

Testing models of object substitution with backward masking

Gregory Francis¹

Purdue University
Department of Psychological Sciences
703 Third Street
West Lafayette, IN 47907-2004

and

Yang Seok Cho

Korea University
Department of Psychology
Anam-dong, Seongbuk-Gu
Seoul, Korea 136-701

October 31, 2005

Revised: May 15, 2005

Key words: backward masking, metacontrast, models, object substitution

Running head: Models of object substitution - P393

¹E-mail: gfrancis@psych.purdue.edu; phone: 765-494-6934. This material is based upon work supported by the National Science Foundation under Grant No. 0108905.

Abstract

A briefly presented visual target stimulus can be difficult to identify if shown in the context of a mask stimulus. There are several quantitative models that have been used to explain many different masking effects. Here we look at modeling what has been called object substitution. We analyze four models and show that three of the models work with a common principle while the fourth behaves quite differently. The three similar models have previously been used to explain many other aspects of visual masking, while the fourth model was designed exclusively to deal only with object substitution masking effects. We identify an experimental test to choose between the models. The experimental results support the behavior of the fourth model.

Introduction

A briefly presented target stimulus can be difficult to see if it is presented in combination with a mask stimulus. This kind of visual masking is a fundamental tool in visual perception and cognitive psychology, where it is used to limit the amount of information processing [see reviews by Breitmeyer & Ögmen (2000) and Enns & Di Lollo (2000)]. Given the ubiquitous role of masking throughout psychology, it is important to develop theoretical accounts of the underlying mechanisms that are involved in masking. In a series of publications, Enns and Di Lollo (1997, 2000) and their colleagues (Bischof & Di Lollo, 1995; Di Lollo, Bischof & Dixon, 1993; Di Lollo, Enns & Rensink, 2000) reported new experimental properties of masking that they argued could not be accounted for by earlier theories of masking. The new properties showed that when attentional focus was not directly cued on the target, strong masking could occur even for a very sparse visual mask (just four dots around the target). Enns and Di Lollo (1997) and Di Lollo *et al.* (2000) proposed that masking in this situation was the result of interactions at an object level of analysis, where the mask stimulus replaced information about the target due to reentrant processing in cortical neural circuits. Enns and Di Lollo referred to this type of masking as object substitution, and Di Lollo *et al.* developed a quantitative model that matched the experimental data very well.

Other research groups have quickly adopted the theoretical basis and experimental techniques proposed by Di Lollo *et al.* (2000) to explore additional properties of object substitution (e.g., Jiang & Chun, 2001a,b; Kahan & Mathis, 2002; Lleras & Moore, 2003; Neill, Hutchison & Graves, 2002). However, Francis and Hermens (2002) simulated several existing quantitative models of visual masking and showed that these older models could also account for the experimental data of Di Lollo *et al.* (2000), provided the models assumed that an effect of visual attention was to modulate the strength of mask inhibition toward the target signal. The match between the models and the experimental data was quite good. As a result of this match it is currently unclear whether new models of masking are really required to account for object substitution effects.

In a reply to the comments of Francis and Hermens (2002), Di Lollo, Enns and Rensink (2002) argued that the models were inappropriately applied to their data. Although we disagree with many of the comments from Di Lollo *et al.* (2002), in this paper we explore a more direct test of the models to explain properties of backward masking. We also consider the model proposed by Di Lollo *et al.* (2000) and explore how well it generalizes to backward masking conditions.

Our analysis of the models reveals that theories of object substitution can be tested by having the mask onset follow the target onset, a condition called backward masking. Although backward masking was used in early studies of object substitution (Enns & Di Lollo, 1997), more recent studies have used common onset masking (Di Lollo *et al.*, 2000), where the target and mask appear simultaneously and the mask remains present after the target disappears. We show that current models of object substitution differ dramatically with regard to backward masking conditions, and we use this difference to test the models.

Model analysis and simulations

Backward masking conditions contrast the model proposed by Di Lollo *et al.* (2000) against the models used by Francis and Hermens (2002). All of these models match the object substitution data reported by Di Lollo *et al.* (2000), which used common onset masking. We briefly explain how this comes about and what sorts of assumptions are needed to make the models account for the experimental data. Additional details of the models can be found in Francis (2000) and Francis and Hermens (2002). Simulations of many of the models can be found on line, as described in Francis (2003).

We are interested in computational principles rather than the specifics of any particular model. Models that have very different descriptions, mechanisms, and proposed neurophysiological interpretations may still behave according to the same computational principles. Verifying that a basic principle is responsible for a certain aspect of model behavior allows for exploration of a large class of models. Likewise, verification that a basic computational principle is unable to match a particular set of data allows for the rejection of a large class of models and can suggest development of alternative models.

U-shaped backward masking

Often times the properties of the target and mask stimuli are held fixed, and the stimulus onset asynchrony (SOA), the time between the target onset and the mask onset, is varied. The resulting set of data is called a masking function. One particularly interesting characteristic of backward masking is that the masking function is sometimes u-shaped. For short SOAs the target is clearly seen, and the required task fairly easy to perform. For middle duration SOAs (often less than 100

milliseconds), the target is harder to see and the task difficult to perform. For long SOAs the task performance is again quite good, perhaps because the target is partially processed before the mask appears.

There are a variety of computational models that account for the existence of both u-shaped and monotonic-shaped masking functions (Weisstein, 1972; Bridgeman, 1971, 1978; Anbar & Anbar, 1982, Francis, 1997, 2003). Francis (2000) identified a simple computational principle that accounts for the appearance of a u-shaped masking function and showed that this principle is used by all of these models. This principle is called mask-blocking. In a mask blocking system, a strong representation of target information can block inhibitory effects generated by the mask.

To demonstrate the principle of mask blocking (see also Francis, 2000; Francis & Cho, 2005), consider a system that generates initial responses $X_T(t)$ and $X_M(t)$, respectively, for the target and mask stimuli. These are single-value variables that characterize some aspect of the visual system's response to the stimuli.

$$\frac{dX_T}{dt} = -AX_T + I_T(t) \quad (1)$$

and

$$\frac{dX_M}{dt} = -AX_M + I_M(t). \quad (2)$$

As is common in writing differential equations, the dependence of X_T and X_M on time is implied but not written. The terms $I_T(t)$ and $I_M(t)$ indicate input from the target and mask stimuli, respectively. Without any input, each variable will decay to a value zero at a rate set by parameter A . The values of these variables do not correspond to perceptual awareness of the stimuli. Instead, these equations contribute to a visual response function (VRF), $v(t)$, that takes excitatory activity from the target response and inhibitory activity from the mask response. A key property of the inhibition is that it is only present if the mask signal is stronger than the target signal. It is in this way that mask-blocking occurs; a strong target signal can block the inhibitory effect of the mask:

$$\frac{dv}{dt} = -Av + X_T - [X_M - X_T]^+. \quad (3)$$

The term $-[X_M - X_T]^+$ in equation (3) describes the interaction of the target and mask. The notation $[]^+$ represents a rectification function so that if the term inside the brackets is not positive, then the function returns the value zero. If the term inside the brackets is positive, the function returns the value unchanged. Thus, the term $-[X_M - X_T]^+$ will be negative if X_M is larger than X_T and will be zero otherwise. Hence, masking occurs when X_M is bigger than X_T . This implements

mask-blocking because if the target response, X_T , is bigger than the mask response, X_M , the mask has no influence on the target VRF.

The value of the target VRF does not correspond to perceptual awareness of the target stimulus. The percept is assumed to be computed from the value of the target VRF by an integration over time:

$$P = \int_0^T [v(t) - G]^+ dt, \quad (4)$$

where G is a threshold parameter, the value zero as the lower limit of the integral indicates the onset of the target, and T is some upper limit of time for the integral. This form of the integral treats as zero any values of $v(t)$ that are below the threshold value. The values of P should be related to experimental data on masking, and changes in the value of P should correspond to changes in the behavioral measure of masking.

