A reentrant view of visual masking, object substitution, and response priming

James T. Enns, Alejandro Lleras, & Vince Di Lollo
University of British Columbia

To appear in: H. Ogmen & B. G. Breitmeyer (Eds.) *The First Half Second: The Microgenesis and Temporal Dynamics of Unconscious and Conscious Visual Processes*. Cambridge, MA: MIT Press.

Correspondence: James T. Enns
Department of Psychology
University of British Columbia
2136 West Mall
Vancouver, B.C. V6T 1Z4
Canada

Phone: 604-822-6634
e-mail: jenns@psych.ubc.ca

Abstract

When a mask follows a briefly presented target there are several consequences. The one that has historically received the most attention is a reduction in the visibility of the target. This is the conventional definition of masking. Yet, another equally important consequence is that errors in target identification are biased toward the identity of the mask rather than being randomly distributed among the target alternatives. This is evidence of object substitution. Finally, when the target is a signal to make a speeded action, this action can be influenced by a prime stimulus that is not even visible to the participant. This is known as masked response priming. In this chapter we review evidence concerning all three of these consequences of viewing rapid visual sequences. We argue that these consequences are difficult to understand, either individually or together, as the consequence of strictly feed-forward processing in the visual brain. In contrast, when these results are considered from the perspective of reentrant visual circuitry, they are easier to understand and to relate to one another. Moreover, predictions derived from a reentrant view of the brain lead to unexpected and novel results that are confirmed when tested against psychophysical data.

Most theories of vision hold to the view that from the time the eye first encounters a scene to the time that conscious perception occurs the scene has been coded at several levels in the visual system. Moreover, most theories propose some form of interaction between the anatomically lower levels of processing involved in registering details of the scene (e.g., retina, lateral geniculate nucleus, area V1) and the anatomically higher levels involved in the meaningful categorization and conscious experience of the scene (e.g., extra-striate cortical regions including the temporal cortex). Some theorists have even referred to this as the distinction between "seeing" and "understanding," or between stimulus-driven and conceptually-driven processing (Broadbent, 1958; Hebb, 1949; Helmholtz, 1866/1967; Hochberg, 1968; Neisser, 1962).

But how do these two levels of visual processing interact? On one option, the influence is unidirectional, meaning that lower levels of processing feed their results forward to the higher levels, where the resulting processes are increasingly abstracted away from the original image (Hubel, 1988; Pylyshyn, 2003). The proposed benefit of such abstraction is that it permits perception to concern itself with the invariant or intrinsic attributes of the objects in the scene (i.e., their true color, their actual size) rather than with the momentary or superficial attributes of the image (i.e., the luminance and wavelength characteristics of light at the eye, their retinal extent).

However, on another option, the influences between levels of processing are bi-directional, such that processing consists of iterative exchanges of neural signals among levels (Damasio, 1994; Di Lollo et al, 2000; Hochstein & Ahissar, 2002; Lamme & Roelfsema, 2000; Mumford, 1992; Zeki, 1993). An initial wave of stimulation ascends rapidly through the system (feedforward processing), followed by descending signals between levels (feedback processing or reentry). Together, the ascending and descending pathways form an iterative-loop system, aimed at noise reduction and hypothesis verification, thereby establishing the most plausible perceptual interpretation of the incoming stimulus. Most notably, in this view, the role played by neurons at the lower levels differ considerably over time, depending on whether they are participating in the signaling of information about the first glance at a scene or whether they are participating in the iterative checking of hypotheses about the scene generated by neurons at the higher levels.

In the research we have conducted over the past few years, we have been increasingly attracted to the reentrant position, largely because of the improved understanding that comes from adopting the iterative reentrant perspective. We have found that when we try to incorporate the vast anatomical evidence for reentrant communication in the brain (Bullier, McCourt, & Henry, 1988; Felleman & Van Essen, 1991; Mignard & Malpeli, 1991; Zeki, 1993) into theories we use to account for psychophysical (behavioral) data, many of the apparent puzzles in these results can be explained by adopting a reentrant stance. Yet this is not to claim that adopting this approach comes without its own potential pitfalls and dangers. Constructing arguments about the structure of the brain on the basis of behavioral data is always an uncertain business.

In what follows we will summarize our recent research in three different areas of visual masking. In each section, we will first describe the ‘standard’ account of a phenomenon, based on taking an approach premised on unidirectional flow of visual information from lower to higher brain regions. We then describe experiments we have conducted that have led us to the reentrant position. In a final section we identify several emerging areas of research in which we hope to broaden the reentrant perspective beyond the scope of visual masking and priming.

A. Metacontrast masking: Thinking outside the box

Metacontrast masking has been defined as “… the reduction in visibility of one briefly presented stimulus, the target, by a spatially adjacent and temporally succeeding, briefly presented second stimulus, the mask.” (Breitmeyer, 1984, p. 4). Masking is known to occur within a narrow band of temporal intervals. When the stimulus-onset asynchrony (SOA) between the target and the mask is either very brief or very long, the target is clearly visible. At intermediate SOAs, however, perception of the target is impaired, leading to a U-shaped function of accuracy over SOA.

The main mechanism thought to be at work in metacontrast masking involves inhibitory interactions between neurons representing the contours of the target and the mask (Breitmeyer & Ganz, 1976; Macknik & Livingstone, 1998; Weisstein, Ozog, & Szoc, 1975). For example, in the model of Breitmeyer and Ganz (1976) the key idea is that the onset of each stimulus initiates neural activity in two channels: a transient channel, which has short latency and responds optimally to low spatial frequencies of fast-changing stimuli, and a sustained channel, which has longer latency and is attuned to figural aspects of the stimulus such as details carried by higher spatial frequencies. Activity in the transient channel is said to inhibit concurrent activity in the sustained channel. Masking occurs when the faster-acting signals triggered by the onset of the mask inhibit the activity of slower signals carrying information about the earlier target.

These "two-channel" theories are predicated on strict temporal and spatial relationships between the target and the mask. In the temporal domain, an optimal interval must elapse from the onset of the target to the onset of the mask: it is the onset of the mask that interferes with the processing of the earlier target. In the spatial domain, there are two main requirements: first, that the mask contain a substantial amount of contours and, second, that the contours of the mask be in close spatial proximity to the contours of the target (Breitmeyer, 1984; Growney, Weisstein, & Cox, 1977).

Common onset: a new masking paradigm

Research carried out over the past several years in our and other laboratories has revealed a new form of masking that is difficult to reconcile with either the temporal or the spatial requirements of the two-channel view. More important, the new form of masking is not predictable from a strictly feed-forward model of visual information processing. Here, we give only a brief introduction to this new form of masking in order to give the reader an indication of what needs to be explained. Interested readers can experience this masking effect first-hand through demonstrations on the Internet (http://www.interchange.ubc.ca/enzo/).

In a typical experiment (Di Lollo, Enns, & Rensink, 2000), the display sequence consisted of an initial brief pattern containing from 1 to 16 rings, each with a gap in one of four cardinal orientations. One of the rings (the target) was singled out by four small surrounding dots that acted as both cue and mask (Figure 1a). The display continued without interruption to a second frame that contained only the same four small dots for durations varying from zero to several hundred milliseconds (ms). Observers reported the orientation of the target's gap. A critical aspect of the display was that the target and the four dots appeared simultaneously in Frame 1. We have referred to this temporal contingency as the common-onset masking paradigm (Di Lollo et al., 2000)

The results, illustrated by the segmented lines in Figure 2b, revealed little or no masking when the display contained only one potential target or when the mask terminated at the same time as the target display, regardless of the number of potential targets. In contrast, pronounced masking occurred provided that both of two conditions were met: first, that the initial display contain multiple potential targets and, second, that the mask remain on view beyond the termination of the target display.

