7 Computational Models of Visual Masking

Gregory Francis and Yang Seok Cho

7.1 Introduction

Backward visual masking refers to impaired performance on some judgment of a target stimulus when it is followed by a mask. Both the target and the mask stimuli are usually very brief (often less than 50 ms), and the target is chosen such that if it is presented by itself, it is easy for observers to perform whatever judgment is required. However, presentation of a mask stimulus, even 100 ms after the target has been turned off, can make the observer's judgment exceedingly difficult. In some cases, observers report not seeing the target at all. Backward masking is a fundamental tool in cognitive psychology and vision research, where it is used to limit the amount of information processing (see recent reviews by Breitmeyer and Öğmen, 2000; Enns & Di Lollo, 2000). Backward masking is also used to investigate aspects of various types of mental diseases (e.g., Braff & Saccuzzo, 1981; Green et al., 1994a; Slaghuis & Bakker, 1995).

Often 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 ms), 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.

U-shaped masking functions are not always observed, however. In some cases the strongest masking occurs when the target and mask are presented at the same time (SOA = 0), and masking effects grow weaker as the SOA between the target and mask increases. In such situations a monotonic-shaped masking function is found. This chapter explores quantitative theories of the appearance of U-shaped and monotonic-shaped masking functions. 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, 2003a). Francis (2000) showed that all but one of these, which was subsequently analyzed further in Francis (2003a), have a common behavioral property that is responsible for the appearance of a U-shaped masking function. This property is called *mask blocking*. The next section discusses a simple system that uses mask blocking and describes how it produces U-shaped or monotonic-shaped masking functions under different conditions. This analysis then suggests an experiment that tests the entire class of models that use mask blocking and also tests some alternative explanations of the shape of masking functions.

7.2 Analysis of Simplified Mask Blocking

Most of the current quantitative models produce a U-shaped masking function with a common computational approach called *mask blocking*. An analysis of mask blocking in the individual quantitative models can be found in Francis (2000). Rather than describe each of the individual models, this discussion will explore a simple system that uses mask blocking.

To make things concrete, consider a system that generates initial responses $X_T(t)$ and $X_M(t)$, respectively, for the target and mask. These are single-value variables that characterize some aspect of the visual system's response to the stimuli.

$$\frac{dX_T}{dt} = -AX_T + I_T(t) \tag{7.1}$$

and

$$\frac{dX_M}{dt} = -AX_M + I_M(t). \tag{7.2}$$

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:

$$\frac{dv}{dt} = -Av + X_T - [X_M - X_T]^+ \tag{7.3}$$

The term $-[X_M - X_T]^+$ in equation 7.3 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:

$$P = \int_0^T [v(t) - G]^+ dt, \tag{7.4}$$

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 7.4 included all nonzero terms. The input for the target was defined as

$$I_T(t) = \begin{cases} 10 & \text{for } 0 \le t < 10 \\ 0 & \text{otherwise,} \end{cases} \tag{7.5}$$

so that it turned on at time zero and off at time 10. The input for the mask was

$$I_M(t) = \begin{cases} I_M & \text{for } \tau_l \le t < \tau_l + 10 \\ 0 & \text{otherwise,} \end{cases} \tag{7.6}$$

so that it turned on at SOA equal τ_l and turned off 10 time units later. The value of I_M is the intensity of the mask. In figures 7.1a–b, $I_M = 5$, and in figures 7.1c–d, $I_M = 20$.