In the simulations described below, the parameters were set as $A = 0.1$ and $G = 0.1$, and T was chosen so that the integral in equation (4) included all non-zero terms. The input for the target was defined as:

$$I_T(t) = \begin{cases} 10 & \text{for } 0 \leq t < 10 \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

so that it turned on at time zero and off at time ten. The input for the mask was:

$$I_M(t) = \begin{cases} I_M & \text{for } \tau_1 \leq t < \tau_1 + 10 \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

so that it turned on at SOA equal τ_1 and turned off ten time units later. The value of I_M is the intensity of the mask. In Figures 1a–b, $I_M = 5$ and in Figures 1c–d, $I_M = 20$.

The effect of strong mask-blocking is indicated in Figure 1a. In this situation, the target and mask stimuli start and end at the same time (SOA=0). The mask input is weaker than the target input, so the mask response is always below the target response. As a result, the mask has no effect on the target VRF; the integral of the target VRF is equivalent to a case where the mask was not present at all. The strong target signal has blocked the mask from producing any masking.

– Figure 1 –

Weaker mask-blocking, and thus stronger masking, is indicated in Figure 1b. Here, the mask onset follows the target onset by an SOA of 30 time units. As result of this time delay, the rising part of the mask response overlaps a fading part of the target response. When the mask response is larger than the target response, it begins to send an inhibitory signal that quickly drives the target

VRF below zero. The area under the target VRF in this case is smaller than for Figure 1a. Thus, increasing the SOA between the target and mask stimuli results in stronger masking. This release from mask-blocking is what produces the downward sloping part of a u-shaped masking function. For much longer SOAs, the mask will have less influence on the target VRF because it arrives too late. This effect produces the upward slope of a u-shaped masking function.

A system that includes mask-blocking will produce a u-shaped masking function if the mask is weak enough that it cannot produce much masking for short SOAs, but can produce some masking when the target response has faded during medium SOAs. For longer SOAs, the later arrival of the mask will always free the target VRF from any masking that might have occurred over medium or short SOAs.

The same system can produce a monotonic increasing masking function if the mask is strong. Figures 1c–d, show simulation plots for a strong mask signal (twice the intensity of the target). In Figure 1c, with SOA=0, the mask response is stronger than the target response at all times, so the target VRF immediately receives strong inhibition and never rises above zero. In Figure 1d the mask is delayed by SOA=30, and the target VRF is not inhibited until the mask's response is larger than the (fading) target response. Thus, for a strong mask, mask-blocking does not occur, or is very weak, and the masking function will be monotonic as SOA varies.

Such a system predicts u-shaped masking functions should appear for relatively weak masks and monotonic-shaped masking functions should appear for relatively strong masks. For the system and parameters of the present simulations, the masking functions are shown in Figure 2 and demonstrate this property. Quite similar behavior can be achieved by other computations (e.g., division of the mask inhibitory signal by the target signal).

– Figure 2 –

Many quantitative models create a u-shaped masking function with a version of masking blocking (Francis, 2000). Although the models differ substantially in terms of mechanistic implementation, equations, parameters and physiological interpretations, analysis of the models reveals that they all follow the same pattern of behavior as the simple mask-blocking system described above. Francis and Herzog (2004) showed through computer simulation that when the mask is weak, each model produces a u-shaped masking function, and when the mask is stronger each model produces a monotonic-shaped masking function. The definition of weak and strong varies across the models,

but the overall behavior of the models was the same. This relationship between the strength of the mask and the shape of the masking function is a fundamental aspect of any system that uses mask-blocking and so is a robust property of these models. The relationship will not disappear with changes in the model parameters, unless they are so substantial as to alter the basic behavior of the model.

Thus, a fundamental characteristic of a system that uses mask-blocking is that the shape of the masking function is related to the *strength* of the target and mask signals. In current models, the intensity and spatial properties of the target and mask stimuli are converted into response strengths and such strengths determine the shape of the masking function. In the experiments below, we use this property to explore model explanations of object substitution effects.

Model explanations of object substitution effects

What is called object substitution refers to a combination of three basic experimental findings. First, masking can appear with a spatially sparse mask (often just four dots). Second, increases in the duration of the mask while keeping the target duration fixed leads to stronger masking. Third, as attention is diverted away from the target (often due to set size variations), masking becomes stronger. We refer to this combination of properties as object substitution masking effects. We are exploring several theoretical explanations of these effects.

Di Lollo *et al.* (2000) did also identify several other properties that they suggested distinguish object substitution effects from other masking phenomena. These properties are not discussed in the present paper because the models are largely silent on their significance. That is, the models do not emulate the visual system with sufficient detail to address those properties. While this does mean the current models are incomplete, we cannot reject the possibility that the models can be further developed to account for these other data sets.

Francis and Hermens (2002) noted that a model that follows the mask blocking principle for producing a u-shaped backward masking function would also account for the first two properties of object substitution. That a spatially sparse mask can produce strong masking is simply due to the mask generating some type of inhibitory response in the visual system. In a system that uses mask blocking, the small inhibitory effect builds up over time. That masking increases with mask duration is likewise due to the mask blocking system's behavior of building up masking effects over extended periods of time. The longer the mask's duration, the longer the inhibitory signal reduces

the target visual response function.

The models of backward masking have not previously included an attentional component. There are surely several ways to introduce an attentional component. Francis and Hermens (2002) suggested that an increase in attentional focus on the target would reduce the strength of the inhibitory signal generated by the mask. This idea is consistent with general ideas that one effect of attention is to modulate the strength of signals corresponding to various stimuli (e.g., Blaser, Sperling & Lu, 1999).

An example of how attentional focus might be implemented in the model described above is to modify equation (3) to include an attentional weight term:

$$\frac{dv}{dt} = -Av + X_T - \alpha[X_M - X_T]^+, \quad (7)$$

where parameter α modulates the inhibitory effects of the mask. When attention is focused on the target, α is small so that there will be little masking. When attention is more distributed, α will be larger so that the mask has a bigger impact on the target VRF.

Figure 3a shows the model's behavior when placed under conditions similar to those used by Di Lollo *et al.* (2000) to explore object substitution. The target and mask stimuli were presented with a common onset and the duration of the mask stimulus continued after offset of the target for various durations. To emulate the use of a weak (four-dot) mask, the intensity of the mask stimulus was one tenth the value of the target intensity. The different curves are for different values of α . A small value of α corresponds to a small set size in the study of Di Lollo *et al.* (2000), which allows for focus of attention on the target. Increasing values of α correspond to larger set sizes, which cause attention to be more distributed.

– Figure 3 –

A physical manifestation of these attention effects would vary dramatically depending on the details of a particular model. We would like to test the basic idea without worrying about the model details. Toward that end, we followed the approach used by Francis and Hermens (2002), who noted that the hypothesized effects of attention can be emulated by reducing the intensity of the mask signal. Figure 3b shows the behavior of the model with $\alpha = 1$ and the mask intensity varied. The effect of varying mask intensity is quite similar to varying α .

Figures 4a, b, and c show that the models of Weisstein (1972), Bridgeman (1978), and Francis (1997) have very similar properties when the mask intensity is varied. Attention effects in these

models can be emulated by a variation in mask intensity. It should be emphasized that we are not claiming that attentional focus actually causes the mask brightness to be reduced. Rather, this is a way to explore the model's behavior without introducing additional complexities. Should the behavior seem to match experimental data, then this would suggest that further development in the physical mechanisms of the models would be warranted. On the other hand, should the model behavior fundamentally not match the experimental data, then this approach can be rejected.

