low level input representation of the two items; the contents of this representation are fed upwards to higher extrastriate visual areas where one or more perceptual hypothesis (of the target+mask) is formed and then sent back to the input level for comparison. If target and mask disappear simultaneously from the screen after only a few tens of milliseconds then awareness of the target may still ensue despite the brevity of the stimulus: Though possibly no longer sustained by current input a residual trace of activity of both target and mask will be retained in the input level representation by cells firing in V1. This means that the input level will still be broadly consistent with the descending perceptual hypothesis based on the earlier feedforward sweep. With a trailing mask, however, the situation is different: here the input level will contain the mask (sustained by current retinal input) and a rapidly decaying trace of the target at the point the re-entrant signal is compared. The mismatch is likely to lead to a rejection of the initial hypothesis (of target plus mask) and an instigation of a new iterative cycle based on current retinal input (i.e. of the mask alone surrounding a blank space). The presence of the trailing mask essentially biases the outcome of the iterative exchange away from perception of the target and towards that of the mask alone. To wit, the trailing mask *substitutes* the target in conscious visual experience (Di Lollo et al. 2000). The longer the trailing mask duration the more probable it becomes that the initial hypothesis is rejected and, consequently, that masking is found.

Thus, the OSTM neatly accounts for the importance of target duration in masking. The theory also accounts for a further ostensive feature of OSM, its seeming dependence on set size. In the original experiments of Di Lollo et al. (2000) the set size of the stimulus array was varied between one (target alone) and sixteen items (target and fifteen distractors), the target denoted within the array by the surrounding four dot mask. With the target alone performance remained largely flat (and near to ceiling) irrespective of mask duration; i.e. no detectable OSM was present. As set size was increased the effect of the trailing mask on performance became increasingly evident. This was observed as a significant interaction between the set size and mask duration variables. The OSTM interprets this interaction as a consequence of the role played by spatial attention in perceiving the target. The iterative process is argued to begin only when the target falls within the window of attention. When only the target is present in a display, attention is argued to be rapidly deployed to its location meaning that iterative processing tends to be completed before the input level information of target and mask is displaced by information about the trailing mask alone. As the number of display items increases the time required for attention to focus on the target location will increase proportionally (a variable described as *time to contact* in the *CMOS* model – the formal mathematical implementation of OSTM, Di Lollo et al., 2000). Thus, with a target located amongst multiple distractors the delayed attentional focus makes the target more vulnerable to substitution by the mask.

Di Lollo et al.’s claim regarding the set size by mask duration interaction has been widely accepted in OSM research. This has been the case despite something of a shift in the theoretical focus of OSM research: some recent interpretations of OSM – while in some cases still explicitly retaining the re-entrant architecture as the underpinning of masking – have tended to focus more on understanding the nature of the object-level interactions involved (Lleras & Moore; 2003; Moore & Lleras, 2005; Enns, Lleras & Moore, 2009; Goodhew et al. 2015; though *cf*. Jannati et al. 2013; Carlson, Rauschenberger & Verstraten, 2007; Weidner, Shah & Fink, 2006; Harris, Ku & Woldorff, 2013, for recent supporting behavioural, brain imaging, and electrophysiological evidence suggesting the role of re-entrant processes in OSM; see also Di Lollo, 2013, for a recent updated theoretical description of OSM presented within a re-entrant architecture). It has been shown that the competition between target and mask which underlies OSM, rather than being between two separate perceptual objects, actually occurs within a single object representation (Lleras & Moore, 2003). Di Lollo et al. (2000) originally claimed that the representation of the mask substitutes that of the target as the focus of conscious perception; Lleras & Moore (2003) argued that the object-based process is better described as one of *updating* rather than substitution (see also Moore and Lleras, 2005; Pilling & Gellatly, 2011; Goodhew et al., 2015) suggesting that OSM occurs under conditions where mask and target come to be represented by the visual system as the same perceptual object because of their close spatiotemporal profile. When the mask trails the target offset, the visual system accordingly updates this token representation according to current input (containing the mask alone), erasing from it the target’s features and thus rendering them inaccessible to conscious vision.[ii]

Despite this shift of theoretical emphasis the object updating account still assumes the basic theoretical predictions of the OSTM to hold true (see Enns, Lleras & Di Lollo, 2006, for a revised account of the OSTM describing object updating within an OSTM framework). In particular the account implicitly assumes that set size interacts with OSM in the manner the OSTM originally claimed. Indeed the vast majority of published studies on OSM – including those done from an object updating perspective – have presented the target within an array of several distractor items, a design feature done to putatively maximise the OSM effect obtained. Recently, however, evidence has come to light which suggests that set size actually has no effect on the masking observed in OSM. This was first noted by Argyropoulos, Gellatly, Pilling & Carter (2013) who closely replicated some of Di Lollo et al.'s original experiments. Despite their attempts at replication in their results across several experiments they failed to reproduce the set size by mask duration interaction. Instead OSM was found to be the same irrespective of whether the target was alone in the array or present with three or fifteen other items. Though set size did not influence OSM, it did affect overall target perceptibility: In all experiments overall accuracy (i.e. irrespective of mask duration) reduced as set size increased.

The authors argue that the difference between their results and those of Di Lollo et al. (2000) reflect, for the most part, differences in the performance range in which OSM was measured. In a large proportion of the studies conducted by Di Lollo and colleagues, average performance for set size one was close to or at ceiling. This meant that masking, which otherwise presumably would have been observed, was not present in the data because the discrimination task was too easy when only one item was present. When more items were introduced this brought performance into a range in which changes in perceptibility were reflected in the accuracy measure. Thus the 'interaction' between set size and mask duration was argued to be artifactual, due not to set size itself but to the restriction on measurable performance when set size was small. The discrimination tasks in Argyropoulos et al.'s experiments were closely modelled on those of Di Lollo et al.'s studies (reporting the missing side of a Landolt square which could be in one of four cardinal orientations). Stimuli were smaller in Argyropoulos et al. (subtending approximately half the visual angle of those of Di Lollo et al.). As a consequence the baseline discrimination task was more difficult and overall accuracy tended to be lower than in the original Di Lollo et al. experiments. This aspect of the experiment brought performance in the set size one condition in particular into a measurable range on the accuracy measure. Under these conditions it was found that, rather than being absent, masking was the same when the target was alone as when distractors were present.

Other experiments in Argyropoulos et al. (2013) replicated the detection paradigm used in one of Di Lollo et al. (2000)'s experiments. Observers reported the presence or absence of a vertical bar on a circle at the location surrounded by the four dots. Here Di Lollo et al.'s data was away from ceiling, even for the set size one condition, and yet they reported a significant interaction between set size and mask duration. However Di Lollo et al. reported accuracy only for those trials in which the target bar was present, meaning accuracy was confounded with any potential response bias. Argyropoulos et al. showed that when a correction was applied to the data to take into account the false alarm rate, the interaction between set size and mask duration was no longer found. Argyropoulos et al. concluded that the interactions reported by Di Lollo et al. were spurious and thus that distractors played no role in OSM.

Other recent evidence has broadly supported Argyropoulos et al. (2013)'s claim regarding the role of set size in OSM. Filmer, Mattingley & Dux (2014a), while finding a significant main effect of set size, similarly failed to find a significant interaction between this variable and mask duration when ceiling effects were controlled for. Again this failure to produce a significant interaction between these key variables was observed across several of their experiments. Notably, no interaction was observed even when mask duration was extended beyond 400 ms. Earlier work by Goodhew Visser, Lipp & Dux, 2011 and Goodhew, Dux, Lipp & Visser , 2012, shows that extending mask duration beyond this limit leads to a partial recovery from OSM showing that the mask duration function is better considered as U-shaped rather than monotonic. Filmer et al. (2014a)'s finding in this regard demonstrated that OSM is unaffected by set size even under conditions where the entire masking function is given. In a further experiment Filmer et al. (2014a) also compared the effect of set size and mask duration in an OSM paradigm in which the stimulus array was arranged in a grid formation. The earlier experiments reported in the Filmer et al. (2014a) paper had – following from those in Argyropoulos et al. (2013) – deviated from the original paradigms of Di Lollo et al. (2000) in presenting the stimuli in a circular arrangement around fixation. This was done as a deliberate attempt to control for eccentricity, though it was arguable that the failure to replicate the set size interaction was a consequence of Di Lollo et al.'s grid arrangement not being used. However the interaction again failed to materialise even under these conditions. Filmer et al. (2014a) also performed a re-analysis of all their data as a further check of the presence of an interaction. As percent correct is a probability measure, it means that two statistically independent variables will be seen as multiplicative in their aggregate effects within this metric. Because of this it is more appropriate to analyse the log transformed scores than the raw scores in order to determine multiplicative relationships between such variables (see Schweikert, 1985 for a proof). However a significant interaction still failed to materialise even with this analysis of log transformed scores.

Pilling, Gellatly, Argyropoulos & Skarrett (2014), also recently presented data which support the position that distractors play no role in OSM, at least in the way originally described in the OSTM model. In this set of experiments distractors were always present in the stimulus array but set size itself was not varied. What was varied instead in four of the experiments was the time of onset of a spatial cue which indicated the target location. The cue could be as early as 150 ms before the onset of the target array or as late as being presented simultaneously with the array. With a non-zero pre-cue accuracy in reporting the target item was substantially increased. In three of the experiments, however, the OSM effect was the same regardless of the cue onset time. Only in one experiment did a pre-cue yield a modest reduction in the degree of OSM produced by a trailing mask. Additionally, a fifth experiment varied the validity of a luminance cue that onset 100 ms before the target display. A valid cue indicated the location of the target (surrounded by four dots) and an invalid cue indicated the location of a distractor. Performance was better with valid than invalid cues but the OSM effect was the same for both. Thus, directing the window of attention away from the distractor locations and towards the target location by either method had in most cases no effect on masking. As with set size, pre-cueing affected target perceptibility but for the most part did so equally for masked and unmasked trials. More recently still, Filmer, Mattingley & Dux (2014b) showed that OSM occurs robustly even with just a single target presented at an attended and foveated location. This evidence in particular demonstrates that diffuse (or misdirected attention) is not a precondition for OSM in the way Di Lollo et al. assumed.

In summary the recent evidence, described above, from Argyropoulos et al. (2013), Filmer et al. (2014a, 2014b), and Pilling et al. (2014) seem inconsistent with the possibility that distractors are important in OSM; at least in the way that Di Lollo and colleagues originally conceived. Despite this we decided to further revisit this question. Our motivation came in part from analysis of data from work done in our lab looking at target and mask preview effects in OSM[iii]. In this work different display sizes were used in order to assess the consequences for these preview effects. While the set size manipulation had relatively little influence on the preview effects it did seem to influence the overall amount of masking. Because of the number of factors involved and because it was not the variable of primary interest, set size was only manipulated as a between participants factor. However these data did suggest to us that the dismissal of set size as a factor in OSM may have been premature.

The current paper starts by using the same digit identification task in which this unexpected set size effect, just referred to, was found. In this first experiment observers had to report the identity of a digit surrounded by four dots presented in a display of distractors. Set size and mask duration were both varied as within-participants factors. To pre-empt the results of this first experiment, contrary to Argyropoulos et al. (2013) and Filmer et al., (2014) a significant interaction was obtained between set size and mask duration. Further experiments determined that the interaction was in no way specific to digit identification; the interaction was reproduced both in target detection and gap discrimination tasks. A final experiment showed that though the set size variable did interact with OSM, set size itself was not the factor driving the interaction. Set size only influenced OSM to the extent that it increased proximity of distractors flanking the target. When set size was decoupled from crowding by designing the stimulus arrays in a way which allowed the two to vary independently, set size influenced performance but only crowding interacted with mask duration.

Experiment 1: digit identification task

In this first experiment a digit identification task was used. Pilling et al. (2014) used digits as stimuli and found robust OSM for this stimulus class and found that for almost all participants, responses across all conditions fell within a measurable range outside of ceiling and floor despite the substantial variation in performance produced by the manipulations . As a method of investigating OSM this task arguably holds an advantage over the Landolt discrimination task used by Di Lollo et al (2000) and Argyopolous et al. (2013). The task involves ten response options (0-9), making the probability of correct random responding .1. The fact that the baseline probability of a correct response is lower makes it easier to distinguish non-random from random responding in a participant under conditions where accuracy is expected to be low. Thus a digit identification task is well suited to the purpose of measuring target perceptibility and exploring the potential interactive effects of set size and mask duration. If OSM occurs independently of set size then set size and mask duration should not interact.