Mask blocking refers to a situation where signals generated by the target block mask-generated inhibition. In the present system this is implemented by a rectified

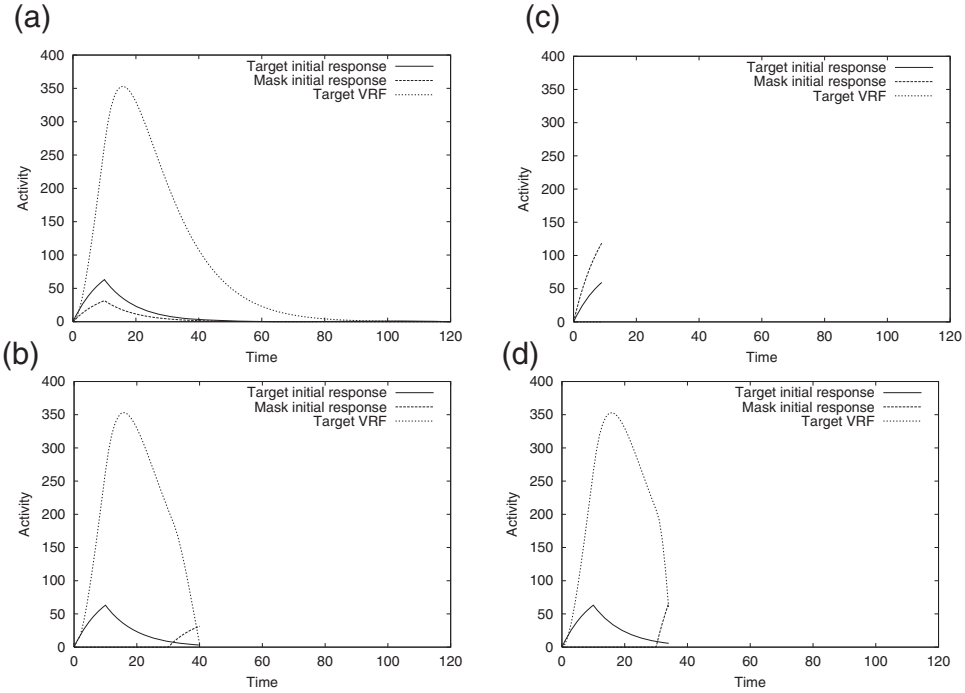

Figure 7.1
Simulations of a system that uses mask-blocking to generate a U-shaped masking function. (*a*)–(*b*) A weak mask at stimulus onset asynchronies (SOAs) of 0 and 30 ms. (*c*)–(*d*) A strong mask at SOAs of 0 and 30 ms. Each simulation stops after offset of the stimuli when the target visual response function (VRF) becomes negative (not shown)

subtraction in equation 7.4. However, quite similar behavior can be implemented using other computations (e.g., division of the mask inhibitory signal by the target signal).

The effect of strong mask blocking is indicated in figure 7.1a. In this situation, the target and the 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 the case in which the mask is not present at all. The strong target signal blocks the mask from producing any masking.

Weaker mask blocking, and thus stronger masking, is indicated in figure 7.1b. Here, the mask onset follows the target onset by an SOA of 30 time units. As a 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 7.1a. 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 monotonically increasing masking function if the mask is strong. Figure 7.1c–d shows simulation plots for a strong mask signal (twice the intensity of the target). In figure 7.1c, 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 7.1d, 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 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 7.2 and demonstrate this property. An online version of this system has been created following the methods described in Francis (2003b). It is available online at http://www.psych.purdue.edu/~gfrancis/Publications/BackwardMasking/.

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, further 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 is 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.