– Figure 4 –

As Francis and Hermens (2002) noted, the behavior of these models under object substitution conditions is essentially the same as the model proposed by Di Lollo *et al.* (2000), shown in Figure 4d. Even though the latter model has a very different interpretation of mechanisms, its overall behavior is essentially the same as the mask-blocking models, and all the models do a good job of matching the experimental data. As a result, these data do not provide a good means of choosing between model types. The next section shows that varying the SOA between the target and mask highlights key differences between the models.

Model predictions with backward masking

In this section we explore how the models behave when the target and mask are the same as in the previous section, but the SOA between the target and mask varies. Figure 5 shows model behavior for the same conditions as in Figure 4, except that the mask duration is fixed to equal the target duration and the SOA between the target and mask varies. All of the mask-blocking models predict the appearance of u-shaped masking functions.

Figures 5a–c show simulation results from models proposed by Weisstein (1972), Bridgeman (1978), and Francis (1997), respectively. These models all predict that, with the target and mask stimuli used to study object substitution effects, the visibility of the target stimulus should be minimized when the mask follows the target by a positive SOA, which is a common finding in many backward masking experiments (Alpern, 1953; Francis, 2000; Breitmeyer & Öğmen, 2000).

This behavior is a direct result of the properties of mask-blocking. In a mask-blocking system, a u-shaped backward masking function appears when the mask signal is weak, relative to the target signal. To emulate the presence of a sparse mask, the models assume the mask signal is relatively weak. Moreover, focused attention is hypothesized to further weaken the inhibitory signal generated

by the mask. Thus, everything about how these models account for the object substitution effects suggests that the mask signals must be relatively weak. As a result, mask-blocking generates a u-shaped masking function. A change of model parameters will not fundamentally change this aspect of the models, unless it is so extreme as to prohibit the models from producing u-shaped masking functions at all.

– Figure 5 –

In contrast, Figure 5d shows results from the model proposed by Di Lollo *et al.* (2000). It predicts that detection of a target's shape should be impaired if the mask is presented with the target at SOA=0, but that detection should be unimpaired by the appearance of the mask after a longer SOA. Consistent with earlier studies and simulations, the model predicts an effect of attentional focus, which is described as set size. When the target is presented amongst distracters, detection of the target shape will be more impaired as set size increases.

Thus, for sparse masking under backward masking conditions, there is a clear difference between the model proposed by Di Lollo *et al.* (2000) and the models proposed by Francis and Hermens (2002). Earlier work on sparse masking under backward masking (Enns & Di Lollo, 1997) does not clearly favor one model type over the other, primarily because these experiments were not designed to identify the shape of the masking functions.

Before describing an experiment to test the different predictions, a comment from a reviewer motivates us to discuss the issue of model testing a bit further. This reviewer noted that data already exists that cannot be accounted for by any of the current quantitative models. For example, Neill *et al.* (2002) found that object substitution masking was attenuated when the mask preceded the target, and none of the models used here can accommodate this effect as currently instantiated. What is the purpose, the reviewer wondered, of comparing the models if we already know they are wrong? The question speaks to some important issues in model development and testing [see Pitt & Myung (2002) and Navarro, Pitt & Myung (2004) for discussions of other issues in model testing].

In a certain sense, all models are wrong if only because they are incomplete. Even the models and theories of physics are wrong because they do not yet provide an explanation of everything. Thus, finding that a model is wrong is not useful unless the way the model fails is understood and can be used to drive further model development. A model could be wrong by having, for example,

poor parameters, limited scope, or implausible mechanisms. All one can reasonably hope to do is discover the conditions where a model fails and identify whether the model can be easily modified to account for the discrepancy or whether major changes to the model are required (perhaps so major that the model should be dropped).

Along these lines, it is true, but not useful, to note that all of the models currently under consideration are wrong because they do not include color perception. Such an observation is not useful because it does not suggest how to develop the models to address this aspect of perception. The task of a modeler is to identify those data sets that press on the weaknesses of the models. The goal is not to show that a model is wrong (this is trivially easy to do), but to show how the wrongness is tied up in the fundamental properties of the model. Such a finding can then be used to guide further model development.

We believe that we have identified a key property of a class of models and that an empirical test of that property will help guide further model development. The empirical test is described next.

Experiments

We ran a total of three experiments to test the contrasting model predictions. The first experiment verified that with the stimuli and task that we used, the pattern of the data exhibited what has been taken as evidence of object substitution. The second experiment verified that the target and task that we used does produce u-shaped backward masking functions if the mask is not sparse. Finally, the third experiment tested the predictions of the models. Figures 6a, 7a, and 8a, schematize (in reverse contrast) the display of stimuli for a trial of each experiment. On every trial, the target frame (presented for one refresh frame of a 60 Hz monitor) contained a target and a variable number of distracters arranged in an imaginary circle (6.49 degrees diameter) around a small central fixation square (three frames of a trial with set size four is schematized in Figures 6a, 7a, and 8a). The target was a half-moon that was oriented with its flat side facing up, down, left, or right, while the distracters were filled circles (0.79 degrees in diameter). The observer's task was to report the orientation of the target's flat side. The location of the target among the distracters varied randomly across trials.

When the set size of the target frame equaled eight, all the elements were arranged at the main

compass points and the main diagonals. When the set size of the target frame equaled four, the elements were arranged in the shape of a square, using either the main compass points or the main diagonals. When the set size of the target frame equaled two, the elements were arranged in a dipole across from each other using one axis at random. When the set size of the target frame equaled one, the target element was drawn at a randomly selected position .

Both the target and mask stimuli were drawn in white (225 cd/m^2) on a black (0.6 cd/m^2) background in a room with standard overhead lighting. Luminance measurements were for white or black fields that completely covered the measurement area of a light meter.

Each experiment was self-paced, with the observer pressing a key on a keyboard to start a trial. Each trial started with a fixation point presented by itself for 500 milliseconds, followed by the target and mask elements. After offset of the mask elements, the fixation point remained on the display until the observer indicated the orientation of the target with another keypress. Feedback was given on whether the observer's report was correct for each trial.

The same two observers participated in all of the experiments. One observer was the second author and the other was naive as to the purpose of the study. Both had extensive practice with the experimental task. Before any data was gathered, they went through a version of all three experiments. The observers sat 55 cm from the display.

Experiment 1: Common onset masking

Experiment 1 was designed to insure that our stimulus set could replicate the common onset findings of Di Lollo *et al.* (2000). The mask consisted of sets of four dots (each 0.2 degrees diameter on an imaginary square of 1.23 degrees on each side) that surrounded both the target and the distracters. This spatial arrangement differs from the studies of Di Lollo *et al.* who used only one mask element to function as both a mask and a cue for which target element to report. Our approach is more similar to the method used by Neill *et al.* (2002) and Lleras and Moore (2003). The mask appeared at the same time as the target, but remained on for a variable duration after the target disappeared. An observer saw each set set size and each SOA 80 times.

Some research groups brightness match the elements of the mask stimuli so that there is not a simultaneous change of the duration of the mask and its perceived brightness (e.g., Enns & Di Lollo, 1997). However, the model predictions are not based on brightness matched stimuli, so stimuli in the experiment were also not brightness matched. Brightness matching appears to be

important in some instances (Di Lollo, von Mühlenen, Enns & Bridgeman, 2004), but many studies of masking do not use the brightness matching technique (e.g., Neill *et al.*, 2002) or find that it has no difference in the overall pattern of results (Lleras & Moore, 2003).

