Research Article

METACONTRAST MASKING AND ATTENTION

John Shelley-Tremblay and Arien Mack
New School University

Abstract—Metacontrast masking is generally considered an effect of preattentive processes operative in early vision. Because of the growing evidence of the role of attention in other phenomena previously considered low level and preattentive, its possible role in masking was explored in three experiments in which meaningful target stimuli known to capture attention (one’s own name or a happy-face icon) were compared with control stimuli identical in spatial frequency and luminance. Not only were these salient stimuli significantly more resistant to masking than the control stimuli, but when one of the meaningful stimuli served as a mask, its strength was increased relative to a less meaningful variant, documenting the strong influence of attention on masking.

A group of phenomena, some of them quite recently described, testify to the centrality of attention for conscious perception. Among these phenomena are inattentive blindness (IB; Mack & Rock, 1998), the attentional blink (Raymond, Shapiro, & Arnell, 1992; Shapiro, 1994), repetition blindness (Kanwisher, 1987; Kanwisher & Potter, 1990), failures to perceive scene changes (Rensink, O’Regan, & Clark, 1997; Simons, 1996; Simons & Levin, 1997), and visual neglect (Rafal, 1998; Rafal & Robertson, 1995; Bisiach, Luzzatti, & Perani, 1979). Each of these is an instance of a failure to perceive suprathreshold stimuli that is attributed to inattention, and they are all consistent with the proposition that there is no conscious perception without attention. If this is the case, as one of us has previously argued (Mack & Rock, 1998), then it seems reasonable to ask whether attention plays a critical role in other instances of perceptual failure not normally attributed to inattention, such as masking or binocular rivalry. These phenomena are of particular interest because they are typically attributed to lower-level, sensory processes assumed to be independent of attention. Consequently, we were research to reveal that attention contributes to these phenomena as well, this would not only further underscore the centrality of attention for perception but also necessitate some revision in their explanations.

The experiments reported here explored the possible role of attention in metacontrast masking (MM), which occurs when the visibility of a briefly presented, suprathreshold stimulus is reduced or eliminated by the prior (forward masking; i.e., paracontrast) or subsequent (backward masking) presence of a surrounding or flanking stimulus. This kind of masking is typically attributed to low-level, preattentive interactions (see review in Breitmeyer, 1984), and therefore seemed well suited to our purpose. However, although the majority of studies of MM reflect this view, there are a few that point to a contribution of higher-level processes as well (e.g., Di Lollo, Enns, & Rensink, 1999; Proctor, Nunn, & Pallos, 1983; Ramachandran & Cobb, 1995). In fact, one recent, ingenious study demonstrated the clear influence of attentional processes on MM and provided the inspiration for the present research.

Ramachandran and Cobb (1995) described a case in which masking either does or does not occur, contingent upon the focus of the subject’s attention. In the display used in this study, the target stimulus was a disc flanked by two identical discs, creating a row of discs, and was followed by four flanking columns, two above and two below the central target. Ramachandran and Cobb found more masking on trials in which the observer attended to the vertical columns constituting the mask than on trials in which the observer attended to the row of discs that included the target. The authors concluded from this finding that “far from being an autonomous front-end process, backward masking can be strongly modulated by top-down influences such as visual attention” (p. 68).

Our efforts to explore the influence of attention on MM were based in part on earlier work on IB, the failure to perceive a suprathreshold stimulus located at or near fixation under conditions of inattention (Mack & Rock, 1998). In the course of studying IB, we found that a few highly meaningful stimuli successfully resisted IB and generally were seen under conditions of inattention even though stimuli defined by simple features like motion, color, and flicker, typically considered to be preattentively processed, were not. Because IB research demonstrated that the few stimuli which resist IB do so because they succeed in capturing attention, we speculated that if attention has a role in MM, these same stimuli might also resist MM.

From the small set of IB-resistant stimuli, we chose the subject’s own name and a happy-face icon (��) as the target stimuli for the current experiments. We used minor modifications of these stimuli as control targets because we found that such modifications (e.g., changing a single vowel in the subject’s name or inverting the face) destroyed their resistance to IB. In the pilot studies and the first experiment, the target stimulus was either a happy-face icon or a previously investigated modified variant (see Fig. 1). In Experiment 2, the target was either the subject’s own name, a scrambled version of the name, or a highly frequent English word previously found to suffer IB.
tial frequency, luminance, and extent. Thus, if the subject’s own name were the mask, its power should be

significantly reduces the effects of a mask, then increasing the meaning-

fulness of the mask should increase its power in a comparable way. We predicted that if attention plays a significant role in MM, then stimuli that are resistant to IB because of their power to capture attention (e.g., the happy-face icon and one’s own name) should be more resistant to masking than modified variants of these stimuli that maintain the spatial frequencies and luminance profiles of the originals. If this is the case, then it might be because the power to capture attention produces an increase in the signal strength of the to-be-masked stimulus, thereby increasing its resistance to MM.

The last experiment extended the investigation into the role of attention by examining whether the meaningfulness of a mask affects its power to mask. If the presence of a highly meaningful target significantly reduces the effects of a mask, then increasing the meaningfulness of the mask should increase its power in a comparable way. Thus, if the subject’s own name were the mask, its power should be increased relative to a modified version of the name with the same spatial frequency, luminance, and extent.