Unidirectional inhibitory models cannot account for masking obtained with the common-onset paradigm because those models were explicitly designed to explain why masking had never been found when the target and the mask were presented simultaneously (Alpern, 1953; Breitmeyer & Ganz, 1976; Matin, 1975; Weisstein et al., 1975). In light of the results in Figure 2b, we can now assert that masking with common onset had never been obtained in earlier studies because the target and the mask not only started together but also ended together. This temporal contingency corresponds to the condition in which the duration of the trailing mask was equal to zero, where no masking was obtained. Strong masking developed, however, as the duration of the trailing mask was increased, provided that the display contained a sufficient number of potential targets (Figure 2b). Thus, the critical factor in metacontrast masking is not whether the target and the mask have a common onset, but whether the mask remains on display beyond the offset of the target. A recent study has even shown that the critical variable is really the temporal factor of mask duration and not simply the additional stimulus energy that comes from prolonging a visual stimulus (intensity x time) (Di Lollo, von Mühlenen, Enns & Bridgeman, in press).

A second reason why inhibitory models cannot account for the results in Figure 2b is that the four-dot mask used in that experiment do not have a sufficient amount of contour to generate the required amount of inhibition. As noted above, the strength of classical metacontrast masking is known to decrease markedly as the amount of contour in the mask is reduced (Breitmeyer, 1984).

Object substitution: a new understanding of masking

We believe that an account of common-onset masking with a four-dot mask can be developed in the broader context of how the visual system handles rapidly sequential stimuli. In keeping with reentrant thinking, we suppose that conscious perception of a stimulus emerges from neural activity triggered not only by feed-forward signals but also by iterative exchanges between cortical regions connected by reentrant pathways. On the first cycle, the input from the initial display is encoded at a low level within the system and then proceeds to higher levels where tentative cognitive representations are produced. These representations are in need of confirmation because they may be incomplete and ill defined or, equivalently, the ascending signals might activate more than one representation at the higher level. This creates an ambiguity for perception that can be resolved by comparing the high-level codes with the initial pattern of activity at the lower level. Reentrant pathways enable such a comparison.

We assume that, on the second and later iterations, the high-level representation reenters the lower level and is compared to what is currently there. If the image on the screen has not changed, a match is found and processing continues. If the screen has been blanked after the initial display, there is nothing on the screen to compete with the decaying initial stimulus representation. Thus, for a short time a match can still be found with this decaying trace. If, however, the reentrant information does not match the current information at the lower level (i.e., if the input has changed), then a new tentative representation emerges and processing continues as for a new stimulus.

We propose that common-onset masking occurs when the initial display (target plus mask) is replaced on the screen by a different configuration (mask alone) before the required processing iterations have taken place for the target to be identified. Thus, the emerging percept of the compound image (target plus mask) is replaced in consciousness with the percept of the mask alone. It is as though the visual system treats the trailing configuration as an updated replacement of the earlier one (Kolers, 1972).

Implicit in this description is a general account of backward masking in which a process of perceptual updating plays a critical role. It is our view that this type of masking occurs when the emerging representation of a target is replaced by the emerging representation of a trailing mask as the object occupying a given spatial location. We refer to this process as masking by object substitution.

CMOS: a computational model for object substitution

We have incorporated these theoretical ideas in a computational model of masking by object substitution, CMOS, in short (Di Lollo et al., 2000). CMOS belongs to a class of models known as closed-loop controllers that are common in several areas of industry and robotics (e.g., Carpenter & Grossberg, 1987). In their most general form, these models describe a process in which some form of input is collected and coded before being sent on to an output device. However, the output device, in addition to generating an output signal, sends a copy of that signal back to the input device. This leads to another round of coding to facilitate the comparison of this feedback signal with the current signals entering the input device. In this way, the output signal is influenced in an ongoing way by both current input and information already processed.

The central assumption in the CMOS model is that visual perceptions emerge from the activity of three-layered modules such as illustrated in Figure 2, arrayed throughout the visual field. Each module can be conceptualized as a circuit involving the connections between cortical area V1 and a topographically related region in an extra-striate visual area. The output of each module is a representation of the spatial pattern within its receptive field.

Here, we give the reader only a general idea of how the model operates (details are given in Di Lollo et al., 2000). The onset of a visual stimulus triggers the first of several cycles of activity in the three layers illustrated in Figure 2. The activity in the Pattern Layer is then fed back to the Working Space by means of a simple overwriting operation. As part of this transfer, pattern information is translated back to the pixel-like codes of the Input Layer, permitting a direct comparison. This comparison is necessary for resolving ambiguities and for achieving fine spatial registration between the reentrant code and the ongoing activity at the lower level.

Masking is produced in this model by the fact that the contents of the Input Layer change dynamically with new visual input. The contents of the Pattern Layer change more slowly because its input is a weighted sum of what is currently in the Input Layer and what was in the Working Space on the previous iteration. This produces a degree of inertia in response to changes in input that is an unavoidable consequence of reentrant processing. If the visual input changes during this critical period of inertia, masking will ensue. We refer to this process as object substitution because the emerging representation of a target in the Pattern Layer is replaced by the emerging representation of the mask as the object occupying a given spatial location.

The continuous lines in Figure 2b illustrate simulations from a formal implementation of CMOS. The model curves provide a remarkably good fit to the empirical data. Notably, the simulation captures the two fundamental factors that mediate object-substitution masking: attention must be distributed over multiple potential targets, and the mask must continue to be visible during the period in which the iterations between higher-level pattern representations and lower-level contour representations are likely to occur.

Several recent investigations have confirmed the occurrence of object-substitution masking under a variety of conditions, and have explored its limits (e.g., Giesbrecht, Bischof, & Kingstone, 2003; Jang & Chun, 2001a, 201b; Lleras & Moore, 2003; Neill, Hutchison, & Graves, 2002; Tata, 2002). Especially revealing is an investigation by Woodman and Luck (2003) who recorded event-related potentials to explore dissociations among attention, perception, and conscious awareness during object-substitution masking. They found that the target is identified by the visual system even when it is masked and, therefore, cannot be reported accurately. Under these conditions, however, the target triggers a shift of attention so that, by the time attention is deployed, only the mask remains visible, leading to impaired accuracy of target identification. These findings buttress the interpretation that object-substitution masking is critically dependent on reentrant processing.

B. Object substitution as a general theory of masking?

Although object substitution theory does a reasonably good job of accounting for masking in the common onset paradigm, and even metacontrast-like masking involving four small dots, a critic might legitimately wonder whether this theory can account for masking obtained in conventional paradigms, including metacontrast involving snugly-fitting surrounding contours and pattern masking involving overlapping shapes and visual noise. According to the reentry hypothesis, there is no difference in principle between masking with common onset and many aspects of classical metacontrast and pattern masking. All forms of backward masking should be subject to the influences of object substitution, in that the emerging representation of a temporally leading target will be replaced in consciousness by that of the mask if it follows the target closely in time and appears before target identification is complete. However, it is also possible that there will be differences in each form of masking, with for example, metacontrast masking producing specific types of contour interactions that are not shared by pattern masking or masking by four dots.

The approach taken to this question in a recent study (Enns, 2003) involved a systematic comparison of the effects of the four-dot mask with the 'classic' masks used in metacontrast, noise, and pattern masking. To accomplish this, a simple letter identification task and identical target-mask sequences were used to compare six different masking stimuli, as illustrated in Figure 3. The distribution of attention was manipulated by varying the potential number of targets randomly between 1, 4, and 7 items. The temporal relation between target and mask was varied from –150 ms (mask before target) to +600 ms (target before mask). In Experiment 1 the mask itself acted as the target 'probe,' indicating to the observer which letter was to be identified. This meant that for all the positive temporal intervals, the letter to be identified was only indicated after the mask had been presented, preventing spatial attention from focusing on the target prior to the arrival of the mask. In subsequent experiments a spatial cue was placed near the target location, either simultaneously with the target display (Experiment 2) or 100 ms prior to its arrival (Experiment 3).