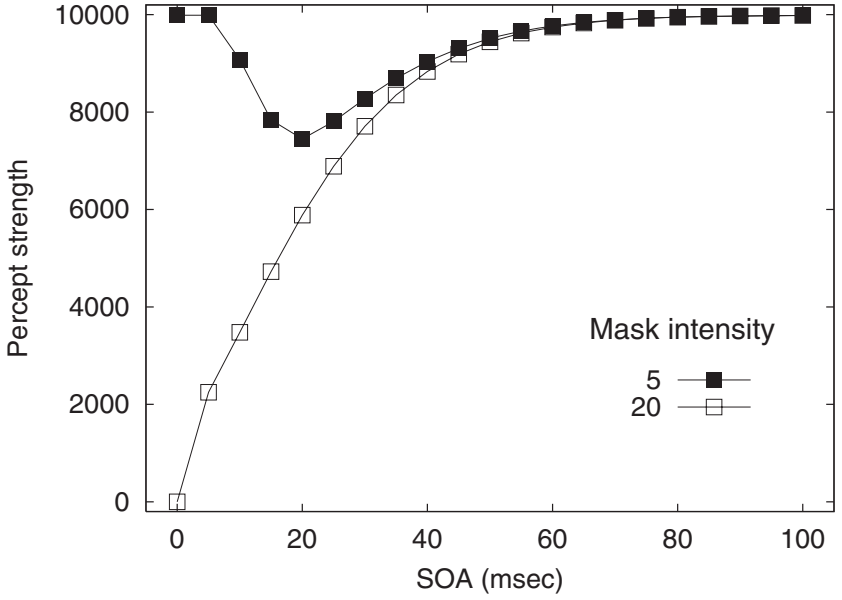

Figure 7.2
Masking functions for the mask-blocking system. The monotonic masking function occurs for the high-intensity mask, while the U-shaped masking function occurs for the low-intensity mask. A property of these systems is that, at each stimulus onset asynchrony (SOA), a monotonic masking function must lie below a U-shaped masking function.

Consistent with the properties of mask-blocking systems, Kolers (1962) noted that monotonic-shaped masking functions appeared for high-energy masks and U-shaped masking functions appeared for masks that were similar in energy to the target. Subsequent work has shown that this observation holds for a variety of mask types, including pattern masks that have contours overlapping the target (e.g., Hellige et al., 1979; Spencer & Shuntich, 1970; Turvey, 1973), metacontrast masks that do not overlap the target (e.g., Schiller, 1965; Weisstein, 1972), and masks consisting of a homogeneous field or disk that covers the target (e.g., Stewart & Purcell, 1974).

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 the mask signals. In a mask-blocking system, 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. Thus, all mask-blocking systems make a common prediction: For a fixed target, task, and SOA, any kind of masking situation that produces a U-shaped masking function must have weaker masking than any kind of masking situation that produces a monotonic-shaped masking function.

7.3 Other Explanations

It may be worthwhile to compare the predicted relationship between mask strength and masking function shape against other explanations of masking function shapes:

1. *Target property.* Weisstein (1972) noted that U-shaped masking functions appeared when the target and mask had roughly the same intensity but that decreasing the target intensity led to monotonic-shaped masking functions. More generally, one could hypothesize that the shape of the masking function is solely related to the target stimulus, with some target stimuli producing U-shaped masking functions and other target stimuli producing monotonic-shaped masking functions.

2. *Type of mask.* Several researchers have noted (e.g., Enns & Di Lollo, 2000) that some types of masks seem to produce U-shaped masking functions and other types of masks tend to produce monotonic-shaped masking functions. For example, noise masks, pattern masks, and light masks often produce monotonic-shaped masking functions, while metacontrast masks (where the mask shape does not spatially overlap with the target shape) tend to produce U-shaped masking functions.

3. *Level of masking.* Turvey (1973) suggested that monotonic-shaped masking functions indicated low-level (e.g., retinal) masking, while U-shaped masking functions indicated high-level (e.g., cortical) masking.

4. *Source of mask inhibition.* Breitmeyer and Ganz (1976) and Breitmeyer (1984) proposed that U-shaped masking functions appeared because of differential delays in sustained and transient streams of the visual system. Strongest masking occurred when the transient inhibition of the mask overlapped with the sustained information of the target. Since the transient system was faster than the sustained system, the ideal overlap occurred when the mask followed the target. However, monotonic-shaped masking functions could occur if the mask had a strong enough sustained signal to allow for sustained-on-sustained inhibition. It follows that in this theory monotonic-shaped masking functions should correspond to stronger masking than U-shaped masking functions because the monotonic-shaped masking functions are the result of additional inhibition. The quantitative model of Purushothaman et al. (2000) uses this approach to explain various masking effects.