Results

Figure 6b plots the percent correct detections of the target orientation against mask alone duration. Different set sizes are plotted as separate lines. There are notable differences between the observers. Observer YS was much more sensitive to the effects of mask duration and set size than observer SY (note the different scales on the y-axis). Nevertheless, both observers show the same general pattern of results. Increasing mask duration tended to make it more difficult for the observer to report the orientation of the mask. The set size effect was fairly weak for short mask alone durations and was stronger for long mask alone durations. The results are qualitatively in agreement with other studies of common onset masking (e.g., Di Lollo *et al.*, 2000)

– Figure 6 –

Experiment 2: Metacontrast masking

Depending on the properties of the target and the task, studies of backward masking sometimes do not find u-shaped effects when varying the SOA between the target and mask. We wanted to be certain that our stimuli and task did allow, in principle, for u-shaped effects to appear. Thus, we ran a backward masking experiment with an annulus mask (1.32 degrees diameter and 0.175 degrees thickness), which is often found to produce u-shaped effects (Breitmeyer, 1984). The experiment was similar to experiment 1, except the mask was presented for two monitor refresh frames, and varied its onset relative to the target frame. The SOA between the target and mask varied from 0 to 60 milliseconds in steps of 15 milliseconds. Each observer saw each set size and SOA 100 times.

Results

Percent correct detections of the target orientation plotted against SOA are shown for different set sizes in Figure 7b. There were again notable differences between observers. Observer YS had a larger set size effect than observer SY. However, the main finding is that both observers produced u-shaped masking functions for almost all conditions. The one exception was for observer YS with

set size two.

– Figure 7 –

The data demonstrate that it is possible, in principle, for u-shaped backward masking functions to appear with this kind of target and judgment. The next experiment can now address what happens when a sparse mask is used in place of the annuli. The models studied by Francis and Hermens (2002) all predict that the masking function should be u-shaped, while the model proposed by Di Lollo *et al* (2000) predicts that strong masking should only appear for common onset of the target and mask stimuli.

Experiment 3: Backward sparse masking

Experiment 3 was the same as experiment 2 except that the annulus masks were replaced by the four-dot masks used in experiment 1.

Results

Percent correct detections of the target orientation plotted against against SOA are shown for different set sizes in Figure 8b. As before, observer YS shows a larger set size effect than observer SY. The main finding, however, is that the strongest masking occurs for common onset of the target and mask. As the SOA between the target and mask increases, percent detections of the target rapidly increase and reach ceiling.

– Figure 8 –

A comparison with the model predictions in Figure 5 indicates that the pattern of data matches the prediction of the model proposed by Di Lollo *et al.* (2000) and is inconsistent with the prediction of the models proposed by Francis and Hermens (2002).

Conclusions

The experimental results definitively show that the explanation of object substitution effects proposed by Francis and Hermens (2002) does not apply when the sparse mask is used under backward masking conditions. Since the models used by Francis and Hermens were originally used to explain properties of backward masking, this finding poses a problem for these models.

As the model analysis demonstrates, the result goes to the heart of the model properties. All of the models use mask-blocking effects to produce a u-shaped masking function. Mask-blocking effects vary the shape of the masking function according to the strength of the mask's inhibitory signal. Weak signals produce u-shaped masking functions, while strong signals produce monotonic shaped masking functions. Following the arguments in Francis and Hermens (2002), the sparse mask with distributed attention should produce a weak signal and thus a u-shaped masking function, contrary to the experimental data. Because mask-blocking effects are at the core of these models, it seems that we need a fundamentally different explanation for these kinds of masking effects.

One might wonder whether the sparse mask is truly weak. All of these models studied by Francis and Hermens (2002), can produce monotonic shaped masking function if the mask is strong. Perhaps the spatial extent and configuration of the dots make the visual representation of the sparse mask stronger than the visual representation of the annulus mask used in Experiment 2. If this were true, it would explain why the sparse mask produces a monotonic shaped masking function and the annulus mask produces a u-shaped masking function.

We can quickly reject this possibility by looking at the actual amount of masking for the different mask types. The target and task of the observer was the same for Experiments 2 and 3. Thus, differences in target detection must be due to the properties of the mask. A comparison of the curves in Figures 7b and 8b shows that the strength of masking is not related to the shape of the masking function. Consider, for example, the data generated by observer YS with a set size of 8. With an SOA of zero, percentage correct target detection for the annulus mask is around 90%, while the percentage correct for the sparse mask is closer to 60%. For SOA zero, one could claim that the sparse mask produces stronger masking than the annulus mask. However, at an SOA of 17 milliseconds, the percentage correct detection of the target with the annulus mask is around 77%, while the percentage correct for the sparse mask is close to 90%. Thus, at SOA 17 milliseconds, the annulus mask is stronger than the sparse mask. This cross-over pattern means that any attempt to identify the relative strengths of the two masks faces the problem that their relative effect on the target differs with SOA. This property poses severe problems for the models of backward masking that are based on mask-blocking (Francis & Herzog, 2004).

Based on its performance in predicting the effects of a sparse mask for backward masking, the model proposed by Di Lollo *et al.* (2000) looks like a good candidate for generating a different explanation. Indeed, Enns (2004) has suggested that the object substitution theory can provide a

coherent explanation of both object substitution effects and backward masking effects. However, the simulation model proposed by Di Lollo *et al.* (2000) is not adequate. The problem is similar to the problem faced by the mask-blocking models. Their model does not have any means of representing different mask stimuli, except by the strength of a mask signal. In that model, all masks lead to a monotonic shaped masking function, but the annuli masks in experiment 2 produced u-shaped masking functions. The model does not have enough degrees of freedom to account for different shapes of the masking function.

It should be emphasized that this criticism only concerns the quantitative model of object substitution masking proposed by Di Lollo *et al.* (2000). There are differences between the object substitution *theory* and the object substitution quantitative *model*. Throughout this manuscript, we have tried to use the term object substitution to refer only to the experimental findings. Di Lollo *et al.* also often use the term to label their theory of how neural information involves reentrant processing across multiple places in visual cortex. The object substitution theory consists of ideas about how information is represented in the visual system, and how information processing proceeds through time. In principle, there may be several different quantitative models that capture the spirit of the object substitution theory. Indeed, some of the models proposed by Francis and Hermens (2002) could be recast as a variation of the object substitution theory.

We cannot test all possible quantitative models of the object substitution theory because the theory does not provide enough guidance to limit the behavior of the models. In the object substitution theory, whether a backward mask produces a monotonic shaped or u-shaped masking function depends on the exact quantitative nature of reentry, stimulus representation, and calculation of the visible percept. Different models that all include reentry and other concepts from the theory may lead to very different predictions of behavior. We see this as a serious problem with the object substitution theory as it currently stands. Perhaps further development of the theory will lead to additional constraints on possible models.

In the meantime, we can only test the quantitative models that have actually been proposed. That we find all current models fail in a fundamental way means that there must be substantial modification to the models. It may be that such improvements are consistent with the object substitution theory.

Along these lines, Enns (2004) argued that some of the differences between different kinds of masks involve how the target and mask stimuli perceptually integrate at the shortest SOAs. Many

researchers have proposed that integrative processes play an important role in masking effects (Eriksen, 1966; Breitmeyer & Ganz, 1976; Michaels & Turvey, 1979; Navon & Purcell, 1981; Reeves, 1982; Bachmann, 1994). Despite this long history, such effects have rarely been incorporated as a part of computational models of masking. The models described by Reeves (1982) and Bachmann (1994) are among the few exceptions. Both the model proposed by Di Lollo *et al.* (2000) and the mask-blocking models proposed by Francis and Hermens (2002) would benefit from an investigation of the involvement of integration effects.