Three critical predictions were derived from the reentrant view. First, under conditions of spatially distributed attention (Experiment 1, where the mask was also the target probe) the reentrant hypothesis predicts that all forms of backward masking should have, at a first approximation, an equal effect on target accuracy. This prediction derives directly from the idea that the contents of the mask will replace those of the target if it has not been identified prior to its replacement on the screen by the mask. Second, the differences among masking effects should be greatest when the target and mask are in close spatiotemporal proximity and are therefore temporally integrated with one another (all three Experiments). Under these conditions, the problems of target identification will not concern those of object substitution. Rather, they will involve problems of camouflage, in the case of noise and pattern masking, and local contour interactions, in the case of metacontrast masking. Third, backward masking will be minimized when attention can be pre-focused on the spatial location of the target (Experiment 3). This prediction follows directly from the idea that if target identification can be completed before only the mask remains on view, object substitution will not occur.

Target identification accuracy in Experiment 1 (Figure 4) showed that all backward masks had very similar effects, provided that attention was distributed (display size = 7). The primary consequence of each of these masks under distributed attention conditions was to reduce target accuracy to the level obtained in the baseline (dot probe) condition, but to do so within a target-mask interval of 50 ms to 150 ms rather than the 600 ms that are required when no backward mask replaced the spatial location occupied by the target.

However, there were two notable exceptions to this general trend of equal backward masking for all mask types. One concerned the four-dot mask. Unlike the other four masks, which had their maximum influence at an interval of 50 ms and beyond, the four-dot mask had its full effect only at intervals of 150 ms and longer. This suggests that although backward masks are equal in their effects at intervals of 100 ms or more, they contain important differences in their effects at shorter intervals, as per the second prediction we derived from the reentry theory for this paradigm. This deviation from the general trend is therefore in keeping with the proposal that masking has at least two distinct ways to influence target identification: an early or fast-acting component associated with object formation and a later or slower-acting component associated with object substitution.

A second deviation from the general pattern was seen in two of the mask types, digits and letters, which reduced target accuracy much more severely than the simple decay of information that occurs without a backward mask (dot probe conditon). Although this appears superficially to be a violation of the prediction derived from reentry, a closer look reveals that the specific errors made in these two cases are actually consistent with it. Recall that the theory states that if only the mask remains on view prior to the complete identification of the target, processing will become focused on the mask. This means that if the mask itself activates target-relevant features or properties, these features may come to control the observer’s response. This is exactly what happened. An examination of the responses made when targets were incorrectly identified indicated that many of them were related to the target-relevant features in the mask rather than being random. For digit masks this meant that, for example, that the digit 4 led to target responses that were visually similar to the digit (e.g., A); for letter masks, the identity of the mask was often incorrectly reported as the target letter. In both cases the mask seemed to be replacing the target as the object of conscious report by the observer.

Target identification accuracy in Experiment 3 (Figure 5) showed that all forms of backward masking have minimal effects if spatial attention can be focused on the target location prior to the target-mask sequence. This prediction derives from the idea that object substitution will not occur if the target shape can be identified prior to the time that only the mask shape remains on view, which will then become the focus of the identification processes. However, as in Experiment 1, substantial masking was observed for all masks other than the four dots in the range of intervals between –50 ms and +50 ms. Strong masking occurred even when there was only a single display item and when a spatial precue indicated with 100% certainty where this item would appear. In this range of intervals, the focus of spatial attention had little influence on target identification. This attention-insensitive aspect of masking is consistent with the idea that temporal integration and local contour interaction are influenced relatively little by the distribution of spatial attention.

This comparison of the effects of spatial attention (display size and spatial cuing) on various types of visual masks (metacontrast, random noise, four dots, and patterns) provides strong support for the idea, derived from the object substitution theory of masking that there are two distinct visual masking processes at work in a typical masking paradigm.

The first process is active in the range of target-mask intervals of 0-100 ms. It interferes with object formation, and at least in the case of pattern masking, this seems to be through the mechanism of temporal integration. Although it is the formal task of the observer to identify the target object (letter), when targets and masks are presented in close temporal proximity, the first ‘object’ formed by the visual system is actually a composite of the target and mask patterns. For some masks, such as the four dots, such a composite object interferes little if at all with the task of identifying the target letter. However, for other masks, such as the random dots, digits and letters, the fusion of target and mask slows down target identification. For this early process, the effects of the mask are essentially those of 'camouflage,' which must be removed or segmented before the target can be identified.

This temporally early masking process is influenced very little by whether spatial attention is widely distributed, equally ready to select any one of eight different possible target locations, or whether it is already narrowly focused on only a single display location. This can be seen by comparing target accuracy for the smallest set size in both experiments (display size = 1 in Figures 4 and 5). The data points for each masking condition are almost all identical when this comparison is made, illustrating the point that the object formation process is uninfluenced by the spatial focus of attention and that it is over by the time 150 ms has elapsed between the presentation of target and mask.

The second process, masking by object substitution, is critically dependent on the existence of a temporal delay between target presentation (a physical event initiating visual processing) and target identification (a mental achievement). This is evident in three critical features of the data. First, when attention was widely distributed at the onset of the target display, there was no backward masking in effect for a single display item beyond a target-mask interval of 100 ms (Figure 4). This held true whether the mask was a metacontrast frame, random noise, or even a competing shape such as another letter. This means that aside from the expected masking effects during object formation, there were no additional masking effects that arose from a delay in selecting the correct display item for perceptual report. Second, the four dot mask yielded no evidence of any masking during the object formation stage, either when attention was distributed (Figure 4), or when attention could be focused earlier on the target location (Figure 5). Evidently, four dots pose no significant 'camouflage' problems for target identification. Finally, the absence of all backward masking effects for target-mask intervals of 150 ms and more when the target location was cued in advance (Figure 5), is consistent with object substitution masking depending critically on a delay between display presentation and the identification of a specific item in that display.

Taken together, these results suggest that a mask will interfere with target identification only when each of two conditions are satisfied: (a) The mask must be presented prior to the completion of target identification, and (b) A temporally integrated target-mask composite must interfere with target identification. Masks that fail to satisfy both of these conditions will be ineffective. This was the case for the common onset four-dot mask that terminated with the target (in Section A of this paper and in the present experiments at target-mask intervals of 0 ms). These stimulus sequences satisfy condition (a) but not condition (b). The converse was the case for experiments described in the present section with digit and letter masks presented 150 ms or more following a precued target letter. These tests satisfied condition (b) but not (a). In summary, the reentrant hypothesis leads to a unification of the principles for many different kinds of masking that have not been anticipated previously when researchers have taken a unidirectional approach to theorizing about visual masking.

C. A cautionary tale of mask-dependent response priming

Backward masking has employed in myriad studies to date and much is known about the relevant factors involved (Breitmeyer & Ogmen, 2000; Enns & Di Lollo, 2000). Yet its most frequent use by vision researchers has not been to better understand why it occurs. Instead, it has been used most often as a tool of convenience to reduce the visibility of a stimulus. But its use as such a tool is premised on masking being understood as a strictly feed-forward process; one in which the mask simply interrupts the processing of the target. In this view, which has become quite standard, processing of the target has simply been abandoned upon presentation of the mask, in favor of processing of the mask.

For example, research in visual priming often uses a backward mask to reduce the visibility of a stimulus referred to as the prime. In a typical masked priming procedure, a prime shape is first presented briefly, followed by a mask shape which renders the prime very low in visibility, and then these shapes are followed by a target shape which remains on view and must be responded to rapidly. The goal of this procedure is to determine whether and how the advance information conveyed by the prime influences the processing of the later-occurring target. Within a feed-forward view of visual processing, it therefore seems reasonable to assume that the chief determinant of priming is the prime shape. The mask shape only indirectly influences target processing by weakening or strengthening the information carried by the prime shape. In what follows, we will show how misguided such an assumption can be and how the concept of object updating through reentrant processing can shed light on a type of masked priming that is often claimed to index the inherently "inhibitory" nature of visual unconscious processing.