5. *Type of masking effect.* Some researchers (e.g., Eriksen, 1966; Haber, 1969) have proposed that there are at least two different masking effects: integration and interruption. *Integration masking* occurs when the target and mask temporally integrate and the target becomes difficult to process. Such masking produces monotonic-shaped masking functions because integration is less likely to occur as the target and mask are separated in time. *Interruption masking* occurs when the mask

interrupts the processing of the information about the target. Such masking has been hypothesized to produce U-shaped masking functions.

6. *Integration effects.* A different interpretation of the effect of integration has been proposed by Navon and Purcell (1981) and Reeves (1982). They proposed that integration of the target and mask's would sometimes still allow the target properties to be judged well. In this view, the decrease in target visibility/identifiability as SOA increases is the result of the target and mask's being less likely to integrate into a single percept. Thus, contrary to the previous explanation, U-shaped masking functions would be evidence of integration at short SOAs, while monotonic-shaped masking functions would indicate that such integration either had no effect, had a deleterious effect, or did not occur at all.

Clearly, these alternative explanations are conceptually quite different from the strength-based explanation that makes up the explanation for all models that are based on mask blocking. The next section describes an experiment that tests the prediction of the mask-blocking systems and simultaneously turns out to investigate many of the alternative explanations as well.

7.4 Experiment 1: Backward Masking

We investigated the relationship between U-shaped and monotonic-shaped masking functions with an experiment that varied the properties of the target and the mask stimuli. Figure 7.3 schematizes the target and mask stimuli (in reverse contrast). The target frame consisted of four elements (one target and three distractors) arranged on the corners of a virtual square measuring 12.06° on each side. There were two types of target frames: sparse and filled. The sparse distractor stimuli consisted of five dots arranged with four dots on the corners of a 0.92° by 0.92° square and an additional dot in the middle. The sparse target stimulus was similar, but was constricted horizontally to make a vertical rectangle (0.57° by 0.92°). The filled stimuli were similar, except an outline square and rectangle were drawn for the distractor and target stimuli, respectively, and there was no central dot. The target frame was displayed for one frame of the 60 Hz monitor (approximately 17 ms).

Around each target and distractor stimulus a mask stimulus was drawn that was either sparse or filled. The sparse masks were four dots drawn on the corners of a 1.43° by 1.43° square. The filled mask was an outline square (thickness of 0.17°) of the same size. The mask stimuli were shown for two refresh frames of the monitor (approximately 34 ms). Both 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 are for white or black fields that completely covered the measurement area of a light meter.

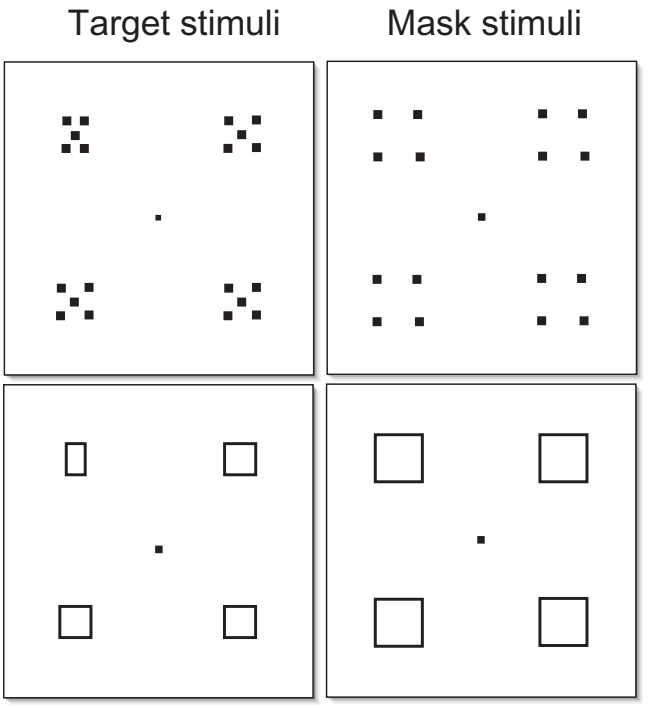

Figure 7.3
The target and mask stimul; used in the backward masking experiment (in reverse contrast). Each target type was paired with each mask type separately.