References

- Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America*, **43**, 648–657.
- Anbar, S. & Anbar, D. (1982). Visual masking: A unified approach. *Perception*, **11**, 427–439.
- Bachmann, T. (1994). *Psychophysiology of visual masking: The fine structure of conscious experience*. Commack, New York: Nova Science Publishers, Inc.
- Bischof, W. F. & Di Lollo, V. (1995). Motion and metacontrast with simultaneous onset of stimuli. *Journal of the Optical Society of America A*, **12**, 1623–1636.
- Blaser, E., Sperling, G., & Lu, Z.-L. (1999). Measuring the amplification of attention. *Proceedings of the National Academy of Sciences, USA*, **96**, 11681–11686.
- Breitmeyer, B. (1984). *Visual masking: An integrative approach*. New York: Oxford University Press.
- Breitmeyer, B. & 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. & Ögmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, **62**, 1572–1595.
- Bridgeman, B. (1971). Metacontrast and lateral inhibition. *Psychological Review*, **78**, 528–539.
- Bridgeman, B. (1978). Distributed sensory coding applied to simulations of iconic storage and metacontrast. *Bulletin of Mathematical Biology*, **40**, 605–623.
- Di Lollo, V., Bischof, W. F., & Dixon, P. (1993). Stimulus-onset asynchrony is not necessary for motion perception or metacontrast masking. *Psychological Science*, **4**, 260–263.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, **129**, 481–507.
- Di Lollo, V., Enns, J. T., & Rensink, R. A. (2002). Object substitution without reentry? *Journal of Experimental Psychology: General*, **131**, 594–596.

- Di Lollo, V., von Mühlenen, A., Enns, J. T., & Bridgeman, B. (2004). Target and mask duration in metacontrast masking: data and models. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 733–745.
- 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. (1997). Object substitution: a new form of masking in unattended visual locations. *Psychological Science*, **8**, 135–139.
- Enns, J. T. & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, **4**, 345–352.
- Eriksen, C. W. (1966). Temporal luminance summation effects in backward and forward masking. *Perception & Psychophysics*, **1**, 87–92.
- Francis, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review*, **104**, 572–594.
- Francis, G. (2000). Quantitative theories of metacontrast masking. *Psychological Review*, **107**, 768–785.
- Francis, G. (2003). Online simulations of models for backward masking. *Behavior, Research Methods, Instruments & Computers*, **35**, 512–519.
- Francis, G. & Cho, Y. S. (2005). Quantitative models of backward masking. In B. Breitmeyer & H. Öğmen (Eds.) *The 1st Half Second*, MIT Press.
- Francis, G. & Hermens, F. (2002). Comment on: Competition for consciousness among visual events: The psychophysics of reentrant visual processes, by Di Lollo, Enns and Rensink (2000). *Journal of Experimental Psychology: General*, **131**, 590–593.
- Francis, G. & Herzog, M. (2004). Testing quantitative models of backward masking. *Psychonomic Bulletin & Review*, **11**, 104–112.
- Jiang, Y., & Chun, M. M. (2001a). Asymmetric object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance*, **27**, 895–918.

- Jiang, Y., & Chun, M. M. (2001b). The spatial gradient of visual masking by object substitution. *Vision Research*, **41**, 3121-3131.
- Kahan, T. A. & Mathis, K. M. (2002). Gestalt grouping and common onset masking. *Perception & Psychophysics*, **64**, 1248-1259.
- 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 & Performance*, **29**, 106-120.
- Michaels, C. F. & Turvey, M. T. (1979). Central sources of visual masking: Indexing structures supporting seeing at a single, brief glance. *Psychological Research*, **41**, 1-61.
- Navon, D. & Purcell, D. (1981). Does integration produce masking or protect from it? *Perception*, **10**, 71-83.
- Navarro, D. J., Pitt, M. A. & Myung, I. J. (2004). Assessing the distinguishability of models and the informativeness of data. *Cognitive Psychology*, **49**, 47-84.
- Neill, W. T., Hutchison, K. A., & Graves, D. F. (2002). Masking by object substitution: Dissociation of masking and cuing effects. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 682-694.
- Pitt, M. A. & Myung, I. J. (2002). When a good fit can be bad . *Trends in Cognitive Sciences*, **6**(10), 421-425.
- Reeves, A. (1982). Metacontrast U-shaped functions derive from two monotonic functions. *Perception*, **11**, 415-426.
- Weisstein, N. (1972). Metacontrast. In D. Jameson & L. Hurvich (Eds.) *Handbook of sensory physiology* (Vol. 7, No. 4, *Visual psychophysics*). Berlin: Springer-Verlag.

Figure Captions

Figure 1. Simulations of a system that uses mask-blocking to generate a u-shaped masking function. (a)–(b) a weak mask at SOAs of 0 and 30. (c)–(d) a strong mask at SOAs of 0 and 30. Each simulation stops after offset of the stimuli when the target VRF becomes negative (not shown).

Figure 2. Masking functions for the mask-blocking system. A monotonic masking function occurs for the high intensity mask while a u-shaped masking function occurs for the low intensity mask.

Figure 3. The behavior of the mask-blocking system under common onset masking and attentional focus. (a) The term α modulates the influence of the mask on the target representation, as might happen under varying focus of attention. (b) Essentially the same behavior pattern occurs if the attention effect is emulated as a reduction in the intensity of the mask input.

Figure 4. Simulation results from four models that account for object substitution effects. All of the models do a good job of matching the basic effects. The simulations are based on models proposed by (a) Weisstein (1972), (b) Bridgeman (1978), (c) Francis (1997), and (d) Di Lollo *et al.* (2000).

Figure 5. New simulation results from four models of visual masking under backward masking conditions. All of the simulations emulate conditions with a four-dot mask under varying levels of attentional focus on the target. The models of Weisstein (1972) (a), Bridgeman (1978) (b) and Francis (1997) (c) predict u-shaped masking functions. (d) The model of Di Lollo *et al.* (2000) predicts that strong masking should only occur for common onset of the target and mask.

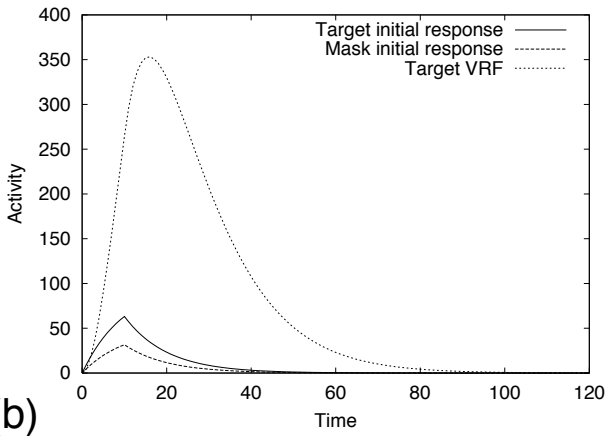
Figure 6. Stimuli (in reverse contrast) and results for experiment 1 with common onset masking. (a) A schematic of three frames from a trial with set size of four. The target and mask elements appeared together and then the target elements disappeared while the mask elements remained. (b) Percentage correct detection against the duration of the mask elements alone. Each graph is for one observer. The pattern of results show an effect of mask alone duration and set size, which agrees with previous research.

Figure 7. Stimuli (in reverse contrast) and results for experiment 2 with backward masking using annuli masks. (a) A schematic of three frames from a trial with set size of four and a positive SOA. The target elements appeared by themselves in the first frame and were replaced by a blank frame. The mask elements then appeared after the appropriate SOA. (b) Percentage correct detection

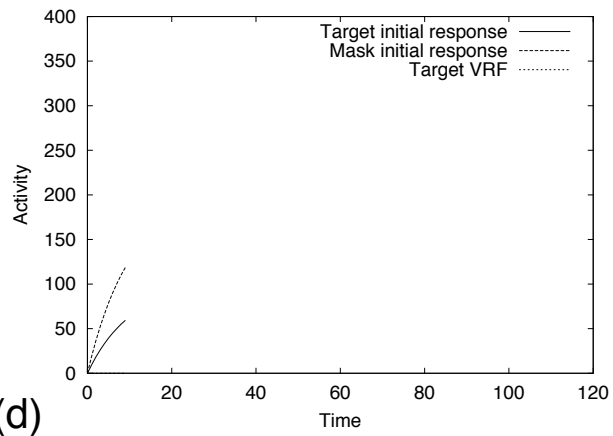
against the SOA between the target and mask frames. Each graph is for one observer. The curves are almost all u-shaped.