The Negative Compatibility Effect (NCE) was first described by Eimer & Schlaghecken (1998) and has since been the subject of much research (Eimer & Schlaghecken, 2001, 2002; Eimer, Schubo & Schlaghecken, 2002; Klapp & Hinkley, 2002; Lleras & Enns, 2004; Schlaghecken & Eimer, 2001, 2002; Verleger, Jaskowski, Aydemir, van der Lubbe & Groen, 2004). In a typical experiment, a brief prime is first presented (a double arrow pointing either right or left), followed by a mask (typically, a pattern mask composed of superimposed right- and left-pointing double arrows), which is then followed by a target (a double arrow pointing either right or left). The participants' task is to indicate as rapidly as possible in which direction (right or left) the target arrow is pointing. NCE refers to the counterintuitive finding that participants are reliably faster to respond to the target when prime and target point in opposite directions (incompatible trials) than when prime and target point in the same direction (compatible trials).

Most interpretations of the NCE have relied on feed-forward models of perception, where the role of the mask is confined to reducing or even eliminating perceptual awareness of the prime (Eimer & Schleghecken, 2002; Klapp & Hinkley, 2002). These authors claim that when perceptual awareness of the prime is limited, the prime is processed by unconscious mechanisms that are intrinsically inhibitory in nature. From this perspective the NCE arises because a low visibility prime activates the unconscious inhibitory processes associated with it, causing the response to an identical visible target to be inhibited relative to a target corresponding to the opposite response.

In a recent paper, we called these unconscious inhibition theories of the NCE into question (Lleras & Enns, 2004). We began by noting that most previous studies had used pattern masks composed of the superimposition of both possible primes. In those studies where different masks were used, the masks still contained what we referred to as "task-relevant features" (i.e., arrow-like features pointing right and left) rather than task-irrelevant features (Klapp & Hinkley, 2002; Schlaghecken & Eimer, 2000). To test what role these kinds of masks may have played in producing the NCE, we conducted a series of experiments in which masks containing task-relevant features and task-irrelevant features were compared. Figure 6 shows a summary of the priming effects observed under both conditions. As can readily be seen, relevant masks systematically produced negative priming effects (replicating previous studies on the NCE), whereas irrelevant masks systematically yielded positive priming effects.

These results directly refute unconscious inhibition theories of the NCE because in such accounts "a mask is a mask is a mask." That is, the type of mask used should not affect the direction of the priming effect, provided that it reduces the visibility of the prime sufficiently to activate the unconscious inhibitory processes associated with the subsequent target. Only prime visibility should determine priming magnitude, with low levels of visibility yielding a NCE and high levels of visibility a positive priming effect. But this was not what we found. We found that completely opposite priming effects could be obtained at equal levels of high or low prime visibility. What determined the direction of the priming in both cases was whether the mask contained task-relevant (negative priming) or irrelevant features (positive priming).

Unconscious inhibition theory falls far short of explaining these data. What we proposed instead was based on the concept of object updating, as presented in the preceding sections. At the core of this proposal is the possibility, derived from our previous studies of masking by object substitution, that the prime and mask may be interpreted by the visual system as changes in the same object over time. As a result, the prime-mask stimulus sequence is susceptible to the normal processes of object updating and a single object representation may represent both stimuli as one object changing over time. When the prime first appears (say a left arrow), an object representation is established and attributes of the prime are encoded and associated with it. Notably, if task-relevant attributes are encoded, these attributes will start priming their associated response (a left response, in this example). When the mask appears afterwards (say a juxtaposition of left and right arrows), the same object representation is updated with information from the mask. Because some of the attributes of the mask are already represented in this object representation (those corresponding to the left arrow embedded within the mask, and also present in the prime), updating will primarily involve those “novel” features that are present in the mask, but are absent from the prime (those corresponding to the right arrow features embedded in the mask).

In accordance with previous literature on masked priming, we therefore proposed that priming is determined by the most recent set of task-relevant features encoded at the object level of representation, provided that strong links already exist between those features and their corresponding response (see Neumann, 1990 and Scharlau & Ansorge, 2003). This assumption, in combination with object updating, is sufficient to account for the NCE observed with relevant masks as well as for the positive priming observed with irrelevant masks. On relevant-mask trials, the most recent set of task relevant features detected prior to target onset are those present in the mask and absent from the prime, therefore creating an opposite-to-prime facilitation. This opposite-to-prime facilitation results in faster responses to the target when the prime is opposite in shape (because the novel features in the mask point in the same direction as the target) and to slower responses to the target when the prime is identical (the novel features in the mask now point the a direction opposite to the target). Although the observed priming effect may superficially appear to be negative with respect to the prime, it actually corresponds to a positive priming effect induced by the novel task-relevant features in the mask. In sum, the observed priming effect reflects the influence not of the prime alone but of the prime-mask bundle on response selection. Correspondingly, on irrelevant-mask trials, the most recent set of task-relevant features encoded prior to target onset are those present in the prime (the mask has none of these features). Thus, the observed priming effect truly reflects the positive influence of the prime on response selection to the target.

We showed that this object updating account could even be applied to masks that did not contain exact replicas of the primes, but merely contained similar features in spatially non-corresponding locations to those in the target (Lleras & Enns, 2004). We also tested this account in an experiment in which the roles of the prime and mask were reversed. For example, in one experiment either the superimposed set of double arrows (task relevant) or the double-pound sign (task irrelevant) were used as a prime and the double-headed arrows were used as a mask. Although neither of these primes could now bias a particular response, the priming observed from the masks was influenced by these response-neutral primes. Task relevant primes led to more priming than task irrelevant primes, consistent with the prime-mask sequence contributing jointly to the priming effects on the target.

Standard feed-forward accounts of masking would predict no influence of the prime type in this experiment. Priming ought to be determined solely by the mask-target sequence, which was identical in both types of trials. However, the object updating account predicts that more priming ought to be observed in the relevant prime condition, than in the irrelevant prime condition. This follows because, if prime and mask are associated with the same object representation in the relevant prime condition, the processing of the task-relevant priming features will begin at the onset of the prime. This is slightly earlier in time than in the irrelevant prime condition. Thus, by the time the “mask” appears in the relevant prime condition, any further updating of the priming features in the mask can only add strength to the representation that has already begun to be formed. In contrast, in the irrelevant prime condition, the object updating process could only begin to represent the target-relevant information at the onset of the mask. The results were consistent with these predictions: task-relevant primes resulted in a significantly larger priming effect than task-irrelevant primes.

In sum, these experiments ofhelp to demonstrate that (1) the behavioral consequences of masking should always be examined directly rather than simply being assumed; (2) object updating is a powerful concept that can be applied to studies of unconscious processing of stimuli that are reduced in visibility, as well as to studies of response priming to objects with strongly associated response tendencies; and (3) the object updating concept produces testable predictions and can help to explain some behavioral phenomena with a simpler set of assumptions than those needed when standard feed-forward theories of perception are applied.

D. The future of reentrant theorizing

In the preceding sections of this paper we have summarized the consequences of our reentrant theorizing in the realm of masking and masked priming. But we are growing more confident every day that the reentrant hypothesis does not apply only to this restricted set of laboratory phenomena. Even in everyday vision, it is our view that the perception we are able to accomplish in a momentary glance is influenced by the perception we have of the larger schema of the scene that we have recently formed or have simply assumed to be true. The processes of perception in a glance will automatically evoke hypotheses about what lies beyond that glance, leading to a bias in where the next glance will be placed. In turn, having the expectation of particular schema will lead to the testing of specific hypotheses in subsequent glances. In other words, vision is almost always influenced as much by what lies in the mind as what lies before the eyes of the observer. Here we will briefly point to two examples of phenomena illustrating the point that these recursive aspects of vision are as relevant in everyday visual functioning as they are in visual masking and priming.