Method

Participants.

Seventeen first year Oxford Brookes Psychology students (14 female) took part in the experiment. All participants reported normal or corrected-to-normal visual acuity. This and all further experiments were approved by the Oxford Brookes University Research Ethics Committee.

All participants gave informed consent and received course credits or payment for taking part in the experiment.

Design.

The experiment had two within-samples independent factors, each with three levels. These were set size (1, 6, or 12 items) and mask duration (0, 60, or 180 ms). The dependent variable was identification performance, measured by the percentage of correct responses.

Stimuli and procedure.

The experiment was conducted in a darkened and sound deadened room with back lighting. Stimuli were presented on a 20 inch Sony Trinitron CRT computer monitor set at a resolution of 1024×768 running at a 100Hz refresh rate. The monitor was controlled by an Intel Pentium 4 (2.66 GHz) PC fitted with a NVDIA GeForce 4 graphics card. The monitor was viewed by the participant from a distance of approximately 110cm. Software written in the *BlitzMax* programming language (BlitzMax V. 1.5; Sibly, 2011) controlled all aspects of stimulus presentation, randomisation and response recording. All stimuli were black (0.03 cd/m2) presented on a white (97 cd/m2) background. The stimulus array consisted of 1, 6 or 12 digits depending on the set size condition. Digits were in Arial font Pt. 32 (0.47° subtended visual angle in height) and were centred on the circumference of a virtual circle (itself with a radius of subtending 3.9° from the centre of fixation to the centre of each digit) with a fixation cross at its centre. Digits were evenly spaced apart from one another on the virtual circle. Participants were required to identify the target digit (indicated by the surrounding mask). The mask consisted of four dots forming a virtual square (subtending .36° in height/width) around the target. The dots comprising the mask were each .10° of visual angle in width/height. The identity of the target digit was randomly determined on each trial with the constraint that each of the ten digits appeared with equal frequency within all trial types. Distractor digits were chosen randomly for each trial in which distractors were present. Each trial began with a blank white screen presented for 500 ms followed by the onset of the fixation cross which was accompanied by a brief alerting tone. After a further 250 ms the stimulus array was presented with the four dot mask surrounding the target digit. The stimulus array remained on screen for 40 ms and was followed by the trailing mask either for 0 (non-masking control condition), 60, or 180 ms. The fixation cross was onscreen throughout these frames and remained until the participant responded. Responses were made on a standard computer keyboard, pressing a key from 0-9 corresponding to the target identity. Immediate aural error feedback was given following a key press for an incorrect response. On a key press the fixation cross disappeared and a new trial was instigated. A schematic depiction of the trial sequence is given in Figure 1. There were 540 randomly ordered trials, 60 for each combination of mask duration and set size, presented in 10 distinct blocks. The computer prompted the participant to have a brief break after every 54 trials. The experimental session was initiated by verbal instructions from the experimenter. Participants were informed that accuracy not speed of response was important for the experiment. Three randomly selected demonstration trials of the experiment with slowed display sequences were then shown to the participant. The participant then completed 30 practice trials which were randomly selected and where the timings were the same as the actual experiment followed by the actual experimental trials. The duration of the entire experimental session was approximately 30-40 minutes.

Insert Figure 1 about here

Results

The average percent correct responses in each factorial condition of mask duration and set size are shown in Figure 2 (A). These data were analysed using a repeated measures ANOVA with two factors, each with three levels: set size (1, 6, 12); mask duration (0, 60, 180). There were significant main effects of set size, $F(2,32)=217.16$, MSerror = 25.53, $p<.001$, partial $\eta^2=.93$, and mask duration $F(2,32)=125.30$, MSerror = 24.58, $p<.001$, partial $\eta^2 = .89$. Importantly, a significant interaction was found between set size and mask duration $F(4,64)= 17.56$, MSerror =17.56, $p<.001$, partial $\eta^2=.52$. Examination of Figure 2 indicates that the interaction was caused by the fact that the masking (as indexed by mask duration) increased with set size.

Simple effects t-tests were conducted and revealed that OSM was produced even for a set size of one. There was a significant difference in performance between a mask duration of 0 and 60ms ($t(16)=5.45$, $p<.001$) and between a mask duration of 0 and 180ms ($t(16)=4.76$, $p<.001$. There was no significant difference in performance between 60 and 180ms mask duration however ($t(16)=.66$, $p=.519$ indicating that substantial masking occurs with a 60ms mask duration. Further simple effects t-tests revealed that set size affects performance even in the absence of OSM. There was a significant difference in performance between set size 1 and 6 ($t(16)=5.38$, $p<.001$), 6 and 12 ($t(16)=4.01$, $p=.001$ and 1 and 12 ($t(16)=9.48$, $p<.001$ at a mask duration of 0ms.

Though we obtained an interaction between setsize and mask duration, caution must be exercised in interpreting it based on ANOVA analysis of raw percentage scores. As noted earlier, analysis of the log transformed scores is a more appropriate analysis for testing multiplicative relationships between variables where percent correct measures are being used. Thus for these data (and that of the other Experiments in this paper where the dependent variable was a percent correct measure) we repeated the ANOVA analysis of the interaction on log10 transformed accuracy scores (the same procedure described in Filmer et al., 2014a). This transformation did not markedly change the basic pattern of the data (see Figure 2, plate B); importantly the interaction between set size and mask duration was retained, $F(4,64)=28.60$, MSerror = .01, $p<.001$, partial $\eta^2=.64$. Thus, Experiment 1 clearly demonstrates that the strength of masking (as indexed by the effect of mask duration) was influenced by set size.

Insert Figure 2 about here

Discussion

The interaction found between set size and mask duration is inconsistent with the findings reported by both Argyropoulos et al. (2013) and Filmer et al. (2014). It is, however, consistent with the original finding reported by Di Lollo et al. (2000; see also Jiang & Chun 2000; Kahan & Mathis, 2002; Kotsoni,Csibra, & Mareschal, 2007); unlike the data in Di Lollo et al however the interaction we observed cannot be explained as an artefact of constraints in the measurement of performance. In all conditions participants were well below ceiling (the maximum participant score in any condition was 83.3%, the minimum 28.3%). One aspect of our current data was consistent with both Argyropoulos et al. and by Filmer et al.: the presence of distractors was wholly unnecessary for OSM to occur. What is different from Argyropoulos et al. and Filmer et al. is that the addition of distractor items did augment the effect of the trailing mask on performance.

The significant interaction is arguably surprising. Given the inconsistency in this regard, we felt it necessary to perform a further study to determine if the interaction would be repeated under different task conditions. It is notable that Argyropoulos et al. (2013), also failed to produce a set size by mask duration interaction with a detection task. In their task observers had to report whether or not there was a vertical bar present on a circle at the location of the four dots. There was no interaction when response bias was taken into account. A second experiment (Experiment 2) was therefore conducted in which the task was to detect rather than identify the target digit.

Experiment 2: Digit detection task

Experiment 2 was essentially the same as Experiment 1 in terms of the display sequence. However a digit was present inside the mask on only half the trials; on the others there was just a blank space. Participants had to make a present or absent response regarding whether there was a digit at the mask location. Pilot work showed that performance was at or near ceiling with the stimulus duration used in Experiment 1 (40 ms); therefore Experiment 2 had a briefer stimulus array (10 ms).

Experiment 2 served to test whether the result found in Experiment 1 was in some way specific to the digit identity task, or whether it could also be obtained under different task demands. We tested whether set size would interact with mask duration under these conditions or whether the two would have only independent effects on detection.

Method

Participants

Fifteen participants (7 female) took part in the experiment. All participants gave informed consent and received £7 remuneration for completing the experiment. All participants reported normal or corrected-to-normal visual acuity.

Stimuli and procedure

The stimulus sequence is shown in Figure 3. Participants were required to report whether or not there was a target digit present within the four dot mask using a corresponding key (Z or M) on the computer keyboard. Three set size conditions were given (1, 6 or 12), and the trailing mask duration was 0, 60 or 180 ms as in Experiment 1. There were 1080 experimental trials. The target digit was present on 50% of all trial types. Equal numbers of trials were given for each of the 18 factorial combinations of conditions. On trials in which a target digit was present each of the ten digits were shown with equal frequency within each factorial combination. The identity of the distractor digits was random. A demonstration and practice trials were given as in the previous experiment. Participants were instructed to emphasise accuracy in responding.

Insert Figure 3 about here

Results

The proportion of hits, was calculated for target present trials and of false alarms on target absent trials (these are shown in Figure 4, plates A and B respectively). From these data a signal detection measure (d-prime [d']) was calculated (Fig. 4, plate C), as was a measure of criterion (*C*; Fig. 4 , plate D[iv]). ANOVA analysis concentrated on the d' scores. It can be seen that performance decreased as both set size and mask duration increased; the lowest performance levels occurring at a set size of 12 and a mask duration of 180ms. A 2-way repeated measures ANOVA of the d' revealed a significant main effect of set size (F2,28) = 21.26, MSerror = .23, p <.001, partial η^2 =.60 and mask duration (F2,28) = 9.13, MSerror = 0.38, *p*=.001, partial η^2 =.40 (using a Greenhouse-Geisser correction), with performance decreasing as set size and mask duration independently increased. A significant interaction was found between set size and mask duration (F4,56) = 4.58, MSerror =.19, *p*= .003, partial η^2=0.25, reflecting the fact that mask duration had a progressively greater effect with increasing set size.[v]

Simple effects t-tests were conducted and revealed that OSM was produced even for a set size of one. There was a significant difference in performance between a mask duration of 0 and 60ms (t(14)=4.77, *p*<.001), a mask duration of 0 and 180ms (t(14)=7.39, *p*<.001 and a mask duration of 60 and 180ms (t(14)=2.85, *p*=.013 when the target was presented alone in the display. Further simple effects t-tests again revealed that set size affects performance even in the absence of OSM. There was a significant difference in performance between set size 1 and 6 (t(14)=3.66, *p*=.003) and 1 and 12 (t(14)=2.92, *p*=.011 at a mask duration of 0ms. However, there was no significant difference in performance between set size 6 and 12 (t(14)=.73, *p*=.476.

A further ANOVA analysis was performed on the criterion data (*C*). A two-way ANOVA was performed in the same manner as for the d' scores described above. The main effect of set size approached significance F(2,28)=3.17, MSerror=.14, *p*=.057, partial η^2 =.19), there was a clear main effect of mask duration, F(2,28)=36.83, MSerror=.16, *p*<.001, partial η^2 =.73); no interaction was found between the two factors, F(4,56)=1.05, MSerror=.04, *p*=.389). Thus these data show a tendency for observers to shift from a moderately conservative to a moderately liberal criterion as mask duration increases. As set size was increased a similar criterion shift towards a more liberal criterion was also observed but to a lesser extent.

Insert figure 4 about here

Discussion

An interaction between set size and mask duration was found for the detection task just as for the digit identification task. Importantly this interaction in detection accuracy could not be explained as the consequence of a ceiling effect nor of a response bias as it was found with a signal detection measure. Thus again the results are inconsistent with the findings reported in Argyropoulos et al. (2013).

The criterion data displayed an interesting pattern. It seems that the effect on the observer of increasing mask duration (and – to a lesser extent – of increasing set size) was to produce a criterion shift towards liberal responding. That is observers became increasingly likely to report a target as present – even on target absent trials – the longer the mask lingered on screen (or the more items were present in the display). A similar criterion shift was noted (but not formally analysed by Argyropoulos et al. in their line detection task). The existence of this criterion shift tells us nothing about its perceptual or cognitive basis (Pylyshyn, 1999). The effect could just reflect a conscious change in strategy under the conditions of uncertainty that masking produces towards reporting a target as being present. For instance, an observer may feel some uncertainty as to whether a digit they have glimpsed was present at the cued location, i.e. inside the four dots, or was at some other location, leading them to adopt a more lax criterion under masked conditions. This account is consistent with the fact that the criterion shift is more pronounced in conditions in which distractors were present (set size 6, set size 12), but has some difficulty explaining why this criterion shift effect, though diminished, is not entirely abolished in the set size 1 condition.

A more interesting possibility is that the change in responding reflects an underlying change in the perceptual experience of the observer associated with the stimulus manipulations. For instance the trailing mask might create the appearance of an object inside the mask location leading to more false positive responses. This seems to go against phenomenal descriptions of the OSM effect in which the area inside the four dots is said to be blank (e.g. Di Lollo et al., 2000); however, the phenomenal consequences of OSM have yet to be systematically investigated in any formal way (though see Koivisto, 2012 for one approach to measuring the subjective consequences of OSM using confidence ratings). Certainly the findings here suggest that this may be a profitable area of research and it is a question we have been pursuing in our laboratory.