The observer's task was to report the location of the (virtual) rectangle in the target frame. Each of the four target–mask combinations was constant for a given session. The experiment consisted of 12 sessions, each with a 20-trial practice block and three 90-trial experimental blocks. Each block included five SOAs (0, 17, 34, 51, and 68 ms) between the target and the mask. Two observers participated. One observer was the second author and the other was naive to the purpose of the experiment.

The choice of stimuli were derived from an earlier study (Cho & Francis, 2003) that compared masking functions with a sparse mask and an outline mask. That study found that a sparse mask produced a monotonic-shaped masking function while the outline mask produced a U-shaped masking function. The current study additionally varied the properties of the target stimuli.

7.4.1 Results

The key question to be resolved is whether the shape of the masking functions is related to properties of the target, the mask, or target–mask combinations. Figure 7.4a–d plots masking functions for the different conditions for each observer. There are substantial overall differences between the observers, which is consistent with earlier studies (e.g., Weisstein & Growney, 1969); nevertheless, both observers show a similar pattern of results.

Figure 7.4a–b plots the masking functions for the sparse (dot) target and distractors. The curve with the filled elements is for the sparse (four-dot) mask, and the curve with the open elements is for the filled (outline square) mask. For both observers, the sparse mask produced a U-shaped masking function and the filled mask produced a monotonic-shaped masking function.

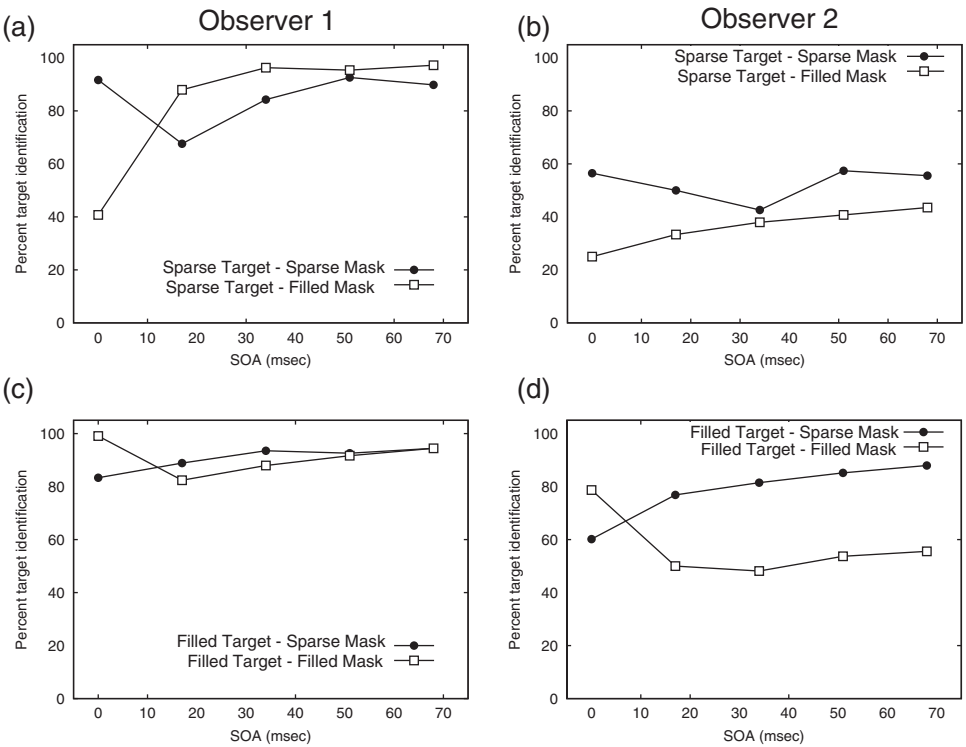

Figure 7.4
Results from experiment 1 for two observers. (*a*)–(*b*) Masking functions for the sparse target and both types of masks. (*c*)–(*d*) Masking functions for the filled target and both types of masks. SOA, stimulus onset asynchrony.