Beyond the limits of the attentional blink

Everyone who watches television is aware that not all the images that flash before our eyes are processed to an equal level of comprehension and awareness. Rather, we are aware that the cost of attending to a specific image in detail is a loss in comprehension of the images that follow, at least for a brief period of time. In the laboratory, this phenomenon is known as the <u>attentional blink</u>, and it is measured by inserting two visual targets into a stream of visual distractors that are displayed in rapid serial visual presentation (RSVP). When the lag between the two targets is varied systematically in steps of 100 ms or so, the first target can be identified very accurately, but at the cost of a severe reduction in second-target accuracy. This is the attentional blink. It is commonly largest when the inter-target lag is short and it diminishes gradually as lag is increased to 500 ms or more.

The standard feed-forward theoretical account of the attentional blink proposes some form of limited attentional resource that is allocated to the leading target at the expense of the trailing target. We have recently proposed an alternative account based not on depletion of a limited resource but on a temporary loss of perceptual control over the processes of perception in a glance (Di Lollo, Kawahara, Ghorashi & Enns, in press). The main idea is that processing of initial items in the visual stream is governed by a hypothesis testing mechanism that is configured to pass target items and exclude non-target items. This 'input filter' is maintained by reentrant signals from a central processor that can perform only one function at a time. As soon as the central processor has detected an item from the target set and has begun to identify it, the maintenance signals for this initial monitoring task are discontinued, and the input filter comes under exogenous control by the next items in the visual stream. Therefore, if the next item belongs to the same category as the first target, the filter's configuration remains unaltered, and the item can be processed efficiently. If, however, the next item belongs to a different category from the first target, it will take longer to process than a same-category stimulus because it does not match the configuration of the input filter. This delay makes the trailing item vulnerable to masking by the next item in the stream. In addition, the configuration of the input filter will be altered, so that even ensuing items belonging to the same category as the first target will not be processed efficiently.

This reentrant account of the perception of rapid visual sequences makes the bold prediction that no attentional blink will occur if two or even three items from the target class are presented in succession. This is in sharp contrast to the standard theories that predict a decrease in accuracy for targets that follow the first in rapid succession. Although some of these theories try to accommodate the result that up to one additional target may occasionally sneak through the "sluggish attentional window" used to process the first target, none of the theories predict that up to three targets in a row will be processed efficiently. Yet this is exactly what the data show.

We conducted a series of experiments comparing report accuracy for items in a visual stream with three successive targets (which was uniformly high) with two targets separated by a single distractor (resulting in the attentional blink for the second target) (Di Lollo et al, in press). The pattern of results revealed that it was the appearance of the distractor, during the period when the participant was identifying the first target, that inadvertently caused the input filter to become reset to the distractor class of items. Thus, the cognitive limit measured in the attentional blink is not one of a failure to be able to identify multiple items appearing in succession, but it is instead a failure to be able to control the hypotheses that are currently being tested in the information seen in a glance.

Using memory to resume an interrupted search

Think of the last time you drove or walked down a busy commercial street while searching for a friend that you had arranged to meet. You didn't know exactly where you would find your friend so you alternated between searching for faces in the crowd and paying attention to the traffic around you. We have recently begun to examine the ability of participants engaged in visual search to resume a search after a forced interruption (Lleras, Rensink & Enns, under review). In contrast to some theories that have regarded visual search as being largely amnesic — on the grounds that observers have little memory for previously inspected locations and that a secondary visual memory task does not interfere with search — we have found that young adults are unusually good at resuming a search after it has been interrupted. This is seen in their exceptionally rapid response times to displays that they have viewed only once very briefly.

Having begun a search that is interrupted, searchers are often able to resume their task at a rate comparable to the speed with which they can respond to a target in the absence of any distracting items. In other words, they no longer need to search among the candidate items; they merely need to test the hypothesis they formed in their initial glance at the display on the second presentation of the same display. We have confirmed that the hypothesis being tested is linked to a particular location in the display and that perturbations of display items that are not near this location do not impair the rapid resumption of the search.

These results are consistent with our general view that visual perception consists of a recursive sequence of hypothesis generation and hypothesis testing. In a single glance at a scene, hypotheses about it must first be generated before they can be tested (confirmed or rejected). Hypotheses based on an initial glance can be tested very rapidly in a second glance, simply because the initial generation step has already been accomplished. On this account only a limited portion of a scene — namely that involving the hypothesis — needs to be remembered during the interruption. This is a very different conceptualization of the 'memorial processes' involved in visual search than have typically been assumed. Rather than memory consisting of some 'finished product' of perception (something like a photograph) at the end of a long chain of neural events, memory in visual search seems to consist instead of a rough map of where detailed processing has already been completed and what one expects to find at the next location to be examined (Hochberg, 1968).

References

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

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

Breitmeyer, B. G. (1984). Visual masking: An integrative approach. New York: Oxford University Press.

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.

Broadbent, D. (1958). Perception and communication. Oxford: Pergamon.

Bullier, J., McCourt, M. E., & Henry, G. H. (1988). Physiological studies on the feedback connection to the striate cortex from cortical areas 18 and 19 of the cat. Experimental Brain Research, 70, 90-98.Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in primate visual cortex. Cerebral Cortex, 1, 1-47.

Carpenter, G., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. Computer Vision, Graphics, and Image Processing, 37, 54-115.

Damasio, A. R. (1994). Descartes' error. New York: Putnam's.

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., Kawahara , J-I., Ghorashi, S.M., & Enns, J. T. (in press). The Attentional Blink: Resource Depletion or Temporary Loss of Control? Psychological Research.

Di Lollo, V., von Mühlenen, A., Enns, J.T. & Bridgeman, B. (in press). Decoupling Stimulus Duration from Brightness in Metacontrast Masking: Data and Models. Journal of Experimental Psychology: Human Perception & Performance.

Eggermont, J. J. (1990). The correlative brain. Berlin: Springer.

Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: Behavioral and electrophysiological evidence. Journal of Experimental Psychology: Human Perception And Performance, 24, 1737-1747.

Eimer, M., & Schlaghecken, F. (2001). Response facilitation and inhibition in manual, vocal, and oculomotor performance: Evidence for a modality-unspecific mechanism. Journal of Motor Behavior, 33, 16-26.

Eimer, M., & Shlaghecken, F. (2002). Link between conscious awareness and response inhibition: Evidence from masked priming. Psychonomic Bulletin & Review, 9, 514-520.

Eimer, M., Schubo, A., & Schlaghecken, F. (2002). Locus of inhibition in the masked priming of response alternatives. Journal of Motor Behavior, 34, 3-10.

Enns, J. T. (2003). Object substitution and its relation to other forms of visual masking. Vision Research, 44, 1321-1331.

Giesbrecht, B. L., Bischof, W. F., & Kingstone, A. (2003). Visual masking during the attentional blink: Tests of the object substitution hypothesis. Journal of Experimental Psychology: Human Perception and Performance, 29, 238-255.

Growney, R., Weisstein, N., & Cox, S. I. (1977). Metacontrast as a function of spatial separation with narrow line targets and masks. Vision Research, 17, 1205-1210.

Hebb, D. O. (1949). The organization of behavior. New York: Wiley.

Helmholtz, H. V. (1866). Handbuch der physiologischen Optik (1st ed.). Leipzig: Voss. (Trans. J. P. C. Southall (1962). Handbook of physiological optics (3rd ed.). New York: Dover).

Hochberg, J. (1968). In the mind's eye. In R. N. Haber (Ed.), Contemporary theory and research in visual perception (pp. 309-331). NY: Holt Rhinehart and Winston.