However the response bias is an aside to the current paper. Let us return to the issue of the interaction and why this aspect of our data is for the second time inconsistent with that of Argyropoulos et al. (2013). One obvious difference with the experiments thus reported is the use of digits compared with the Landolt stimuli used in Argyropoulos et al. It is possible that this factor may underlie the difference in results. Digits are more heterogeneous as a stimulus class than the circle with a present or absent vertical bar or the Landolt stimuli used in Argyropoulos et al. (2013) and Filmer et al. (2014). Digits are also special in being, for most observers, heavily overlearned as a class of visual stimuli. Overlearned stimuli have sometimes been found to produce different patterns of results from other stimulus types on attentional tasks (e.g. Rotte, Heinze & Smid, 1997; Kawahara, Zuvic, Enns & Di Lollo, 2003; Martens, Korucuoglu, Smid & Neuwenstein, 2010; Kahan & Enns, 2014). It is therefore possible that the observed set size effect on masking has something to do with the attributes of digits as stimuli. To test this we therefore decided to employ the Landolt stimuli used in Argyropoulos et al. and FIlmer et al. but using the parameters of our first experiment.

Experiment 3: Landolt square discrimination

Experiment 3 was conducted to establish if there is something fundamentally different about the digit stimuli that have been used so far in this series of experiments. Experiment 3 was designed as a replication of Experiment 1, except for the fact that Landolt squares were used for identification in replacement of the digits. The Landolt squares were the same in height as the digits used in Experiment 1. Set size and mask duration were varied. Participants had to report the missing side of the target Landolt square defined in the array by the surrounding mask.

Method

Participants

Seventeen participants (17 female) from the Oxford Brookes Psychology student panel took part in the experiment. All gave informed consent and received course credits for completing the experiment; all reported normal (or corrected-to-normal) visual acuity.

Stimuli and procedure

The manner and procedure of the experiment were the same as Experiment 1 except for the stimuli being Landolt squares rather than digits. Participants were required to report which side of the target item had a missing segment using one of the four arrow keys on a conventional keyboard. The Landolt squares were .52° of visual angle in width/ height. The missing segment was .31 ° in size. The mask was the same as in the previous experiments. The stimuli were presented on a notional circle the same dimensions as in the previous experiments positioned around a fixation cross (see Figure 5). There were three set sizes (1, 6, 12 items) and three mask duration conditions (0, 60, 180 ms) factorially combined as in the previous experiments. Participants were given a demonstration and practice trials as in previous experiments. There were 540 experimental trials. The target gap position occurred equally often in each of the four cardinal positions in the experiment within each of the six factorially combined conditions. The gap position in the distractor stimuli was randomly determined for each stimulus.

Insert Figure 5 here

Results

Mean percent correct data for each combination of set size and mask duration were examined and are shown in Figure 6. The data were analysed using a two-way repeated measures ANOVA. The results showed a significant main effect of mask duration and set size respectively: $F(2,32) = 34.28$, $MSerror = 103.27$, $p<.001$, partial $\eta^2=.68$ and $F(2,32)= 55.67$, $MSerror=117.09$, $p<.001$, partial $\eta^2=.78$ (both using the Greenhouse-Geisser correction). Importantly, again there was a significant interaction between mask duration and set size: $F(4,64) = 4.78$, $MSerror = 30.61$, $p = .002$, partial $\eta^2=.23$. However it is arguable that the interaction we observe here reflects constraints in the measureable performance due to ceiling effects, rather than being a genuine interaction. Indeed one participant was at 100% in the 0 ms condition of set size 1. To test this, following the procedure adopted in Filmer et al. (2014a) in their Experiment 1 where a similar problem was encountered, we repeated the analysis of the interaction, but this time including only set sizes 6 and 12. The same pattern of significances was obtained including the set size × mask duration interaction $F(2,32) = 4.57$, $MSerror = 30.45$, $p = .018$, partial $\eta^2=.22$.

Simple effects t-tests were conducted and revealed that OSM was produced even for a set size of one. There were significant difference in performance between a mask duration of 0 and 60ms (t(16)=2.19, *p*=.044); 60 and 180ms (t(16)=2.60, *p*=.019 and 0 and 180ms (t(16)=4.44, *p*<.001). Further simple effects t-tests revealed that set size affects performance even in the absence of OSM. There was a significant difference in performance between set size 1 and 6 (t(16)=3.66, *p*=.002), 6 and 12 (t(16)=2.35, *p*=.032 and 1 and 12 (t(16)=4.21, *p*=.001 at a mask duration of 0ms.

These data, as with Experiment 1, were log transformed to give an additional test of the interaction. This showed a clear significant set size by mask duration interaction (F[4,64] = 4.20, MSerror =.003, *p* =.004, partial η^2=.21); this interaction remained significant with the log transformed scores even when only the 6 and 12 set size conditions were analysed, F(2,32) = 4.60, MSerror =.003, *p* =.017, partial η^2=.22.

Discussion

Thus, for the third time we have found an interaction between set size and mask duration in OSM. This time it was demonstrated with stimuli closely comparable to those in Argyropoulos et al (2013) and Filmer et al. (2014a), as well as those in the original Di Lollo et al. (2000) experiments.

Our findings therefore present us with an enigma. How can we account for the seemingly robust set size by mask duration interaction we observe across three different experiments when two previous series of studies (one of which came from our own lab), largely failed to produce this interaction. We shall return to this important question in the General Discussion. However the finding of a robust interaction prompted a further important question. Namely, what aspect of the set size manipulation could be responsible for affecting masking? Di Lollo et al. (2000) argued that set size influences OSM because it increases the time to contact of the focus of spatial attention with the target item. Yet it is unclear why set size should have this effect. In the standard OSM paradigm – that used in the current experiments – the target is defined by the surrounding mask, which is unique within the stimulus array. Thus the nature of the target in the OSM paradigm is unlike one in which the target would produce a serial search function (Treisman & Gelade, 1980). There are also empirical reasons to doubt that time to contact is the important variable. As we noted earlier Pilling et al. (2014) explicitly manipulated spatial attention using a cue. The cue was either presented simultaneously with the stimulus array or presented up to 150 ms before the onset of the array (meaning that attention was already pre-focused at the target location). In most cases this attentional manipulation had no discernible effect on OSM at all, while having a clear effect on overall accuracy; where in one of five experiments an effect was found on masking, it was modest at best. These facts suggest t*ime to contact* is a poor candidate explanation for our interaction.

Set size may influence masking in another way. The presence of distractors is known to increase internal noise within the visual system (Ekstein, 1998; Santhi & Reeves, 2004; Smith & Ratcliff, 2009; Magyar, Van den Berg & Ma, 2012); varying set size may vary internal noise and thus increase a target's vulnerability to OSM. However, there is a further possibility. Set size has always been confounded by the spatial proximity of distractors to the target. This confound is common across most studies in which set size is varied. It is also true of our Experiments 1-3. Increasing the proximity of distractors to a target can lead to what is described as *lateral masking* (Wolford & Chambers, 1983; Pöder, 2004) or, more commonly, *crowding* of the target (Korte, 1923; Pelli, 2008). In order to elucidate the nature of the set size effect we have repeatedly observed, it is therefore necessary to disentangle the effects of set size per se from those related to crowding. It was for this reason that Experiment 4 was conducted.

Experiment 4: Set size vs. crowding

Experiment 4 used the stimuli (digits) and task (digit identification) used in Experiment 1. Experiment 4 was comprised of two parts. In the first part (Experiment 4a) we demonstrate that crowding effects do occur when distractors flank the locations adjacent to the target (the target being defined, as in our previous experiments by four surrounding dots). Having established that crowding does occur in our displays in the next part (Experiment 4b) a series of conditions was given in which crowding and set sizes were independently manipulated. This was done under both masked and unmasked conditions.

Experiment 4a

Method

Participants

Ten participants (8 female) took part in the experiment. These were recruited from staff and students at Oxford Brookes University. All participants reported normal or corrected-to-normal visual acuity.

Stimuli and procedure

The stimulus array consisted of a target item and 3 distractors positioned on a virtual circle of the same dimensions and viewing conditions as Experiment 1. The target (identified by the surrounding four dots) was presented at a random position on the virtual circle with one distractor directly opposite. This basic arrangement was presented under crowded and uncrowded conditions, the difference between the two being the location of two further distractor items with respect to the target. On crowded trials the two distractors flanked the target on either side (target and distractor were separated by a circumferential distance of 1.22° visual angle between the centres of the respective stimuli; see in Figure 7 where the 5 and 3 flank the 7). On uncrowded trials it was the distractor opposite the target which was flanked by the two distractors and those positions adjacent to the target were left empty (see in Figure 7 where the 4 is flanked by the 3 and 7). This arrangement ensured that there was always symmetry across crowded and uncrowded trials in the stimulus array with the consequence that the distribution of spatial attention was likely to be comparable across the two condition types. Participants had to report the identity of the digit surrounded by the four dots. Unlike previous experiments mask duration was not varied: the four dots always offset with the stimulus array. A demonstration and practice trials were given before commencing the experiment. The experiment consisted of 120 crowded and 120 uncrowded trials.

Insert figure 7 about here

Results and Discussion

The mean percent correct performance for the uncrowded and crowded conditions is shown in Figure 8 (plate A). For reasons of consistency with the other Experiments we also present these as the mean of the log transformed scores (Figure 8, plate B). As can be seen when the target was closely flanked by distractors performance was substantially lower than when the distractor directly opposite the target was flanked, $t(9)=4.90$, $p<.001$. This fact demonstrates, as we suspected, that distractors produce crowding when located near the target under the conditions of our stimulus displays. Having established this fact Experiment 4b compared the effects of this crowding and distractor set size on OSM in a paradigm in which the two were independently varied.

Insert figure 8 about here

Experiment 4b

Experiment 4b followed the basic paradigm of Exp. 4a. Unlike Exp. 4a the number of items in the display was varied, as was mask duration. There were always a minimum of four items in the stimulus array (one target, three distractors). The stimulus arrays appeared in exactly the same manner as in Exp. 4a. With the larger set sizes additional distractors were added in unfilled locations while still adhering to the basic conditions of the smallest set size of 4. . Three set size conditions were given (4, 8, 12 items). With a set size of 8 or 12 the additional (4 or 8) distractors were positioned at empty locations on the circle.

If set size is itself the relevant factor in determining the effect of distractors on OSM then an interaction should be found between set size and mask duration as in the previous three experiments. If distractor proximity with respect to the target is the relevant variable in the previous experiments then an interaction should be found between this factor and mask duration. It was also recognised that the two predictions were not necessarily mutually exclusive: a further possibility is that set size *and* crowding could both exhibit interactions with masking.

Method

Participants

Thirty participants recruited from the Oxford Brookes Participant Panel (19 female) took part in the experiment. All gave informed consent. Participants received either £7 remuneration or course credits for completing the experiment. All participants reported normal or corrected-to-normal visual acuity.[vi]

Stimuli and procedure

The stimulus array consisted of a target item and 3, 7 or 11 distractors positioned on a virtual circle. The arrangement of the distractors depended on the trial type. On all trials the target was presented at a random position on the virtual circle with one distractor directly opposite. On crowded trials the target was flanked on either side (a circumferential distance of 1.22 °); on crowded trials the locations adjacent to this distractor were always empty (see Figure 9 where the target digit 7 is flanked by the digits 5 and 3 in all set size conditions); on uncrowded trials this distractor opposite the target was itself flanked by two distractors and the positions adjacent to the target were empty (see Figure 9 where the target is the digit 4 located opposite the above mentioned digit 7 in all set size conditions). With a set size of 8 or 12 items, the additional 4 or 8 distractors were presented at unoccupied locations on the virtual circle. This placement was done with two constraints. Firstly there was always a minimum circumferential distance of 1.22 ° between each additional distractor on the virtual circle. Secondly, on uncrowded trials, there was a minimum circumferential distance of 3.66 ° between the additional distractors and the target; on crowded trials there was always a minimum circumferential distance of 3.66 ° between the additional distractors and the distractor positioned opposite the target. Due to the added crowding conditions the number of mask duration conditions was reduced to two: 0ms (simultaneous mask offset) and 180ms (delayed mask offset). There were 600 randomly ordered trials, 50 for each of the twelve factorially combined conditions of crowding, set size and mask duration. The experiment was conducted in 10 blocks and the computer prompted the participant to take a brief break after each 60 trial increment. A demonstration and practice trials were given, as previously described.