Figure 8. Stimuli (in reverse contrast) and results for experiment 2 with backward masking using four-dot masks. (a) A schematic of three frames from a trial with set size of four and a positive SOA. The target elements appeared by themselves in the first frame and were replaced by a blank frame. The mask elements then appeared after the appropriate SOA. (b) Percentage correct detection against the SOA between the target and mask frames. Each graph is for one observer. The curves are monotonic, with the strongest masking occurring for SOA of zero.

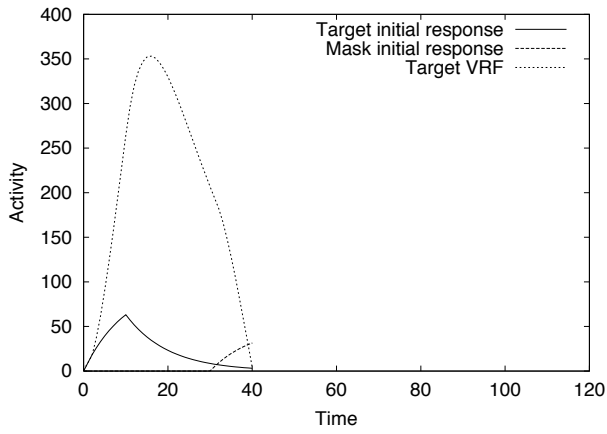
(a)



(c)



(b)



(d)

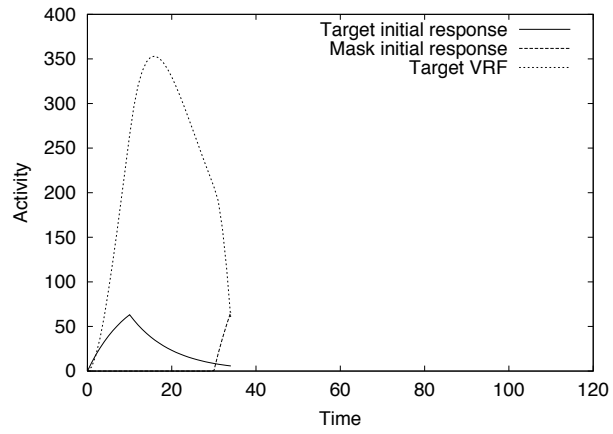


Figure 1:

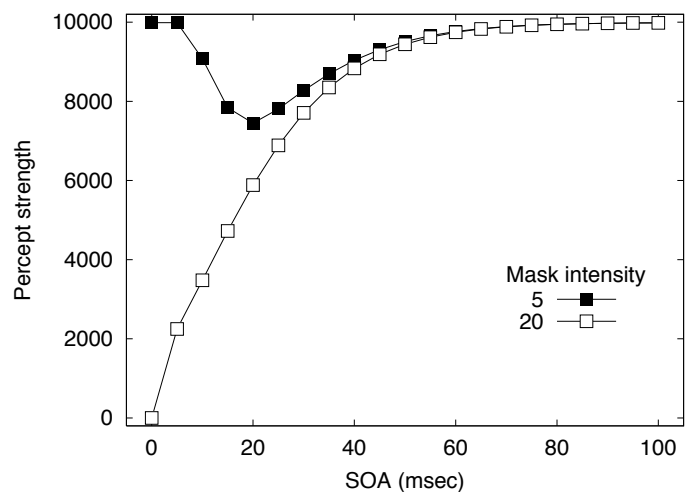


Figure 2:

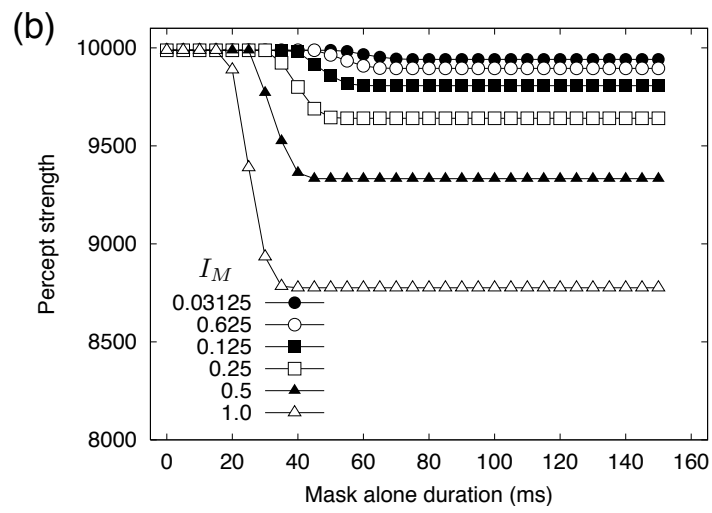
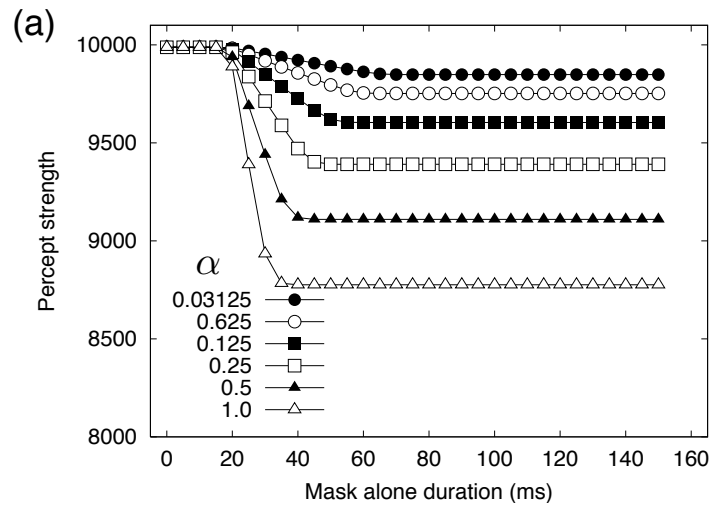


Figure 3:

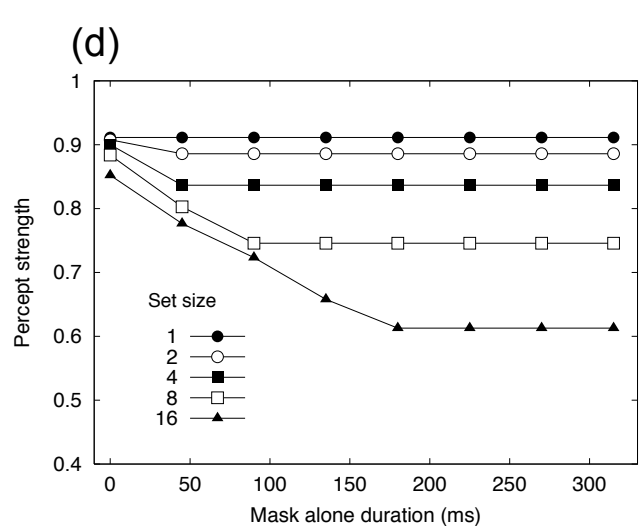
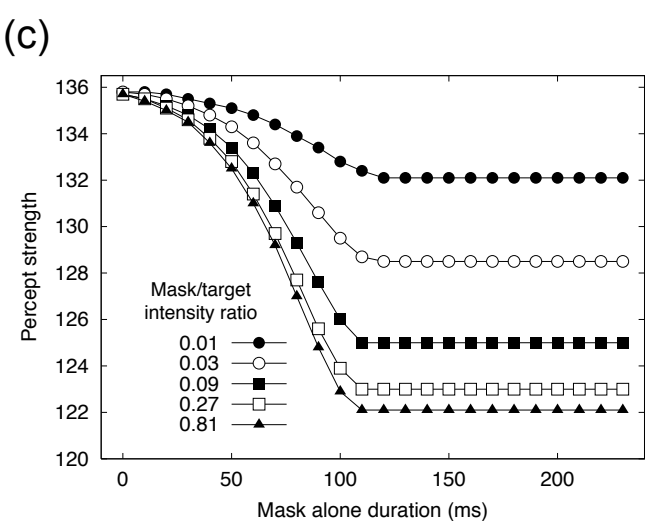
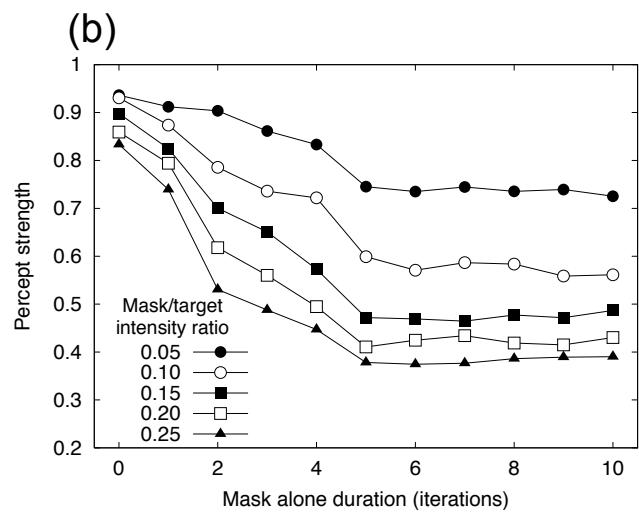
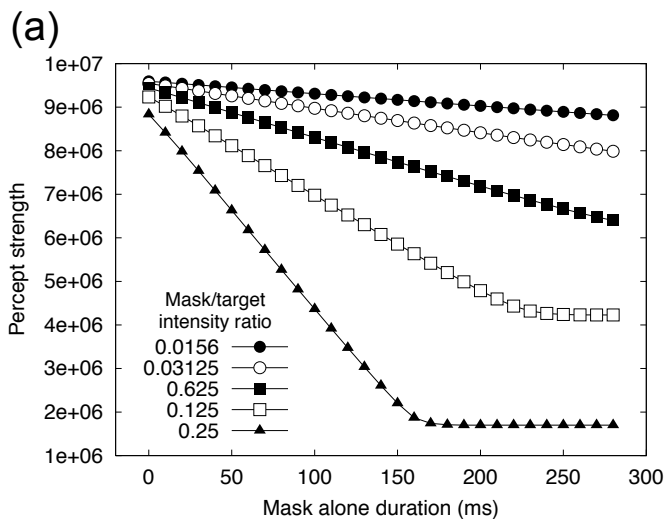


Figure 4:

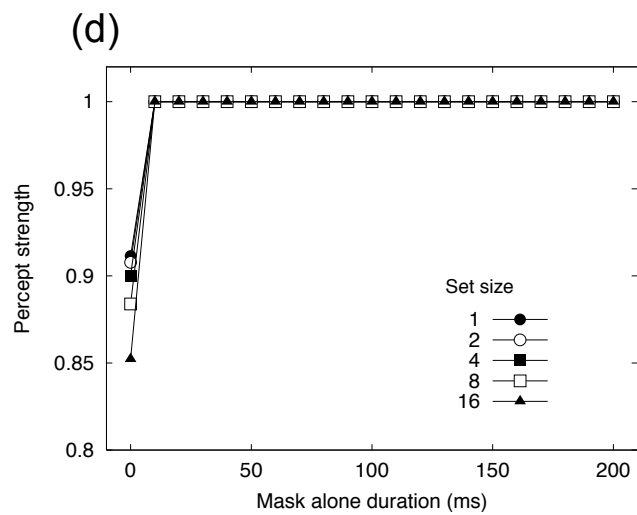
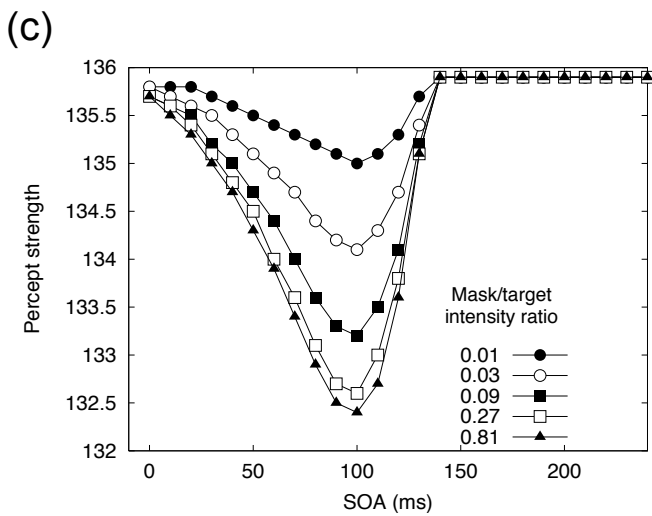
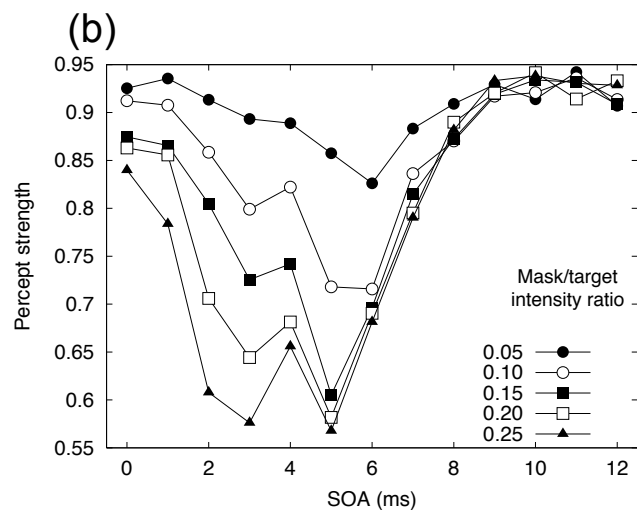
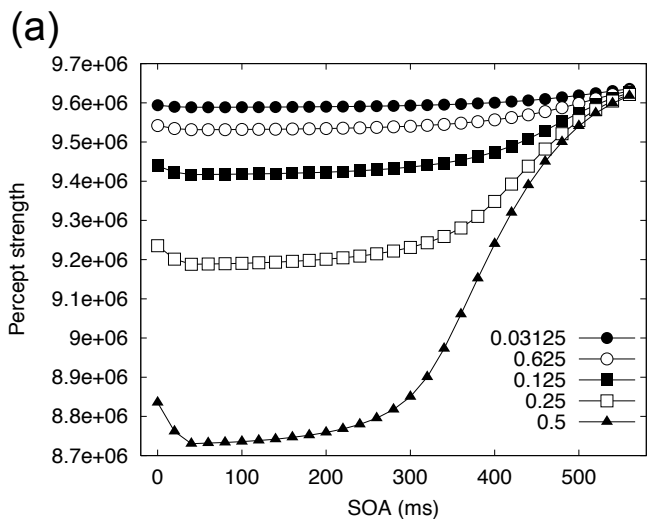
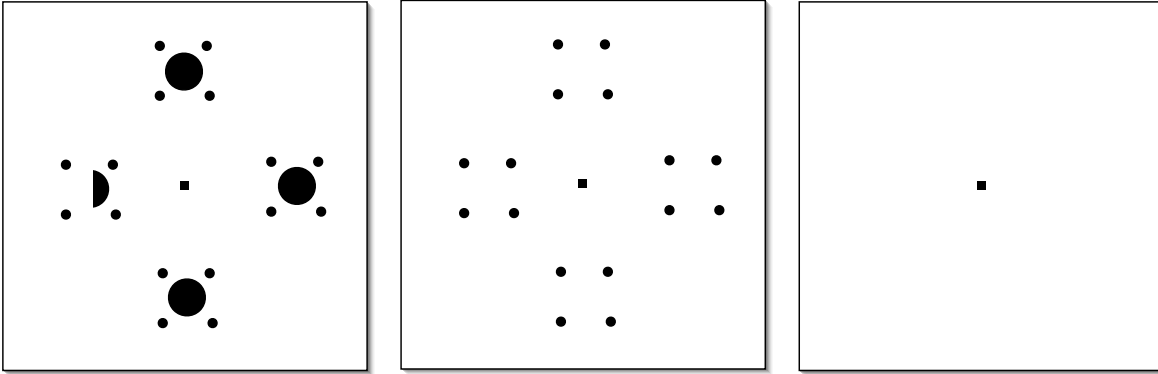
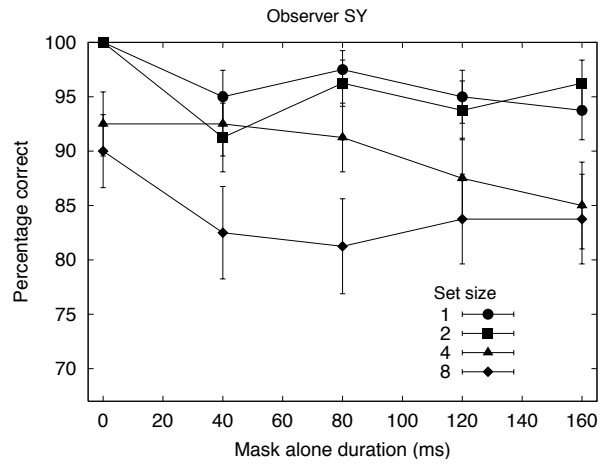
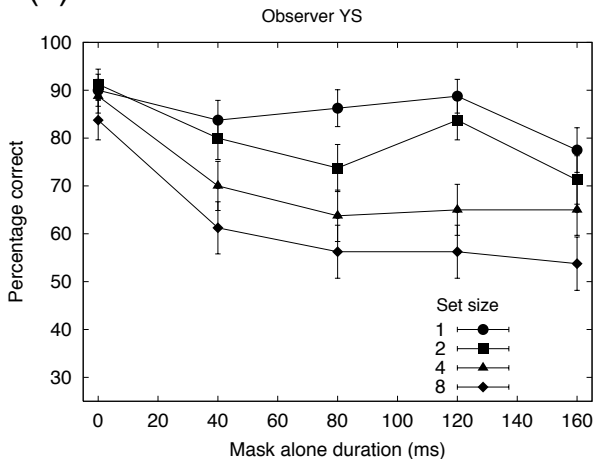