Figure 7.4c–d plots the masking functions for the filled (outline square) target and distractors. The curve with the filled elements is for the sparse (four-dot) mask, and the curve with the open elements is for the filled (outline square) mask. For both the observers, the sparse mask produced a monotonic-shaped masking function and the filled mask produced a U-shaped masking function.

7.4.2 Discussion

The experimental data do not agree with the prediction of the mask-blocking systems. Moreover, the data are also incompatible with many alternative explanations of the appearance of U-shaped and monotonic-shaped masking functions.

With regard to the mask-blocking systems, the models that use mask blocking predict that, with a fixed target and task, at each SOA a monotonic-shaped masking function should always lie below a U-shaped masking function. This is violated for both observers (see Figure 7.4a, c, and d) where the monotonic-shaped masking function is sometimes above the U-shaped masking function. The one case consistent with the prediction (see Figure 7.4b) does little to obviate the conclusion that the models are invalid because the prediction is that the relationship between masking strength and masking function shape must *alway*s hold. The existence of a counterexample is enough to negate the prediction. Moreover, the frequency of the violation of the prediction suggests that the violation does not involve a particularly unusual circumstance. This conclusion agrees with and extends the experimental findings of Francis and Herzog (2004).

As noted above, the data are also incompatible with some of the alternative explanations of the appearance of U-shaped and monotonic-shaped masking functions. For much the same reasons that the mask-blocking systems are rejected, the transient–sustained source of inhibition explanation is rejected because monotonic masking functions sometimes produce weaker masking than U-shaped masking functions.

The data are incompatible with the idea that masking function shape is related solely to the properties of the target. A sparse/filled target can produce either a monotonic-shaped masking function or a U-shaped masking function, depending on the type of mask that is used. Likewise, the data are incompatible with the idea that the masking function shape is related solely to the properties of the mask. A sparse/filled mask can produce either a monotonic-shaped masking function or a U-shaped masking function, depending on the type of target that is used.

The data also challenge the level-of-masking hypothesis. Whenever there is a monotonic-shaped masking function, this hypothesis would hold that low-level masking effects were occurring. However, if strong low-level inhibition exists for a

given mask, it is difficult to explain how changing the target stimulus would somehow make such low-level inhibition ineffective (which would be necessary to produce a U-shaped masking function).

It should be emphasized that this result does not mean that none of the explanations ever apply to backward masking effects. It is possible that they do apply under some conditions. The data do allow us to conclude that none of these explanations apply to the particular conditions of this experiment, and so these explanations do not provide a *complete* description of backward masking.

Thus for the conditions of this experiment, the only remaining viable explanations involve integration and interruption masking. There are two such explanations. One is based on the hypothesis that temporal integration of the target and the mask produces a type of masking by camouflaging the target. Such camouflage can produce masking at the shortest SOAs and thus a monotonic-shaped masking function. The other explanation suggests that temporal integration of the target and the mask does not produce strong masking but leads to visibility of some features of the target at short SOAs. In this explanation integration of the target and mask corresponds to a U-shaped masking function.

We prefer a generalization of these two proposals. When the target and the mask temporally integrate, we imagine that the unified percept can sometimes render the target more difficult to identify and sometimes render the target easier to identify. Thus, we hypothesize that performance on the masking task at the shortest SOAs, which determines whether a U-shaped or monotonic-shaped masking function is found, is related to the ability of the observer to perform the experimental task on the integrated target and mask combination. We tested this hypothesis in a second experiment.