Insert Figure 9 about here

Results

One participant had to be excluded from the analysis as their overall performance was at chance. Mean percent correct data for each combination of set size and mask duration were examined and are shown in Figure 10.

A 3-way repeated measures ANOVA was conducted on the data from the remaining participants. This revealed a significant main effect of all three factors: crowding ($F[1,28] = 176.61$., MSerror =106.72, $p<.001$, partial $\eta^2 = .86$), mask duration ($F[1,28] = 91.90$, MSerror =89.56, $p<.001$, partial $\eta^2 = .7$), and set size ($F[2,6] = 18.03$, MSerror = 30.98, $p<.001$, partial $\eta^2 = .39$). Thus, crowding, set size and mask duration all independently influenced target perceptibility. The interaction between crowding and mask duration was significant ($F[1,28]=5.70$, MSerror =41.46, $p=.024$, partial $\eta^2 = .17$). The interaction between set size and mask duration was non-significant $F[2,56]= .59$, MSerror =36.07, $p=.592$, as was that between crowding and set size, $F[2,56]= .53$, MSerror =35.49, p=.591. The 3 way interaction (set size × crowding ×mask duration) was also non-significant ($F[2,56]= .50$, MSerror =29.22, $p=.610$).

With the log transformed data the crowding by mask duration interaction was even more pronounced ($F[1,28]= 9.26$, MSerror <.01, $p=.005$, partial $\eta^2 = .25$), while the set size by mask duration interaction remained non-significant; $F[2,56]= .50$, MSerror <.01, $p=.607$. As with the untransformed scores none of the other interaction terms approached significance. Observation of

Figure 8 shows that the significant crowding by mask duration interaction reflects the fact that OSM was stronger when the target was crowded by distractors compared to when it was not.

Insert figure 10 about here

Discussion

Experiment 4 shows that set size itself does not influence OSM. Masking – as indexed by the effect of mask duration – was of similar magnitude irrespective of whether there were four, eight, or twelve display items. At the same time set size did affect overall performance: accuracy declined significantly with the number of distractors present. The crowding manipulation, by contrast, did influence OSM. OSM was augmented when the target was flanked by distractors compared to when the adjacent locations were left empty. Thus, Experiment 4 suggests that the ostensible set size effects on OSM reported for Experiments 1- 3 were not actually due to set size itself (i.e. the number of distractor items in the display).

General Discussion

Argyropoulos et al. (2013) and Filmer et al. (2014a) both reported – against the original claim of Di Lollo et al. (2000) – that distractors do not influence OSM. They showed, firstly, that OSM occurs even when the target is alone in the stimulus array (see also Filmer et al., 2014b). They secondly reported that where distractors were added to the stimulus array in varying numbers this did not influence OSM. Taken together our experiments give unequivocal support to the first of these claims, but lend only qualified support to the second. Experiments 1-3 show that distractors did influence the magnitude of OSM. Experiment 4 showed the distractor effect on OSM was spatially dependent, i.e. dependent on whether or not distractors were present at locations near to the target. Set size itself did not influence OSM, though it did influence overall target perceptibility.

The findings in Experiment 1 to 3 can be interpreted in light of the results of Experiment 4. We can assume that the interactions between set size and mask duration found in the first three experiments were actually driven by distractor proximity to the target– which in the first three experiments was conflated with set size – not by set size itself.

Interestingly Di Lollo et al. did note the possibility of crowding occurring in the OSM paradigm. However they argued that distractor crowding would only have a main effect on overall perceptibility, and that this effect itself would be modest in size when compared with that of set size; only set size was viewed to be likely to interact with mask duration (Di Lollo et al. 2000, see pp 488). Thus what was observed in Experiment 4 was exactly the opposite of the predictions regarding the relative effects of crowding and set size: It was set size which was modest in overall effect and only crowding which interacted with OSM.

Crowding and OSM

Why is it that crowding and OSM interact? Crowding has similarities with masking as a phenomenon: both have deleterious effects for stimulus perceptibility. Consistent with this, previous work has documented crowding effects on classical forms of masking (spatial frequency masking, pattern masking, metacontrast), (Chung, Levi & Legge, 2001; Huckauf & Heller, 2004; Vickery, Shim, Chakravarthi & Luedeman, 2009). How then do we interpret the interaction between crowding and OSM? Crowding processes have been argued to occur earlier within the visual processing cascade than OSM, but occur later than metacontrast and pattern masking (Chakravarthi & Cavanagh, 2009; see Breitmeyer, 2014). Chakravarthi & Cavanagh (2009) demonstrated this using a modified crowding paradigm in which a single unmasked target was flanked by four crowding items present at cardinal positions around it. Crowding items were backwards masked either through a pattern, metacontrast, or a four dot (OSM) mask. Crucially only the pattern and metacontrast mask reduced crowding; the four dot mask had no effect. This failure of the four dot masks to reduce crowding by the flankers was argued to show that crowding must precede OSM.

One widely held interpretation of crowding is that it occurs as a consequence of the *spatial integration* or *pooling* of the target and flanker features within peripheral receptive fields of the visual cortical areas (Levi, 2008; Parkes, Lund, Solomon & Morgan, 2001; Pelli & Tillman, 2008; Greenwood, Bex, & Dakin, 2009; Dakin, Cass , Greenwood, & Bex, 2010). The consequences of this spatial integration process are that target and flanker percepts are merged, resulting in a 'jumbled' (Greenwood, Bex & Dakin, 2010; Anderson, Dakin, & Schwarzkopf, Rees & Greenwood, 2012; Kahan & Enns, 2014; Kahan & Enns, 2010), 'smudged' (Korte, 1927; Tyler & Likova, 2007) or otherwise degraded target appearance. Pooling models assume crowding interactions to be feed-forward in nature. However some aspects of crowding are difficult to incorporate within a strict feedforward model: In particular crowding is affected by non-local factors such as grouping and global configuration with other display items (Banks, Larson & Prinzmetal; 1979; Livne, D. Sagi, 2010; Manassi, Sayim & Herzog, 2012, 2013). Based on this and other findings recent work has attempted to incorporate crowding–as with OSM – within re-entrant architectures (Jehee, Roelfsema, Deco, Murre, & Lamme, 2007; Foley, Grossberg & Mingolla, 2012; Herzog & Manassi, 2015).

To recap, in the OSTM model (Di Lollo et al. 2000), masking occurs as a consequence of the visual system's failure to achieve correspondence between the re-entrant signal containing a (hypothesised) representation of the target and the lower level representation of the current activation pattern caused by the trailing mask alone. Crowding may affect masking by increasing the number of iterative cycles required for a successful match to be achieved between the re-entrant signal and input layer representation. When the target location is crowded the quality of the input representation is likely to be degraded. As such a successful match is less likely to occur between the descending re-entrant signal and input representation within the brief time that the target is present. Under such circumstances the target would have an increased vulnerability to substitution by the trailing mask.

The current results can also be interpreted within a purely feedforward model of object substitution. Pöder (2013, 2014), for instance, argues that OSM occurs as a consequence of temporal attentional gating: the trailing mask being selected along with the target and adding noise to the target representation, reducing the target's *signal-to-noise ratio* (SNR). If the flanking distractors are pooled together with the target in the manner earlier described then these could act as an additional source of noise, further reducing the target SNR. On this account crowding and the trailing mask can be considered as respectively spatial and temporal sources of perceptual noise that serve to reduce the target SNR. If these two sources of internal perceptual noise combine multiplicatively it would produce the interaction we observed.

Comparison to previous set size OSM experiments

A question we posed earlier regarded why Argyropoulos et al. (2013) and Filmer et al. (2014) repeatedly failed to produce the 'set size' effects on masking we found to be so robust across our Experiments 1-3. Let us first compare our displays with those of Argyropoulos et al. One difference between Argyropoulos et al. and the current study was in the eccentricity of the stimulus array. In most of the experiments in Argyropoulos et al., stimuli were positioned 2.9° radially from fixation; in their Experiment 5, stimuli were positioned only 1.8° from fixation. In all four of our experiments stimuli were 3.9° degrees from fixation. Thus the stimuli in our experiment were presented more peripherally. As crowding effects scale with eccentricity (Bouma, 1970; Levi, 2008), crowding of the target was likely greater in the displays of the current experiments than in those in Argyropoulos et al.[vii]

Filmer et al. (2014a) also failed to find a significant effect of set size on OSM. Interestingly, when one scrutinises the data presented in Filmer et al. it can be seen that in most cases their data show trends towards masking increasing with set size, this was particularly evidence when their accuracy data were log transformed (see esp. their Fig.3) though these trends were non-significant. Another issue regarding the Filmer et al. study, is the fact their stimulus array presentation times were atypically long: the stimulus array was presented for 100 ms ,compared for instance to 10-45ms in Di Lollo et al (2000), 50 ms in Argyropoulos et al. (2014), and 10-40 ms in the current experiments. OSM diminishes as target duration increases (Gellatly et al., 2010; Guest et al., 2012); the rather weak main effects of OSM observed in this study (e.g. in Filmer et al.'s Exp. 3 their OSM effect was only marginally significant at p=.08 following guessing correction and removal of participants performing at chance) may be a consequence of this. The fact that OSM was weak may have reduced the possibility of observing an interaction with the set size variable. Thus it seems our findings are less contradictory to the results of these papers than might appear on first reading.

Set size and target perceptibility

Manipulation of set size had no detectable effect on masking when crowding was controlled for.[viii] It did have an overall main effect on target perceptibility, though this was somewhat less than that of crowding. One might ask why set size should have any effect at all. As stated earlier *time to contact* seems an unlikely relevant intermediate variable given the distinctiveness afforded to the target by the presence of the surrounding mask. It is more likely that the additional distractors mode of interference occurs by increasing noise in the visual system. This claim has been made by a number of authors based largely on evidence from the visual search paradigm and formalised in signal detection models (Palmer, Ames & Lindsey, 1993; Eckstein, Thomas, Palmer & Shimozaki, 2000; Baldassi & Burr, 2000; Davis, Shikano, Peterson & Keyes Michel, 2003; Cameron, Tai, Eckstein & Carrasco, 2004). Importantly some of this work shows that set size effects persist – albeit in diminished form – even when attention is focused on the target location (Cameron et al., 2004), demonstrating, as in our data, that distractors are processed even when task irrelevant and outside of attentional focus.

That set size influences performance in the absence of any influence on OSM indicates something about the locus of influence of this variable within the visual processing hierarchy. According to additive factors logic (Sternberg 1969, 1998) set size must be influencing a processing stage discrete from that in which OSM occurs. One possibility is that the presence of non-local distractors (as is the case when set size is decoupled from crowding) influences *perceptual decision processes* (Shaw, 1980; Palmer et al., 1993) concerning the target. Perceptual decisions will necessarily occur after the completion of the object formation/consolidation processes with which OSM is associated. The obligatory processing of the distractors may lead to competition between the target and non-local distractor items for decision processes. The more of these items there are the greater this competition, and mutual interference, will be. This source of interference would reduce accuracy in identifying the target but this would be proportionally the same whether the trial was masked or unmasked.

The role of distractors in OSM

Experiment 4 is the first study to demonstrate that crowding modulates OSM. Given what is known about crowding we can speculate what other factors regarding the relationship between target and flankers are likely to be relevant. We describe some of these below.

Crowding occurs when distractors flank a target's location. Target- flanker distance is obviously critical; crowding diminishes with distance and is abolished outside of the crowding window. A number of factors have been identified in determining the size of this window and the amount of crowding that will result. A major factor is *eccentricity* (Bouma, 1970; Pelli & Tilman, 2008). With foveated targets crowding is largely absent; as a target is moved into the visual periphery, as noted earlier, crowding effects become more pronounced and the critical distances proportionally larger (Bouma, 1970). We would predict that our observed crowding × OSM interaction will exhibit these same characteristics: flankers being minimally effective on masking with a foveated target (Dux, Visser, Goodhew, & Lipp, 2010; Filmer et al. 2014b) and having an increasing effect as target eccentricity is increased. Interestingly, the crowding window is not regular in shape but displays a *radial-tangential anisotropy*. Crowding is markedly stronger – and the critical spacing wider – when flankers are positioned radially with respect to fixation compared to when they are tangentially located (Toet & Levi, 1992). In the current experiments the target and distractors were circumferentially arranged, as they are in most OSM studies. They were therefore chiefly tangential with respect to the central fixation, suggesting the crowding would be more greatly marked with radially positioned distractors. With such a stimulus arrangement the modulation of OSM may be even more pronounced than we observed.