Hochstein, S., & Ahissar M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. Neuron, 36, 791-804.

Hubel, D. H. (1988). Eye, brain, and vision. New York: Scientific American Library.

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

Jiang, Y., & Chun, M. M. (2001a). The spatial gradient of visual masking by object substitution. Vision Research, 41, 3121-3131.

Klapp, S.T., & Hinkley, L.B. (2002). The negative compatibility effect: Unconscious inhibition influences reaction time and response selection. Journal of Experimental Psychology: General, 131, 255-269

Kolers, P. A. (1972). Aspects of Motion Perception. New York: Pergamon Press.

Lamme V.A.F., Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. Trends in Neuroscience, 23, 571-579.

Lamme, V. A. F. (2000). Neural mechanisms of visual awareness: A linking proposition. Brain and Mind, 1, 385-406.

Lleras, A., & Enns, J. T. (2004). Negative compatibility or object updating? A cautionary tale of mask-dependent priming. Journal of Experimental Psychology: General.

Lleras, A., & Moore, C. M. (2003). When the target becomes the mask: Using apparent motion to isolate the object-level component of object substitution masking. Journal of Experimental Psychology: Human Perception and Performance, 29, 106-120.

Lleras, A., Rensink, R.A., & Enns, J. T. (under review). Rapid resumption of interrupted search reveals new role for memory in human vision.

Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. Nature Neuroscience, 1, 144-149.

Matin, E. (1975). The two-transient (masking) paradigm. Psychological Review, 82, 451-461

Mignard, M., & Malpeli, J. G. (1991). Paths of information flow through visual cortex. Science, 251, 1249-1251.

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

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 and Performance, 28, 682-694.

Neisser, U. (1967). Cognitive Psychology. New York: Appleton-Century-Crofts.

Neumann, O. (1990). Direct parameter specification and the concept of perception. Psychological Research, 52, 207–215.

Pylyshyn, ZW (2003). Seeing and visualizing: It's not what you think. Cambridge, MA: MIT Press/Bradford Books.

Scharlau, I., & Ansorge, U. (2003). Direct parameter specification of an attention shift: Evidence from perceptual latency priming. Vision Research, 43, 1351–1363.

Schlaghecken, F., & Eimer, M. (2001). Partial response activation to masked primes is not dependent on response readiness. Perceptual Motor Skills, 92, 208-222.

Schlaghecken, F., & Eimer, M. (2002). Motor activation with and without inhibition: Evidence for a threshold mechanism in motor control. Perception & Psychophysics, 64, 148-162 .

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

Weisstein, N., Ozog, G., & Szoc, R. (1975). A comparison and elaboration of two models of metacontrast. Psychological Review, 82, 325-343.

Woodman, G. F., & Luck, S. J. (2003). Dissociations among attention, perception, and awareness during object-substitution masking. Psychological Science, 14, 605-611.

Zeki, S. (1993). A Vision of the Brain. Oxford: Blackwell.

Figure captions

Figure 1. (A): Schematic diagram of the display sequence in common onset masking. The sequence began with a display containing between 1 and 16 rings, one of which (the target) was singled out by four surrounding dots. The sequence continued without interruption with a display of the same four dots for durations between zero and 320 ms. (B): Target identification accuracy in a four-dot masking study (Di Lollo et al., 2000). No masking occurs when attention can be rapidly deployed to the location of the target, as occurs when set size is equal to one. Accuracy is also affected little by increments in set size, provided that the four-dot mask terminates with the target display. Pronounced masking occurs, however, as both set size and mask duration is increased. According to the reentrant hypothesis, this occurs because the representation of the unattended target has been replaced by the mask representation before target identification could be completed.

Figure 2. A computational model for object substitution (Di Lollo et al., 2000). A large number of three-layered modules, such as the one shown here, are arrayed over the visual field. Stimuli from the visual field arrive at the Input Layer, where the receptive fields are small, the feature coded are simple, and activation decays rapidly unless maintained by continued external input. The contents of the Input Layer are summed with the current contents of the Working Space, and sent to the Pattern Layer, where the receptive fields are much larger and code for more complex patterns.

Figure 3. (A) Example target display in the comparison of masks (Enns, 2003). The cross at the center indicates the fixation point; the black dot indicates the target letter to be reported, in this case the letter B. (B) The target letter was preceded by, presented concurrently with, or followed by, one of these masks.

Figure 4. Target identification accuracy when the mask is the probe (Enns, 2003). (A) Dot Probe Baseline, (B) Metacontrast, (C) Noise, (D) Four Dots, (E) Digits, and (F) Letters.

Figure 5. Target identification accuracy with a spatial precue (Enns, 2003). (B) Metacontrast, (C) Noise, (D) Four Dots, (E) Digits, and (F) Letters.

Figure 6. Mean priming effects (Incompatible RT minus Compatible RT) as a function of prime type (Lleras & Enns (2004). Primes and targets were double-headed arrows. Primes were presented for 15 ms, 1.5 degrees above or below fixation. Masks were presented for 120 ms at both possible prime locations. Targets were presented at fixation 100 ms after the mask's offset. The prime-mask interval was 45 ms in conditions A and B, 15 ms in condition C and 90 ms in condition D. A and B represent priming data for which prime visibility was higher on irrelevant-mask trials than on relevant-mask trials (for irrelevant and relevant masks respectively, A: 87% and 72%, B: 88% and 74%). C and D show data for which prime visibility was equal on irrelevant-mask trials and on relevant-mask trials (for irrelevant and relevant masks respectively, C: 58% and 66%, D: 94% and 89%). Examples of the masks used in each condition are shown next to the corresponding bar.

Figure 1

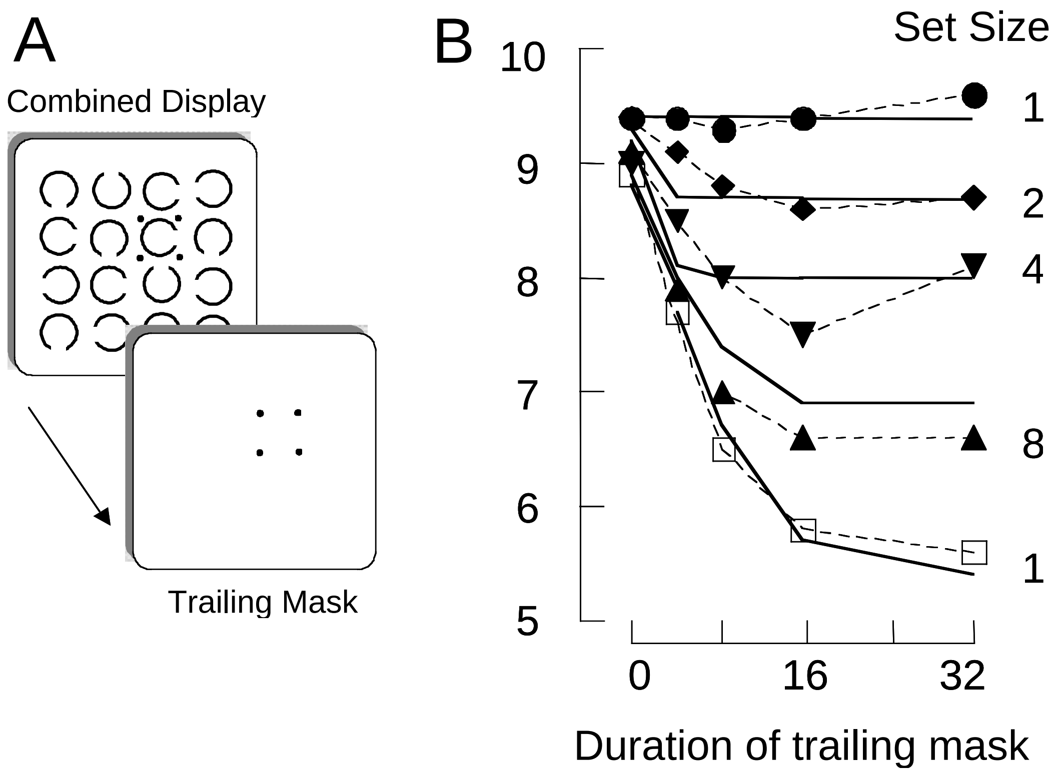

A
Combined Display
Trailing Mask
B
Set Size
10
9
8
7
6
5
1
2
4
8
1
0
16
32
Duration of trailing mask