7.5 Experiment 2: Visual Search

To validate the hypothesis that performance on the masking experiment at the shortest SOAs is related to the temporal integration of the target and mask stimuli, we conducted a visual search experiment with similar stimuli. The four possible combinations of target and mask stimuli were used to make a set of displays for a visual search experiment. Figure 7.5 shows (in reverse contrast) samples of the displays for each target–mask combination. In addition, displays were created that did not include a target but instead consisted of four distractor elements. The observer's task was to judge as quickly as possible whether or not a target element was present. On each trial one of the displays appeared and remained visible until the observer made a target-present or target-absent response by pressing the appropriate key on a keyboard. Reaction time was recorded, and incorrect responses were

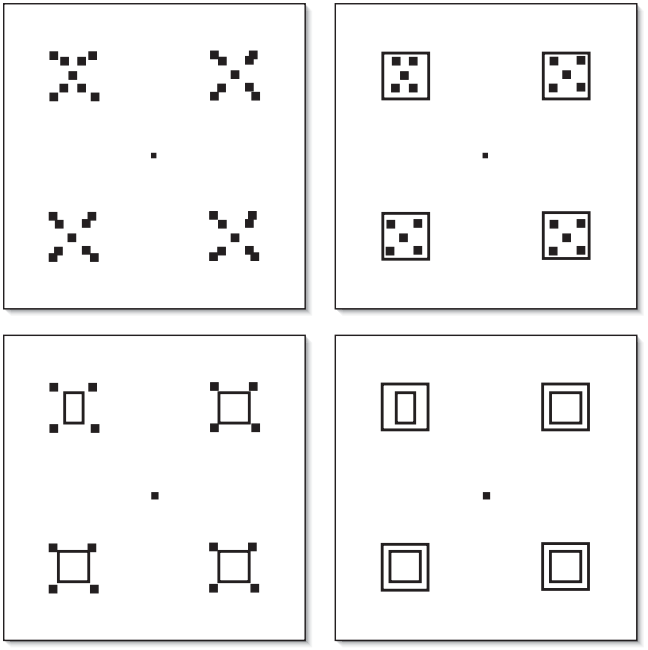

Figure 7.5
Sample displays for experiment 2 (in reverse contrast). The displays are the same as for the backward masking experiment, except the target and mask are superimposed.

discarded from further analysis. The same observers as in experiment 1 participated in experiment 2.

If performance at the shortest SOAs in the backward masking experiment was determined by the temporally integrated percept of the target and mask stimuli, one would expect that reaction times would be longest for those target–mask combinations that produced monotonic-shaped masking functions and shortest for those target–mask combinations that produced U-shaped masking functions. On the other hand, if the difference in masking function shape was due to some other form of interaction between the target and mask, then there seems to be no reason to think that performance on the visual search task would be related to performance on the backward masking task.

7.5.1 Results

Figure 7.6 plots the percentage correct identifications of the target for SOA = 0 in the masking experiment against the reaction time in the visual search experiment. The individual points are for the different target–mask combinations. The points lie