Crowding is also sensitive to the visual similarities of stimuli in terms of the similarity between the target and flankers (t*arget-flanker similarity*), and between the flankers and other display items (*flanker grouping*). That is, crowding tends to be weakened when flankers are visually distinct from the target, for instance because they differ in colour, shape, size, or orientation (Kooi, Toet, Tripathy, & Levi, 1994; Hariharan, Levi & Klein, 2005; Chung, Levi, & Legge, 2001; Bernard &

Chung, 2011). Crowding is also reduced if flankers form a homogenous perceptual group with other distractors in a display (Manassi, et al., 2013). Together, these phenomena show that when a target tends to ‘pop out’ from the rest of the display items crowding is suppressed. Interestingly, OSM has also been reported to be affected by target ‘pop out’. With a target which is visually distinct from the distractors OSM is reduced when the target does ‘pop out’ from other distractors within the display (Di Lollo et al., 2000; Tata, 2002; Gellatly et al. 2006; Pilling et al., 2014). This pop out effect in OSM has tended to be attributed to attentional factors (Di Lollo et al. 2000; Tata, 2002) however it may be that these target ‘pop out’ effects in OSM are, wholly or in part, a consequence of the release from crowding of adjacent flanking distractors which would ensue under these conditions. This and the other suggestions discussed above await further research.

Conclusion

Our experiments demonstrate that the presence of distractors can influence the level of OSM that is observed. However the nature of this influence is rather different to that originally theorised by Di Lollo and colleagues (Di Lollo et al., 2000). What seems critical for distractors to have an effect is for them to crowd the target; under such conditions the OSM effect is amplified. Thus, our work stands with a number of other observations of interactions between crowding and masking (e.g. Huckauf & Heller, 2004; Vickery et al., 2009). However this is the first time such interactions have been demonstrated in OSM. The effect of crowding cannot be attributed to the reduced performance which results from crowding: When set size was varied this also affected performance but did so without impacting on masking. This suggests that there is something about crowding and its neurocognitive underpinnings which lead to it affecting OSM in the manner shown.

In Experiment 4 masking (as indexed by the difference between the 0 ms and 180 ms duration conditions) was increased by about 40% in the crowded, compared to the uncrowded, condition. However we suspect that these effects may be even greater under conditions, described earlier, which increase crowding effects, e.g. increased target eccentricity combined with radially positioned flanking distractors. In any case the fact that crowding and OSM interact at all indicates that the two share common mechanisms. Further exploration of the interaction between crowding and OSM may therefore be fruitful in advancing our understanding of these visual phenomena and, by extension, of mid-level vision itself.

References

Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America, 43,* 648-657.

Anderson, E. J., Dakin, S. C., Schwarzkopf, D. S., Rees, G., & Greenwood, J. A. (2012). The neural correlates of crowding-induced changes in appearance. *Current Biology*, *22*(13), 1199-1206.

Argyropoulos, I., Gellatly, A., Pilling, M., & Carter, W. (2013). Set size and mask duration do not interaction object-substitution masking. *Journal of Experimental Psychology Human Perception and Performance,* 39, 646-661.

Bachmann, T. (2006). Microgenesis of Perception: Conceptual, Psychophysical, and Neurobiological Aspects. In Öğmen, Haluk (Ed);and Breitmeyer, Bruno G. (Eds.), (2006). The first half second: *The microgenesis and temporal dynamics of unconscious and conscious visual processes* (pp. 11-33). Cambridge:MA. MIT Press.

Baldassi, S., & Burr, D. C. (2004). "Pop-out" of targets modulated in luminance or colour: the effect of intrinsic and extrinsic uncertainty. *Vision Research*,*44*(12), 1227-1233.

Banks, W. P., Larson, D. W., & Prinzmetal, W. (1979). Asymmetry of visual interference. *Perception & Psychophysics*, *25*(6), 447-456.

Bernard, J. B., & Chung, S. T. (2011). The dependence of crowding on flanker complexity and target–flanker similarity. *Journal of vision*, *11*(8), 1.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*, 177-178.

Breitmeyer, B. G. (2014). Contributions of magno- and parvocellular channels to conscious and non-conscious vision. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 269,* 20130213.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. *Psychological Review, 83,* 1-36.

Breitmeyer, B. G. & Ogman, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception and Psychophysics, 62,* 1572-1595.

Bridgeman, B. (2006). Contributions of lateral inhibition to object substitution masking and attention. *Vision Research, 46,* 4075-4082.

Cameron, E. L., Tai, J. C., Eckstein, M. P., & Carrasco, M. (2004). Signal detection theory applied to three visual search tasks-identification, yes/no detection and localization. *Spatial vision*, *17*(4), 295-326.

Carlson, T. A., Rauschenberger, R., & Verstraten, F. A. (2007). No representation without awareness in the lateral occipital cortex. *Psychological Science*, *18*(4), 298-302.

Chakravarthi, R., & Cavanagh, P. (2009). Recovery of a crowded object by masking the flankers: Determining the locus of feature integration. *Journal of Vision, 9,* 1-9.

Chung, S. T. L., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research, 41,* 1833-1850.

Coltheart, M. (1975). Iconic memory: A reply to Professor Holding. *Memory & Cognition, 3,* 42-48.

Dakin, S. C., Cass, J., Greenwood, J. A., & Bex, P. J. (2010). Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *Journal of Vision*, *10*(10), 14.

Davis, E. T., Shikano, T., Peterson, S. A., & Keyes Michel, R. (2003). Divided attention and visual search for simple versus complex features. *Vision research*, *43*(21), 2213-2232.

Dell'Acqua, R., Pascali, A., Jolicoeur, P. & Sessa, P. (2003). Four-dot masking produces the attentional blink. *Vision Research, 43,* 1907-1913.

Di Lollo, V. (2014). Reentrant processing mediates object substitution masking: comment on Põder (2013). *Frontiers in Psychology, 5,* 1-5.

Di Lollo, V., Enns, J. T. & Rensink, R. A. (2002). Object substitution without re-entry? *Journal of Experimental Psychology: General, 131,* 594-596.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reenrant visual processing. *Journal of Experimental Psychology: General, 129,* 481-507.

Dux, P. E., Visser, T. A. W., Goodhew, S. C., & Lipp, O. V. (2010). Delayed reentrant processing impairs visual awareness: An object substitution masking study. *Psychological Science*, *21*, 1242-1247

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & psychophysics*, *62*(3), 425-451.

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research, 44,* 1321-1331.

Enns, J. T. & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4,* 345-352.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8,* 135-139.

Enns, J. T., Lleras, A., & Di Lollo, V. (2006). Three aspects of visual masking are informed by reentrant modeling: object substitution, feature migration, and response priming. In Öğmen, H. and

Breitmeyer, B. G. (Eds.). *The first half second: The microgenesis and temporal dynamics of unconscious and conscious visual processes* (pp. 127-147). Cambridge:MA. MIT Press.

Enns, J. T., Lleras, A., & Moore, C. M. (2010). Object updating: A force for perceptual continuity and scene stability in human vision. In Nijhawan, R. (Ed.), *Problems of space and time in perception and action* (pp. 503–520). Cambridge, MA: Cambridge University Press.

Felleman, D. J. & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1, 1-47.*

Filmer, H., Mattingley, J., & Dux, P. (2014a). Size (mostly) doesn't matter: the role of set size in object substitution masking. *Attention, Perception & Psychophysics, 76,* 1620-1629.

Filmer, H. L., Mattingley, J. B., & Dux, P. E. (2014b). Object substitution masking for an attended and foveated target. *Journal of Experimental Psychology: Human Perception and Performance.* Advance online publication.

Foley, N. C., Grossberg, S., & Mingolla, E. (2012). Neural dynamics of object-based multifocal visual spatial attention and priming: Object cueing, useful-field-of-view, and crowding. *Cognitive Psychology*, *65*(1), 77-117.

Francis, G., & Hermens, F. (2002). Comment on "competition for conscious among visual events: The psychophysics of re-entrant visual processes" (Di Lollo, Enns, & Rensink, 2000). *Journal of Experimental Psychology: General, 131,* 590-593.

Gellatly, A., Pilling, M., Carter, W., & Guest, D. (2010). How does target duration affect object substitution masking?. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1267

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.

Goodhew, S. C., Dux, P. E., Lipp, O. V., & Visser, T. A. (2012). Understanding recovery from object substitution masking. *Cognition*, *122*(3), 405-415.

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,* 102-110.

Goodhew, S. C., Visser, T. A., Lipp, O. V., & Dux, P. E. (2011). Competing for consciousness: Prolonged mask exposure reduces object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 588-596.

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 the United States of America, 106,* 13130-13135.

Guest, D., Gellatly, A. & Pilling, M. (2011). The effect of spatial competition between object-level representations of target and mask on Object Substitution Masking. *Attention, Perception & Psychophysics, 73,* 2528-41.

Hariharan, S., Levi, D. M., & Klein, S. A. (2005). “Crowding” in normal and amblyopic vision assessed with Gaussian and Gabor C’s. *Vision Research*,*45*(5), 617-633.

Harris, J. A., Ku, S., & Woldorff, M. G. (2013). Neural processing stages during object-substitution masking and their relationship to perceptual awareness.*Neuropsychologia*, *51*(10), 1907-1917.

Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: a fresh look at object recognition. *Current Opinion in Behavioral Sciences*, *1*, 86-93.

Huckauf, A. & Heller, D. (2004). On the relations between crowding and visual masking. *Perception & Psychophysics, 66,* 584-595.

Jannati, A., Spalek, T. M., & Di Lollo, V. (2013). A novel paradigm reveals the role of re-entrant visual processes in object substitution masking. *Attention, Perception & Psychophysics, 75,* 1118-1127.

Jehee, J. F., Roelfsema, P. R., Deco, G., Murre, J. M., & Lamme, V. A. (2007). Interactions between higher and lower visual areas improve shape selectivity of higher level neurons—Explaining crowding phenomena. *Brain Research*, *1157*, 167-176.

Jiang, Y., & Chun, M. M. (2001). Asymmetric object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, *27,* 895-918.

Kahan, T. A., & Enns, J. T. (2014). Long-Term Memory Representations Influence Perception Before Edges are Assigned to Objects. *Journal of Experimental Psychology: General*, *143*, 566-574.

Kahan, T. A., & Enns, J. T. (2010). Object trimming: When masking dots alter rather than replace target representations. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 88-102.

Kahan, T. A., & Mathis, K. M. (2002). Gestalt grouping and common onset masking. *Perception and Psychophysics*, *64*, 1248-1259

Kahneman, D. (1968). Method, finding, and theory in studies of visual masking. Psychological Bulletin, 70, 404-425.

Kawahara, J., Zuvic, S. M., Enns, J. T. & Di Lollo, V. (2003). Task switching mediates the attentional blink even without backward masking. *Perception & Psychophysics, 65,* 339-351.

Kinsbourne, M. & Warrington, E. K. (1962). A disorder of simultaneous form perception. *Brain, 85,* 47-66.

Koivisto, M. (2012). Is reentry critical for visual awareness of object presence?. *Vision research*, *63*, 43-49.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial vision*,*8*(2), 255-279.

Korte, W. (1923). Uber die Gestaltauffassung im indirekten Sehen. *Zeitschrift für Psychologie, 93*, 17–82.

Kotsoni E, Csibra G, Mareschal D, et al. (2007). Electrophysiological correlates of common-onset visual masking. *Neuropsychologia*, 45(10), 2285-2293.

Lamme, V. A., Super, H. & Spekreijse, H. (1998). Feedforward horizontal and feedback processing in the visual cortex. *Current Opinion in Neurobiology, 8,* 529-535.

Lefton, L. A. (1973). Metacontrast: A review. *Perception & Psychophysics, 13,* 161-171.

Levi, D. M. (2008). Crowding- an essential bottleneck for object recognition: a mini-review. *Vision Research, 48*, 635-654.

Lim, S. W. H., & Chua, F. K. (2008). Object substitution masking: When does mask preview work?. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1108.

Livne, T., & Sagi, D. (2010). How do flankers' relations affect crowding? *Journal of Vision*, *10*(3), 1-14.

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.