Figure 2

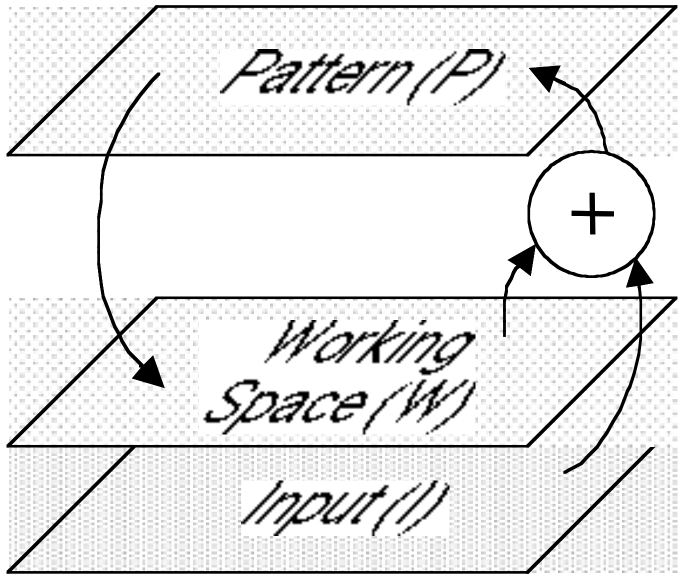


Figure 3A

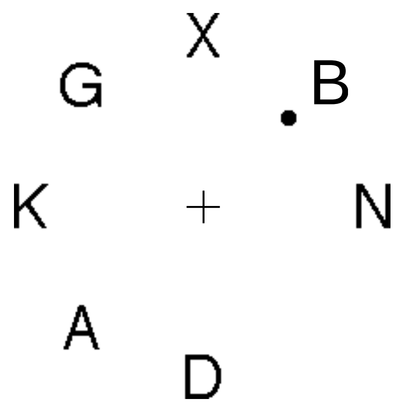


Figure 3B

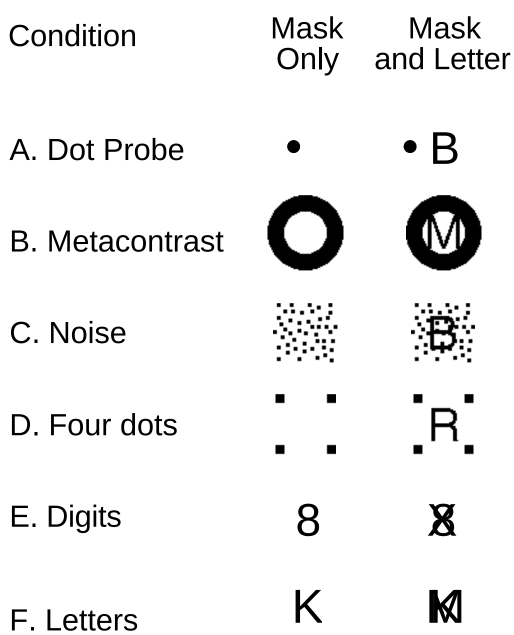


Figure 4

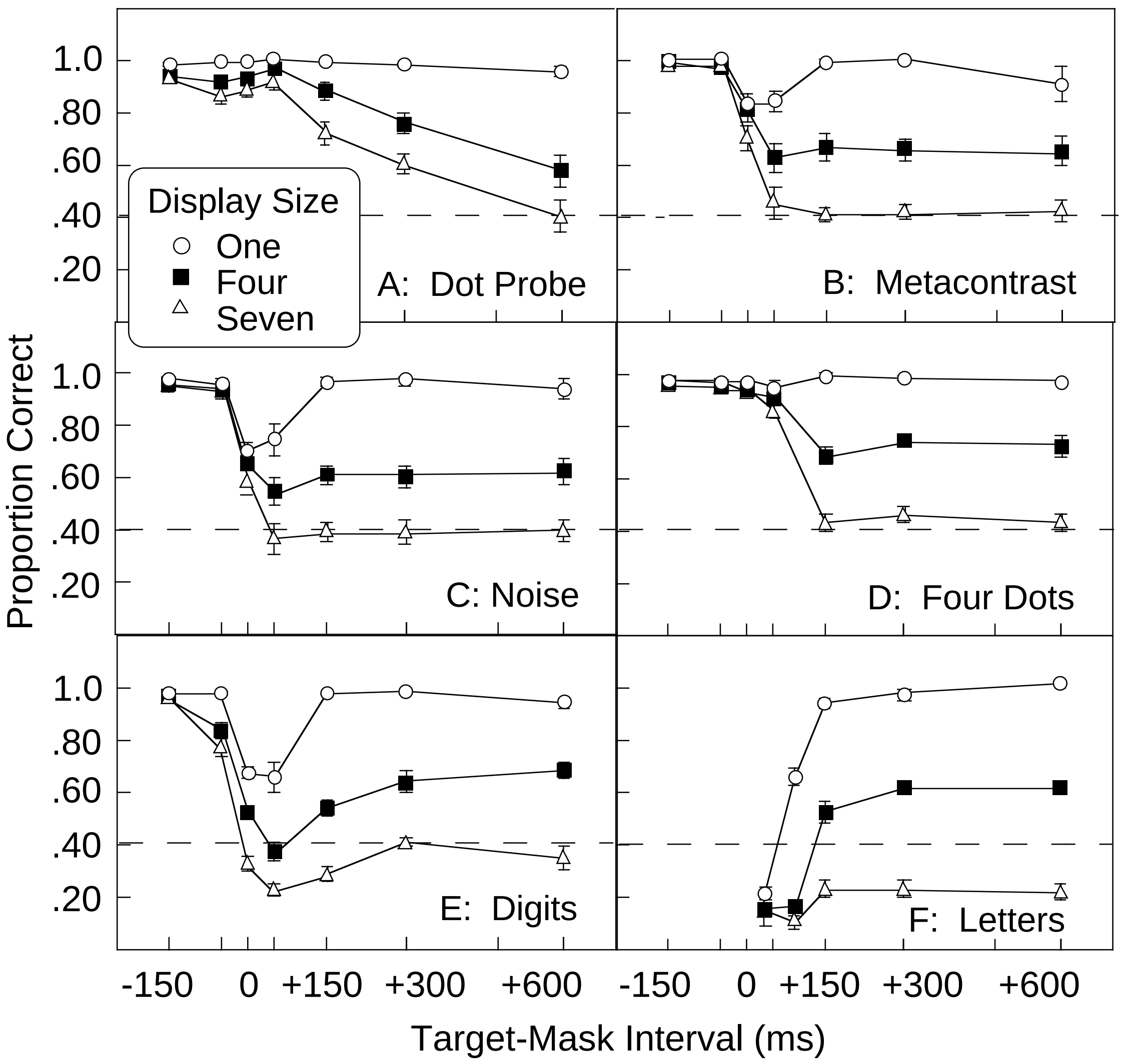
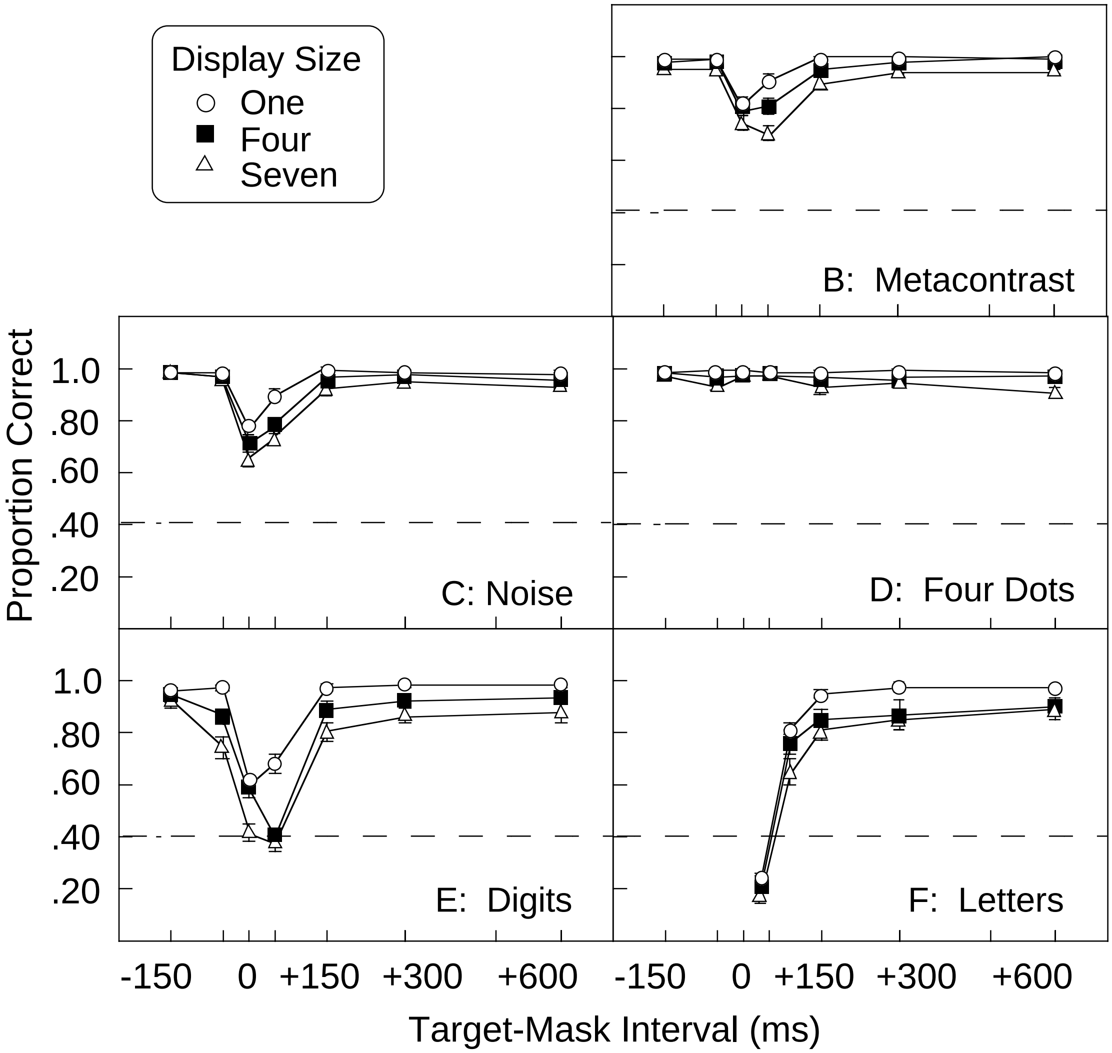
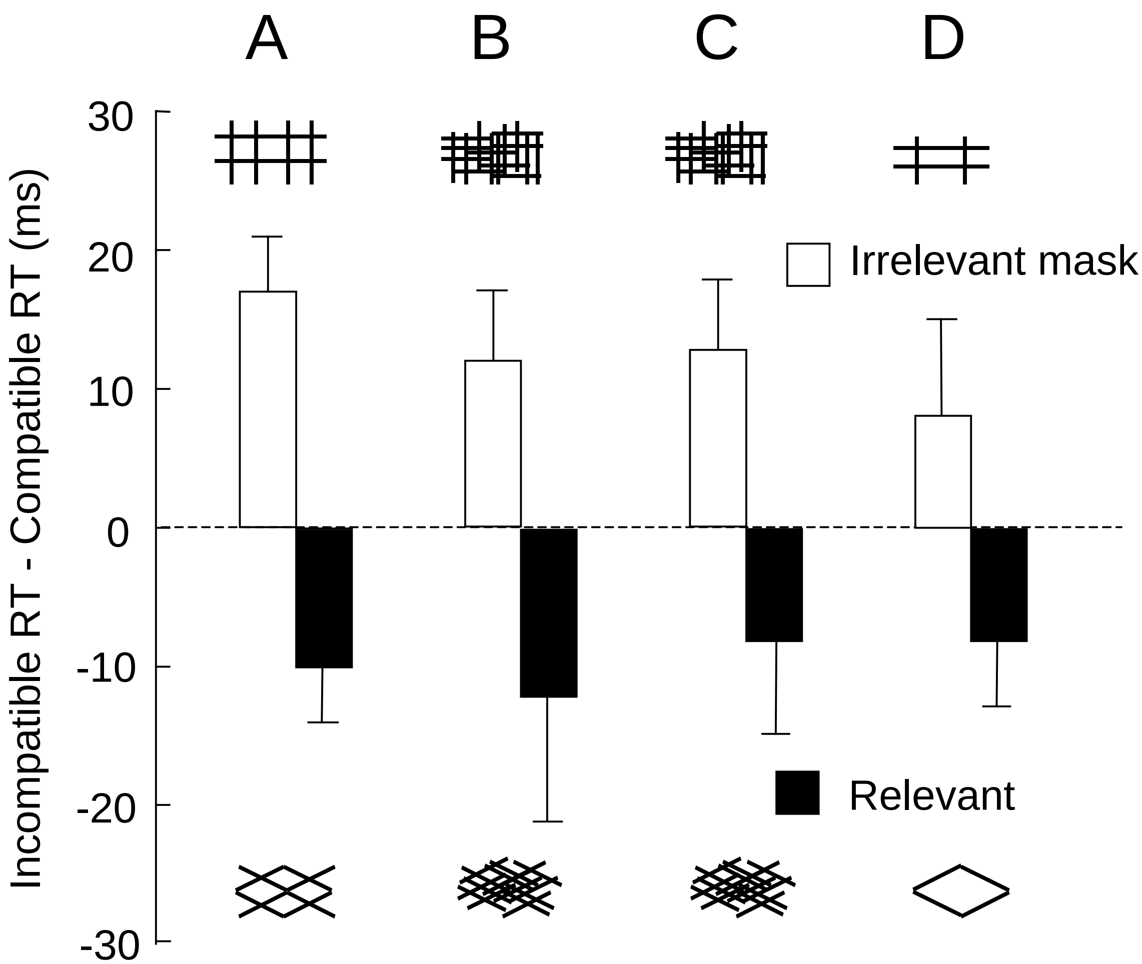

Display Size
One
Four
Seven
A: Dot Probe
B: Metacontrast
C: Noise
D: Four Dots
E: Digits
F: Letters
Proportion Correct
1.0
.80
.60
.40
.20
-150
0
+150
+300
+600
Target-Mask Interval (ms)

Figure 5



Display Size
One
Four
Seven
B: Metacontrast
C: Noise
D: Four Dots
E: Digits
F: Letters
Proportion Correct
1.0
.80
.60
.40
.20
-150
0
+150
+300
+600
Target-Mask Interval (ms)

Figure 6.


A
B
C
D
30
20
10
0
-10
-20
-30
Incompatible RT - Compatible RT (ms)
Irrelevant mask
Relevant