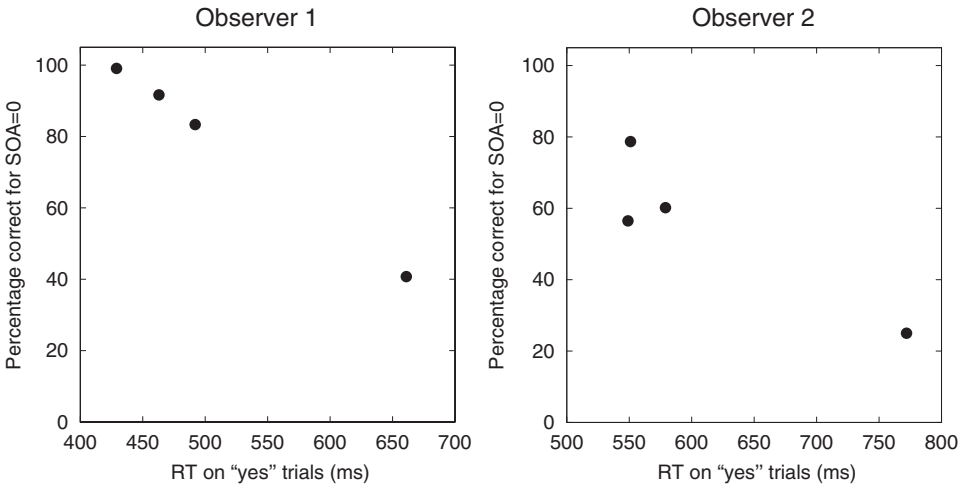

Figure 7.6
Comparison of target identification in the backward masking task and reaction time (RT) in the visual search task for different target–mask combinations. SOA, stimulus onset asynchrony.

nearly on a straight line, which indicates that the two variables are closely related. The correlation between reaction time and percentage correct was –.99 for observer 1 and –.91 for observer 2. Both of these correlations are statistically significant ($p < .005$ for observer 1 and $p < .05$ for observer 2).

7.5.2 Discussion

The data from the visual search experiment are correlated with the data from the backward masking experiment at the shortest SOA. This result is consistent with the hypothesis that performance on the backward masking experiment is determined by the ability of the observer to extract information about the target when it temporally integrates with the mask.

7.6 Conclusions

An analysis of quantitative models of backward masking based on mask blocking led to a prediction regarding how variation in the properties of the mask might change the shape of the masking function. An experiment to test this prediction demonstrated that it was not valid. This is a significant finding because these models have successfully explained a large variety of data on backward masking. The failure of the prediction to hold also suggests that the key limitation of these models is that

the interactions between the target and mask representations are fundamentally strength based. It is this characteristic that leads to the prediction that monotonic-shaped masking functions should correspond to stronger masking than U-shaped masking functions.

Indeed, other computational models that do not use the mask-blocking approach turn out to face similar problems with this set of data. The model of Purushothaman et al. (2000) does not use mask blocking computations, but it still makes the same flawed prediction regarding masking function shape. A quite different model proposed by Francis (2000, 2003a) also does not use mask blocking but also makes the same flawed prediction. Finally, if parameters are set appropriately, a quantitative model of reentrant processing proposed by Di Lollo et al. (2000) can also produce a U-shaped masking function (Francis, 2003b). However, with the parameters used so far, the model makes the same prediction regarding masking strength and masking function shape as the other models. This model has not undergone a full analysis, though, so there may be parameter sets that produce a different pattern of results.

The experimental result is particularly interesting because the various quantitative models of backward masking have been characterized as being fundamentally different from each other. Some of the models involve recurrent or reentrant feedback, while other models have a strictly feedforward flow of information. Some of the models are based on complicated neural network architectures, while other models are described with a single differential equation. Some of the models are hypothesized to work at an object-level description of stimuli, while other models are hypothesized to describe low-level interactions involving lateral inhibition. Despite these substantial differences, all of the models make essentially the same prediction about the shape of the masking function and the overall strength of masking, and all of the models are wrong for essentially the same reason.

The experimental data simultaneously cause us to reject a variety of other explanations for the differences between U-shaped and monotonic-shaped masking functions. The only alternative that seems to remain viable is to consider the role integration of the target and the mask might have at the shortest SOAs. We explored this possibility with a second experiment that provides an independent measure of the ease of identifying the target when the target and the mask stimuli produce a combined percept. The percentage of correct identification at the shortest SOA in experiment 1 correlates with reaction time in experiment 2, which suggests that whether one obtains a U-shaped or a monotonic-shaped masking function depends on whether integration of the target and mask at the shortest SOAs leads to easy visibility of the target or camouflage of the target, respectively.

What remains unclear is whether the role of integration accounts for the shape of the masking function in every situation, or whether other mechanisms (such as mask blocking) sometimes also apply. A goal of future research will be to identify experiments that can test between these quite different methods of producing a U-shaped backward masking function.

Acknowledgments

Gregory Francis and Yang Seok Cho were supported by National Science Foundation Grant 0108905.