Luiga, I., & Bachmann, T. (2008). Luminance processing in object substitution masking. *Vision Research*, 48(7), 937-945.

Magyar, H., Van den Berg, R. & Ma, W. J. (2012). Does precision decrease with set size? *Journal of Vision, 12,* 1-16.

Manassi, M., Sayim, B., & Herzog, M. H. (2013). When crowding of crowding leads to uncrowding.*Journal of Vision*, 1*3*(13), 10

Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding.*Journal of Vision*, *12*(10), 13.

Martens, S., Korucuoglu, O., Smid, H. G. O. M. & Nieuwenstein, M. R. (2010). Quick minds slowed down: Effects of rotation and stimulus category on the attentional blink. *PLoS ONE, 5*, 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.

Mumford, D. (1992). On the computational architecture of the neo-cortex: II. The role of the cortico-cortical loops. *Biological Cybernetics, 66,* 241-251.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(1), 108

Parkes, L., Lund, J., Angelucci, A. Solomon, J. A. & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4,* 739-744*.*

Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in Neurobiology, 18,* 445-451.

Pelli, D. G. & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11,* 1129-1135.

Pilling, M., Gellatly, A., Argyropoulos, Y., & Skarratt, P. (2014). Exogenous spatial precuing reliably modulates object processing but not object substitution masking. *Attention, Perception and Psychophysics, 76,* 1560-1576.

Pilling, M., & Gellatly, A. (2010). Object substitution masking and the object updating hypothesis. *Psychonomic bulletin & review*, *17*(5), 737-742.

Põder, E (2014). The changing picture of object substitution masking: reply to Di Lollo (2014). *Frontiers in Psychology, 5,* 1004.

Põder, E. (2013). Attentional gating models of object substitution masking. *Journal of Experimental Psychology: General, 142,* 1130-1141.

Põder, E. (2004). Effects of set-size and lateral masking in visual search. *Spatial Vision, 17,* 257-268.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*, 341–423.

Rotte, M., Heinze, H. J. & Smid, H. G. O. M. (1997). Selective attention to conjunctions of color and shape of alphanumeric versus non-alphanumeric stimuli: A comparative electrophysiological study. *Biological Psychology, 46,* 199-221.

Santhi, N. & Reeves, A. (2004). The roles of distractor noise and target certainty in search: A signal detection theory: *Vision Research, 44,* 1235-1256.

Schweickert, R. (1985). Separable effects of factors on speed and accuracy: Memory scanning, lexical decision, and choice tasks. *Psychological Bulletin*, *97*, 530-546.

Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. *Attention and performance VIII*, *8*, 277-295.

Sibly, M. (2011). *BlitzMax* [Computer software]. Blitz Research Ltd. Auckland, New Zealand.

Smith, P. L. & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review, 116,* 283-317.

Snowden, R. J., & Hammett, S. T. (1998). The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination. *Vision research*, *38*, 1935-1945.

Sternberg, S. (1969) The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II. Acta Psychologica*, 30, 276-315.

Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science: Vol. 4. Methods, models, and conceptual issues* (pp.703-863). Cambridge, MA: MIT Press.

Tata, M. S. (2002). Attend to it now or lose it forever: Selective attention, metacontrast masking, and object substitution. *Perception & Psychophysics, 64,* 1028-1038.

Tata, M. S., & Giaschi, D. E. (2004). Warning: Attending to a mask may be hazardous to your perception. *Psychonomic bulletin & review*, *11*), 262-268.

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision research*, *32*(7), 1349-1357

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 417-425.

Tripathy, S.P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research*, *42*, 2357-2369.

Turvey, M. T. (1973). On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 81,* 1-52.

Tyler, C. W., & Likova, L. T. (2007). Crowding: A neuroanalytic approach.*Journal of vision*, *7*(2), 16.

Vickery, T. J., Shim, W, M., Chakravarthi, C., Jiang, Y. V., & Luedeman, R. (2009). Supercrowding: Weakly masking a target expands the range of crowding. *Journal of Vision, 9,* 1-5.

Vroomen, J. & Keetels, M. (2008). Sounds change four-dot masking. *Acta Psychologica, 130,* 58-63.

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.

Werner, H. (1957). *Comparative psychology of mental development*. (Rev. ed.). New York:International Universities Press.

Wolford, G. & Chambers, L. (1983). Lateral masking as a function of spacing. *Perception & Psychophysics, 33,* 129-138.

Footnotes

---

[i] Though the four dot mask is the most commonly used type of mask, it is unnecessary for the phenomenon to occur; a mask comprising of a ring or square surrounding the target, similar to those in MM, will also produce OSM (Di Lollo et al. 2000; Tata & Giaschi, 2004; Gellatly, Pilling, Cole & Skarratt, 2006). The value of the four-dot mask is that it experimentally isolates OSM processes from those associated with metacontrast and other local inhibitory processes, such as surround suppression (Snowden & Hammett, 1997).

[ii] This object updating account has received wide empirical support in the fact that manipulations which influence the tendency for target and mask to be perceived as the same or different perceptual objects (e.g. varying whether mask and target are the same or different colours, varying whether mask and target have a common motion path or whether mask and target onset in the same temporal window) influences the amount of OSM which results (Lleras & Moore, 2003; Moore & Lleras, 2005; Gellatly, Pilling, Cole & Skarret, 2006; Gellatly, Pilling, Carter, & Guest, 2010; Guest, Gellatly, & Pilling, 2010; Pilling & Gellatly, 2010; Luiga & Bachmann, 2007; Tata & Giashi, 2004; Lim, & Chua, 2008; Goodhew, et al., 2015).

[iii] This work was presented at ECVP (2013), Bremen, Germany (see Pilling M, 2013, Target and mask preview effects in object substitution masking. *Perception, 42 ECVP Abstract Supplement, page 26).* The work will form part of a forthcoming paper on target and mask preview effects in OSM.

[iv] For the criterion measure *C* zero indicates no bias, positive values a conservative bias (i.e. here a tendency to respond 'absent' under conditions of uncertainty), negative values a liberal bias (a tendency to respond 'present' under conditions of uncertainty).

[v] d' is not a probability measure and is therefore exempt from the requirement to perform log transformation before ANOVA analysis of interactions.

[vi] It should be noted that the number of participants recruited for this study was larger than in previous experiments. This was done, given the complexity of the experiment, to ensure that there was enough statistical power to reveal interactions under these conditions. Experiment 4b was more complicated than previous experiments in two main ways. The first is that three independent variables were being manipulated; each with at least the potential to interact with the other variables. It was deemed that the larger sample size would give an opportunity for any potential interaction effects. Secondly, and more importantly, it was hoped that the use of an increased sample size would give the opportunity for a three way interaction between crowding, set size and masking to reveal itself in our analysis. It was accepted that crowding and set size may not be mutually exclusive in their influence on OSM and as such may result in a greater OSM effect when combined.

[vii] The stimuli in Argyropoulos et al. were also only about half the size of those in the current study in terms of the subtended visual angle. However this is unlikely to be a factor as far as crowding is concerned. In the visual periphery at least crowding seems be largely size invariant, dependent more on the centre-to-centre stimulus spacing than the nearest edges (Tripathy & Cavanagh, 2002).

[viii] It could be argued that the failure to find an interaction between set size and mask duration in Exp. 4b reflects the fact that set size was not varied to the same extent as it was in previous experiments in this paper. In Exp. 1-3 set size was varied between a minimum of 1 item (target alone) and a maximum of 12 items (target + 11 distractors). In Exp. 4a however set size was varied between a minimum of 4 and and 12 items. This larger minimum set size condition is a necessary consequence of having to vary crowding independently of set size (since one cannot have crowding without the presence of distractors and, as we argued earlier in the paper, in order to balance attention across the display it is necessary to have a design in which a distractor is always opposite to the target). Could this restricted set size range explain our failure to observe an interaction between set size and mask duration? We think not: though the set size variable was more restricted it still showed a substantive main effect on performance suggesting that even with this curtailed range there was still ample opportunity for an interaction with mask duration to reveal itself. We think that this shows that set size does not interact with mask duration when this factor is isolated from crowding. We therefore cannot rule out

the possibility that set size could interact with mask duration where a larger set set size range is given; however from the data we have there is no evidence to suspect that this would be the case.

Figure headings

Figure 1. A schematic depiction of the trial sequence in Experiment 1 (digit identification task).

Figure 2. Performance in Experiment 1 (digit identification task). Accuracy (% correct; plate A) and transformed accuracy (Log10; plate B) are shown for the three set sizes (1, 6, 12) by each mask duration condition (0, 60, 180 ms.). Error bars represent +/- 1 standard error.

Figure 3. A schematic depiction of the trial sequence in Experiment 2 (digit detection task). Figure shows example of a target absent trial. On target present trials a digit (0-9) would be present inside the four dot mask, as in Figure 1.

Figure 4. Performance in Experiment 2 (digit detection task). Proportion of hits (*p[Hit]*), proportion of false alarms (*p[FA]*), d-prime (d') and response bias (*C*) are shown respectively in plates A, B, C and D. Error bars represent +/- 1 standard error.

Figure 5. A schematic depiction of the trial sequence in Experiment 3 (Landolt square discrimination task).

Figure 6. Performance in Experiment 3 (Landolt square discrimination task). Accuracy (% correct; plate A) and transformed accuracy (Log10; plate B) are shown for the three set sizes (1, 6, 12) by each mask duration condition (0, 60, 180 ms.). Errors bars represent +/- 1standard error.

Figure 7. A schematic depiction of the trial sequence in Experiment 4a. An example of the stimulus array for crowded and uncrowded trials can be seen respectively in the left and right frames of the stimulus array. In the left frame note that the target (a ‘7’) is flanked by two distractor digits while the opposite distractor (‘4’) is unflanked. In the right frame the target (‘7’) is unflanked while the opposite distractor (‘4’) is flanked on either side by two distractor digits.

Figure 8. Performance in Experiment 4a (uncrowed vs. crowded). This is shown as the mean percentage correct scores (plate A) and as the mean Log10 transformed scores (plate B). Error bars represent +/- 1 standard error.

Figure 9. A schematic depiction of the trial sequence in Experiment 4b (Set size vs. crowding). The three stimulus array frames depict examples of trials for (from left to right) set size 4, set size 8, and set size 12. As in Fig. 7 in the example frames the target digit is a ‘7’ and the opposite distractor is a 4. In all the examples the target is crowded by two flanking distractors (‘5’, ‘3’); on uncrowded trials the locations occupied by the flankers would be empty and the flankers would surround the opposite distractor (in this case the ‘4’).

Figure 10. Performance in Experiment 4 (Set size vs. crowding). Uncrowded trials are shown on the left half of the graph, crowded trials are shown on the right half. The figure shows accuracy (% correct; plate A) and transformed accuracy (Log10; plate B for the three set size conditions (4, 8, 12) by each of the two mask duration conditions (0, 180 ms). Error bars represent +/- 1 standard error.

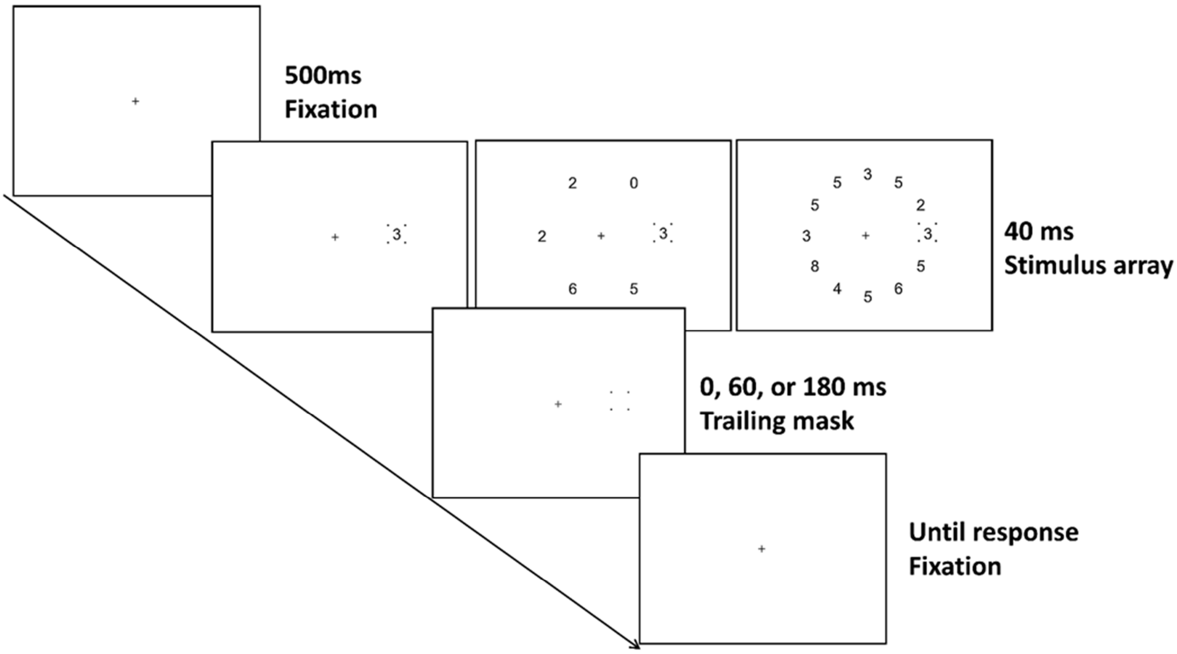
500ms
Fixation
2 0
2 3
6 5
5 3 5
5 2
3 3
8 5
4 5 6
40 ms
Stimulus array
0, 60, or 180 ms
Trailing mask
Until response
Fixation

Figure 1.