Figure 5:

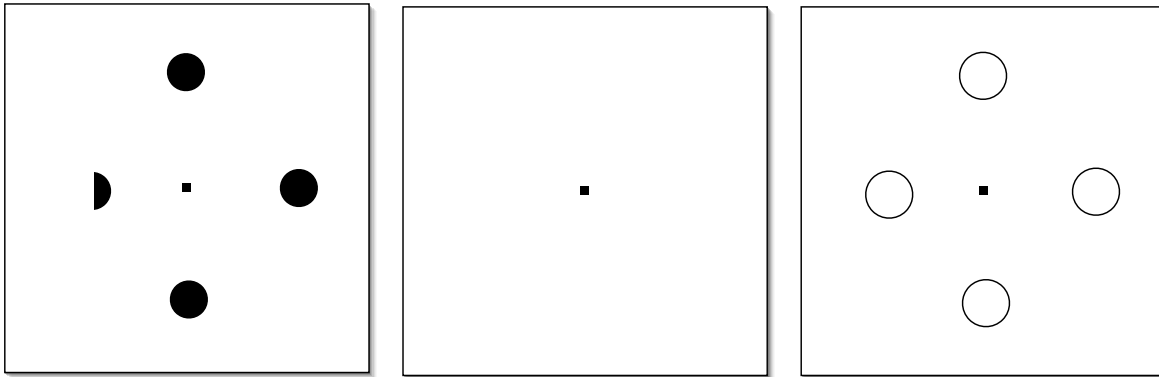
(a)



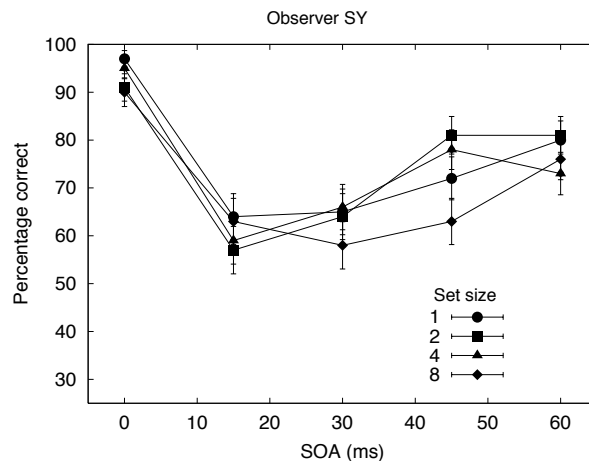
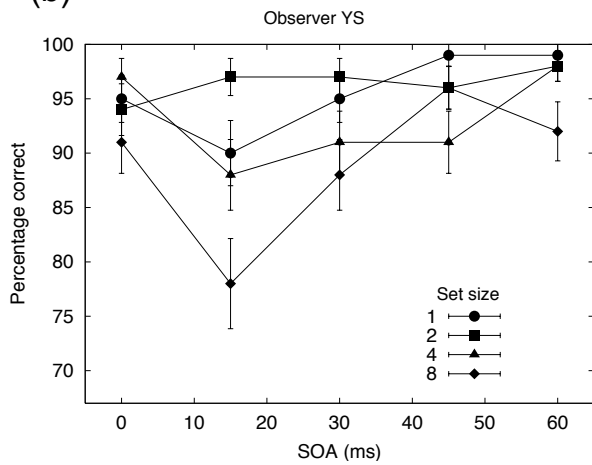
(b)



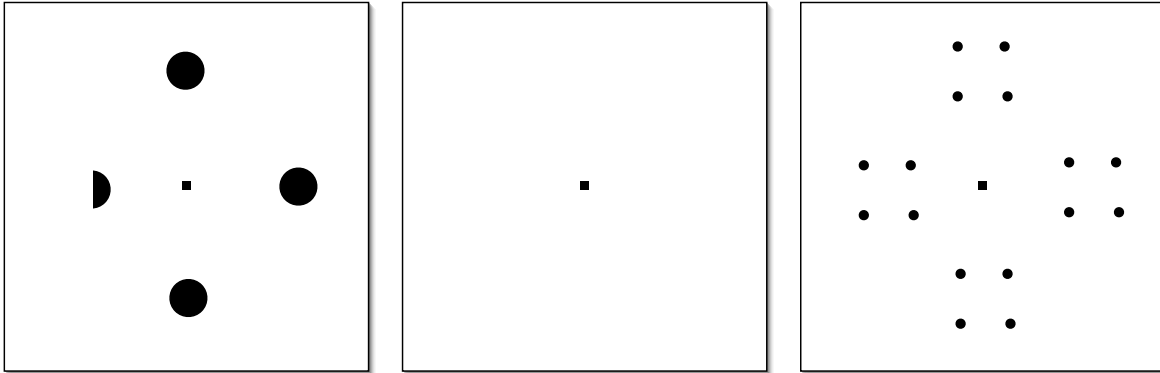
(a)



(b)



(a)



(b)