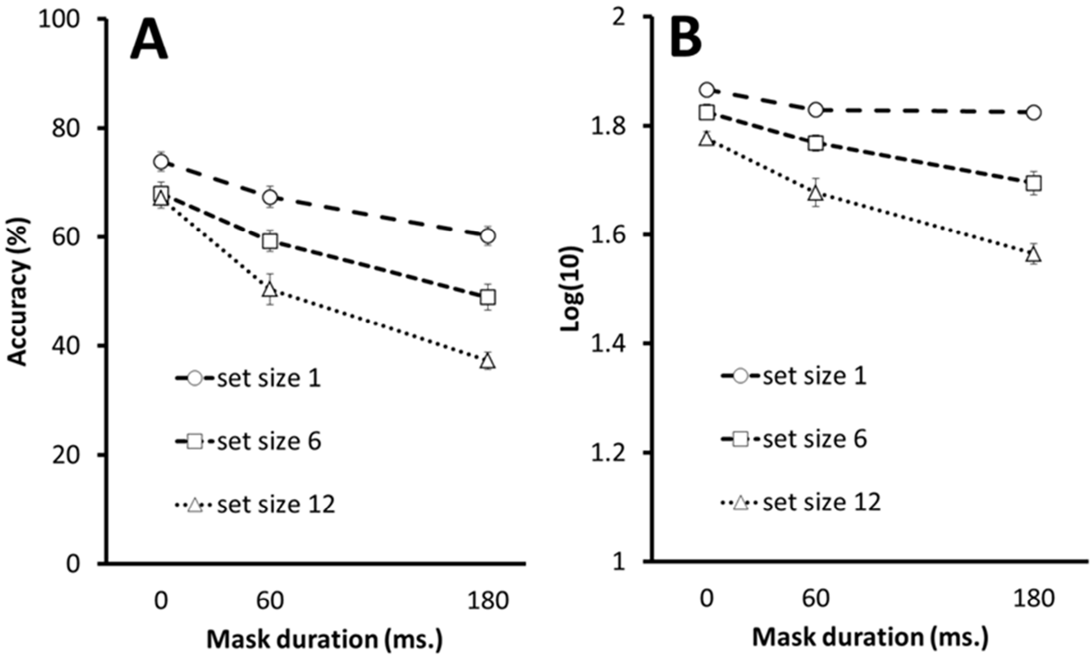
A
B
Accuracy (%)
100
80
60
40
20
0
Log(10)
2
1.8
1.6
1.4
1.2
1
set size 1
set size 6
set size 12
0
60
180
Mask duration (ms.)

Figure 2.

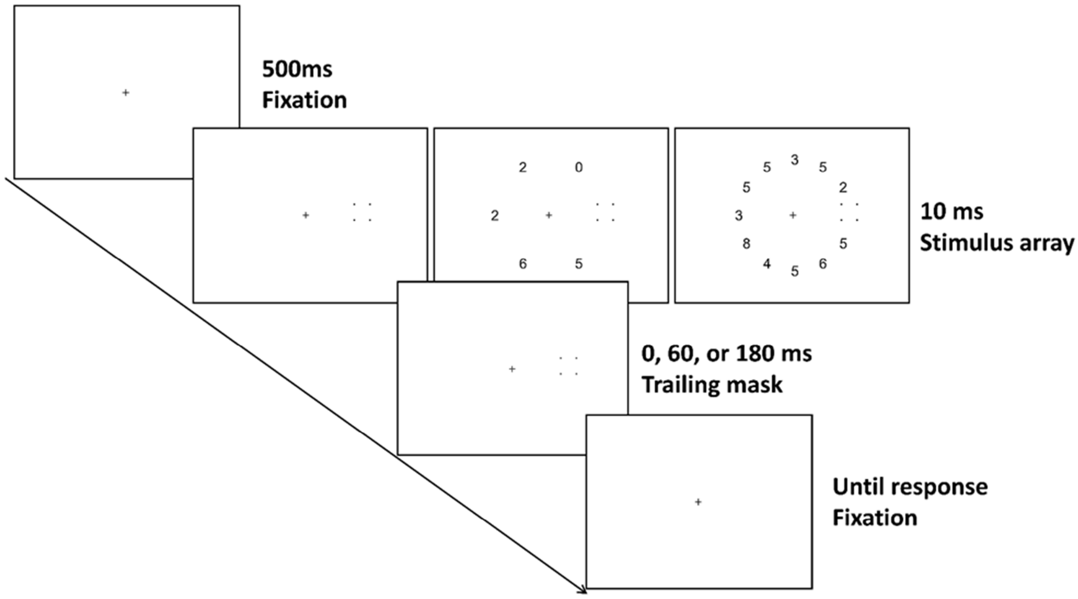
500ms
Fixation
2
0
2
6
5
5
3
5
5
2
3
8
5
4
5
6
10 ms
Stimulus array
0, 60, or 180 ms
Trailing mask
Until response
Fixation

Figure 3.

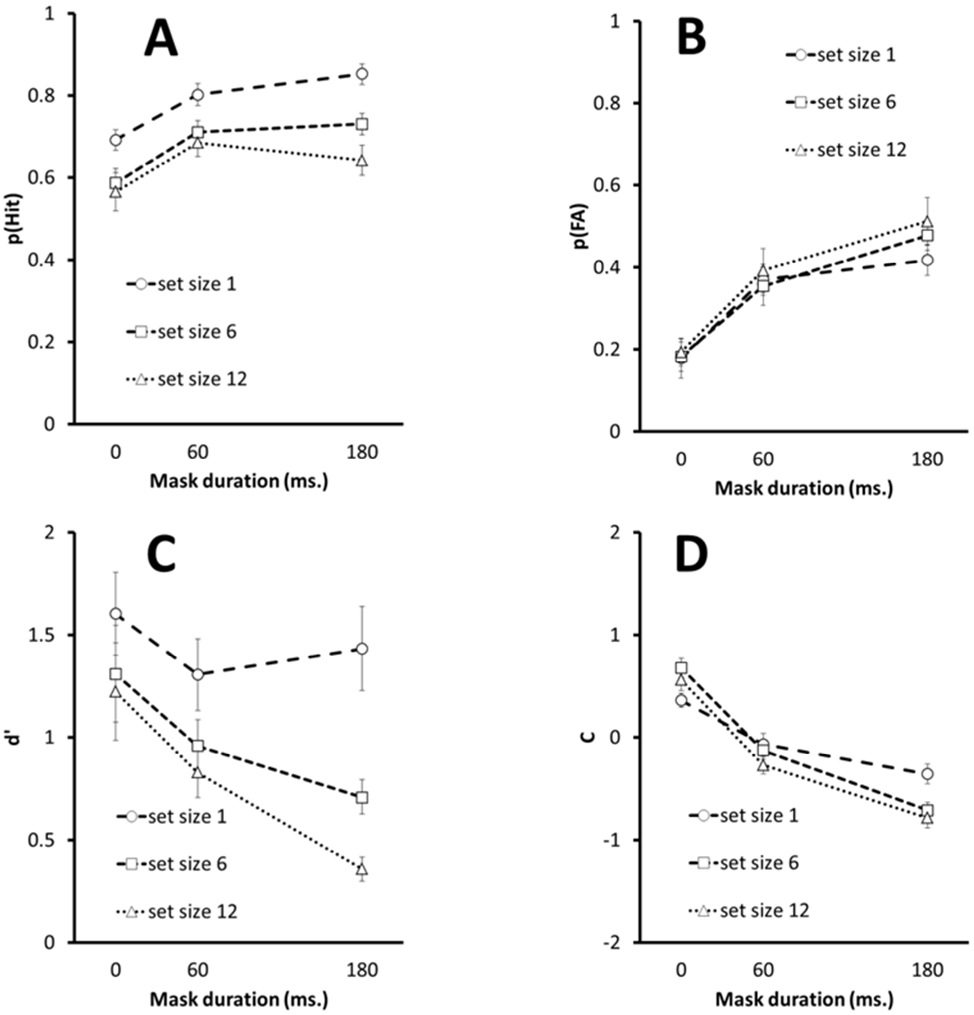
A
p(Hit)
Mask duration (ms.)
set size 1
set size 6
set size 12
B
p(FA)
Mask duration (ms.)
set size 1
set size 6
set size 12
C
d'
Mask duration (ms.)
set size 1
set size 6
set size 12
D
C
Mask duration (ms.)
set size 1
set size 6
set size 12

Figure 4.

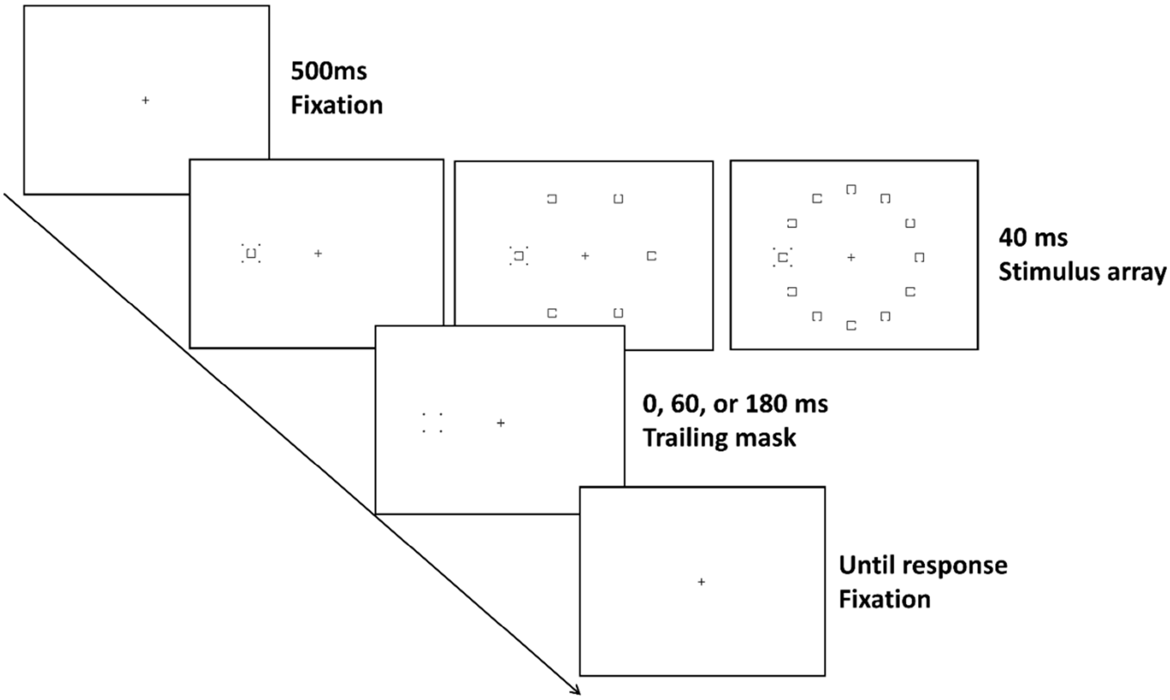
500ms
Fixation
40 ms
Stimulus array
0, 60, or 180 ms
Trailing mask
Until response
Fixation

Figure 5.

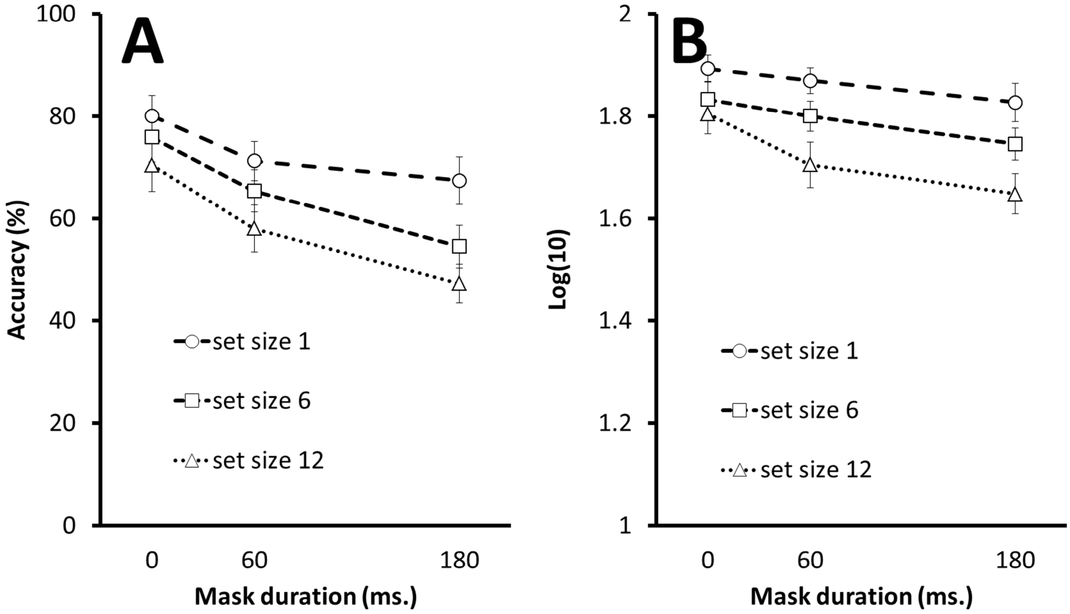
A
B
Accuracy (%)
100
80
60
40
20
0
Log(10)
2
1.8
1.6
1.4
1.2
1
set size 1
set size 6
set size 12
0
60
180
Mask duration (ms.)

Figure 6.

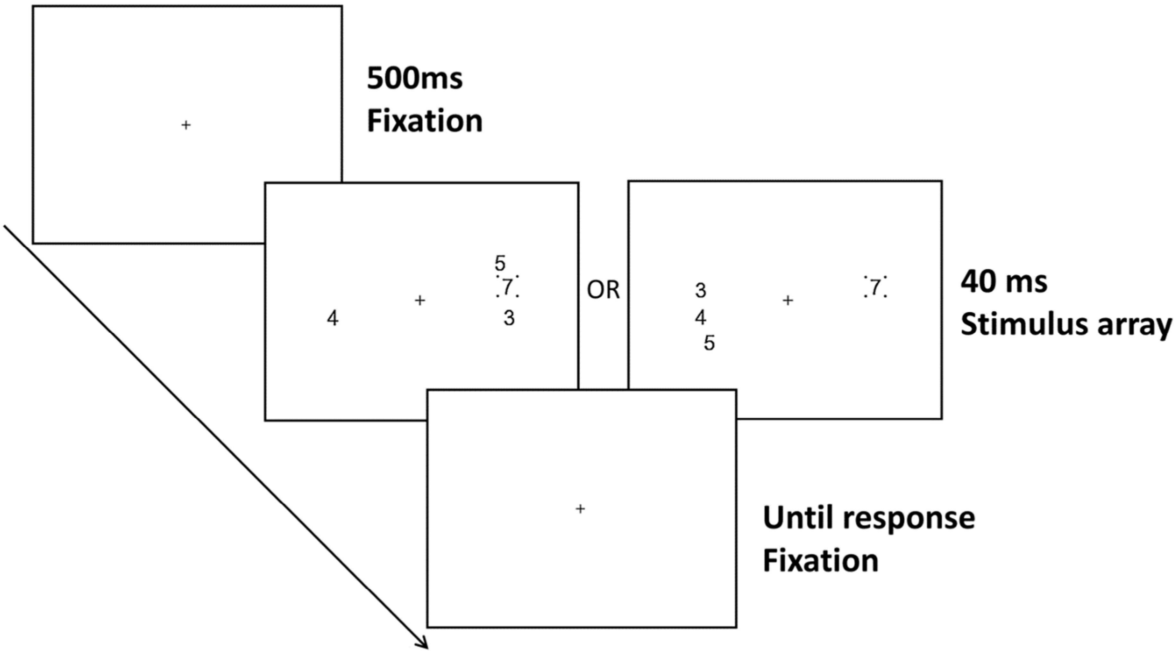
500ms
Fixation
5
7
4
3
OR
3
4
5
7
40 ms
Stimulus array
Until response
Fixation

Figure 7.

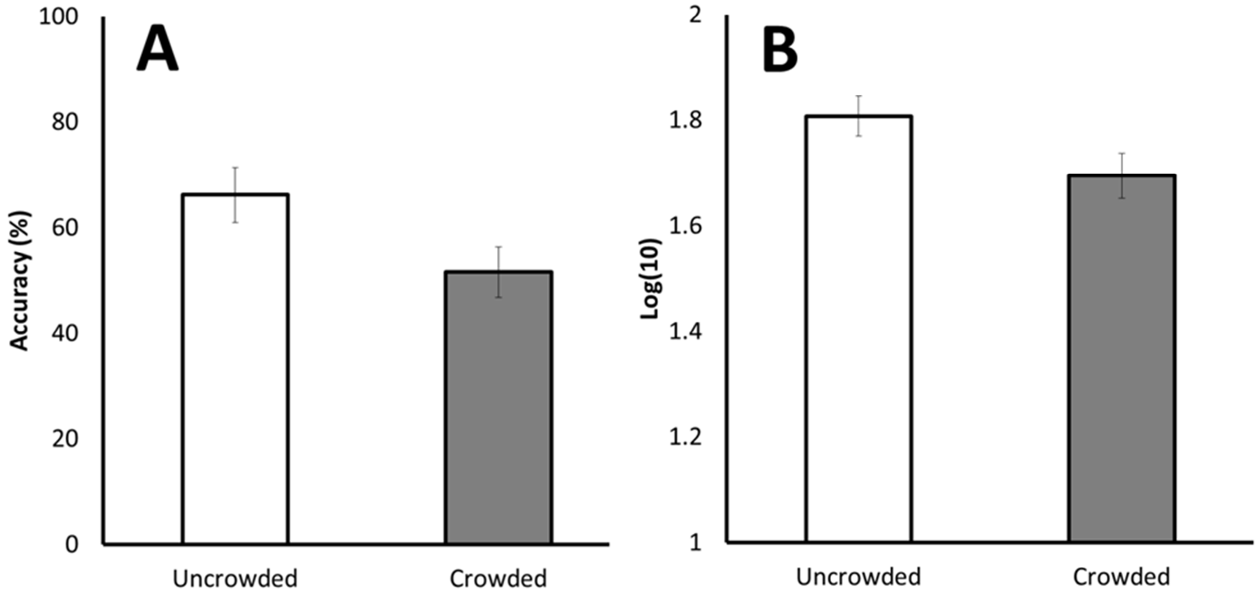
A
Accuracy (%)
100
80
60
40
20
0
Uncrowded
Crowded
B
Log(10)
2
1.8
1.6
1.4
1.2
1
Uncrowded
Crowded

Figure 8.

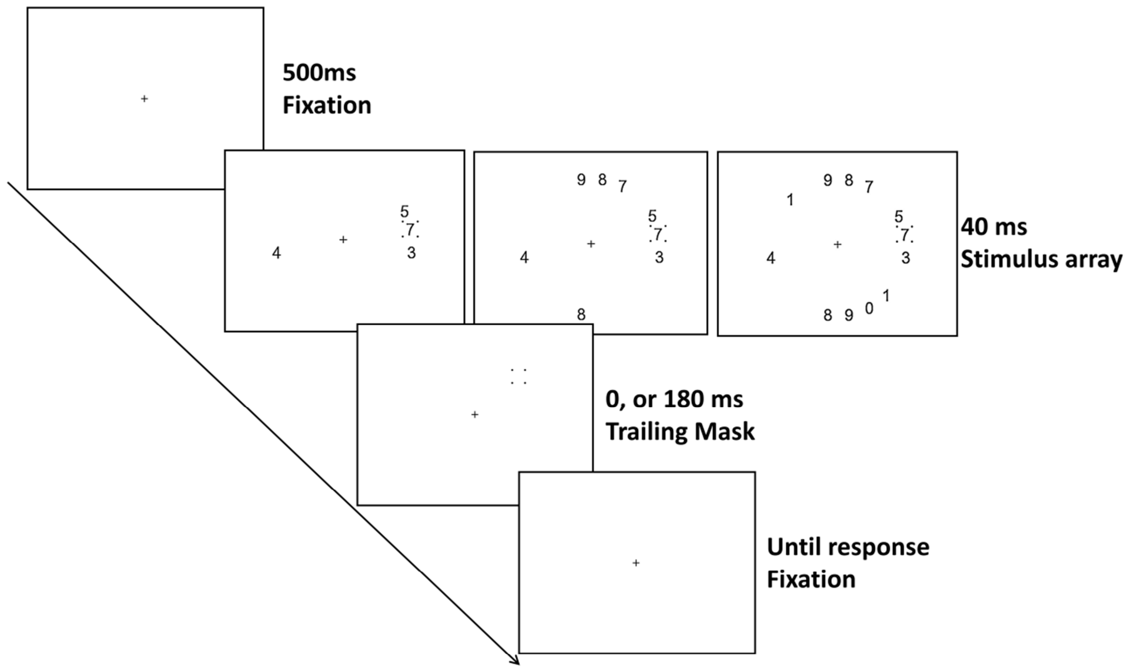
500ms
Fixation
40 ms
Stimulus array
0, or 180 ms
Trailing Mask
Until response
Fixation

Figure 9.

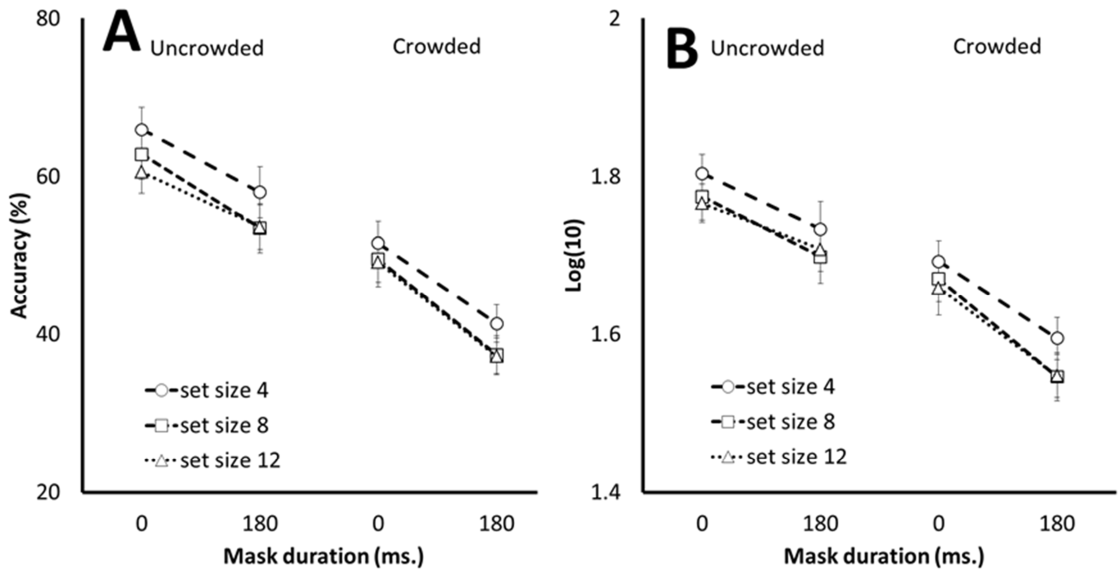
A
Uncrowded
Crowded
Accuracy (%)
80
60
40
20
set size 4
set size 8
set size 12
0
180
0
180
Mask duration (ms.)
B
Uncrowded
Crowded
Log(10)
2
1.8
1.6
1.4
set size 4
set size 8
set size 12
0
180
0
180
Mask duration (ms.)

Figure 10.