PILOT STUDIES

Two pilot studies were performed on a standard Macintosh Quadra/Powerbook computer with a CRT screen with a slow refresh rate (66 Hz). In the first pilot study, the target stimulus was a happy-face icon or one of its two scrambled variants (see Fig. 1). Half the subjects were shown one variant, and the remaining subjects were shown the other. In the second pilot study, the target stimulus was the happy face or an inverted happy face (see Fig. 1). The figures were outlined in white and presented on a black screen. A centered fixation mark appeared on a black screen at the beginning of each trial. The target stimulus was then presented at fixation for 16.6 ms, followed by an interstimulus interval (ISI) of 16.6 ms, an annulus mask for 33.2 ms, and an intercycle interval of 282.2 ms. This sequence was repeated four times per trial, and each subject viewed four trials per stimulus, resulting in a total of 16 possible target detections for each of the two target stimuli. Following each trial, subjects reported whether they had seen anything other than the mask; if they did, they rated its clarity on a scale from 1 to 5, with 5 indicating maximum clarity. Prior to testing, each subject was asked to rate the clarity and brightness of the annulus when presented alone. All subjects rated it as 5 (i.e., as clear and bright) at 1.55 lumen/m². This then served as the standard against which the brightness of the target was judged.

Twelve subjects were tested in the first pilot study. Because there was no difference in the detectability of the two scrambled variants, they were treated as equivalent. Subjects detected the presence of 46% of the happy faces but only 14% of the scrambled variants, *t*(11) = 7.33, *p* < .001. Clarity ratings also were consistent with the prediction that a meaningful stimulus is more resistant to masking: The mean rating was 3.4 for the happy face but 2.07 for the scrambled variants.

In the second pilot study, involving 14 subjects, an equal number of trials in which no target preceded the mask was added so that we could assess the false alarm rate (i.e., the probability that subjects were only guessing that a target had been present). Again, there were four target-mask or mask-only cycles per trial and four trials per type of presentation (happy face, inverted face, or no target), resulting in 16 exposures of each presentation type. Happy-face, inverted-face, and no-target trials were randomly intermixed.

Subjects detected the presence of the happy face significantly more often (56%) than the inverted face (24%), *t*(13) = 9.3, *p* < .001, confirming the results of the first pilot study. No false alarms were report-
ed, indicating that the results could not be attributed to guessing.

The results of both pilot studies provided the grounds for pursuing the investigation using more precisely controlled displays. The major difference between the pilot studies and the main experiments is that in the main experiments a tachistoscope was used to present stimuli, permitting greater control over luminance and timing.

EXPERIMENT 1

Method

Subjects

Thirteen subjects were paid $10 for their participation. All subjects reported normal or corrected-to-normal vision. Two subjects were excluded for making a large number of false-positive responses, as well as difficulty in reporting the unmasked stimulus.

Apparatus

A four-channel Gerbrands G1135 T4A tachistoscope with an independent lamp control and timing apparatus was used. Timing was accurate to 1 ms between 1 and 100 ms, and to 10 ms between 100 and 1,000 ms. Four lamps projected images independently, producing a composite image seen through a binocular viewer.

Stimuli

The three possible target figures were an upright happy face, an inverted happy face, and a scrambled variant. They were drawn in white and centered on matte black 4-in. x 6-in. cards. The figures were 0.67 cm in diameter and subtended 0.5° viewed from 78 cm. The mask was a white annulus with an inner diameter of 0.67 cm and an outer diameter of 1.0 cm. Its inner contour appeared to slightly overlap the outer contour of the target stimulus, as shown in Figure 2.

---

**Fig. 1.** Stimuli used in the pilot studies and Experiment 1. Stimuli A, C, and D were used in the first pilot study. Stimuli A, B, and E were used in the second pilot study. Stimuli A, B, D, and E were used in Experiment 1.

We predicted that if attention plays a significant role in MM, then stimuli that are resistant to IB because of their power to capture attention (e.g., the happy-face icon and one’s own name) should be more resistant to masking than modified variants of these stimuli that maintain the spatial frequencies and luminance profiles of the originals.
Fig. 2. Spatial relationship between the target and mask in Experiment 1.

Stimulus luminance was constant and set at approximately 1.55 lumen/m², which is considered within the dim photopic range. This was the lowest luminance at which subjects consistently gave the annulus presented alone the top rating of 5, “clear and bright,” on the clarity scale in the pilot studies.

Procedure

Subjects placed their faces firmly against the viewer, and testing was carried out in a dimly lit room. The stimulus sequence is shown in Figure 3. There was an approximately 5-s interval between trials while the stimuli were reset.

Each block of trials contained a randomly ordered mixture of each of the three stimulus types plus an equal number of trials in which no stimulus preceded the mask. Each display type was presented at five different stimulus onset asynchronies (SOAs: 0, 20, 60, 100, and 136 ms), for a total of 20 trials per block. Ten blocks were presented, so that each subject had a total of 200 trials. The testing procedure was identical to the procedure in the pilot experiments, except that each trial consisted of only one stimulus-mask presentation, rather than four. Subjects were asked to report the presence or absence of anything other than the annulus following each trial. If they reported seeing anything other than the mask, they were asked to rate its brightness on the 5-point scale used in the pilot studies.5

Results

The results, summarized in Figure 4, confirmed the general trend of the pilot experiments. It is clear that the procedure was effective in producing masking and in generating the familiar U-shaped metac-contrast function. Maximum masking occurred with a 20-ms SOA, which is expected when target and mask are of equal luminance (Proctor et al., 1986). A repeated measures analysis of variance (ANOVA) of the brightness ratings revealed a significant main effect of stimulus type, F(2, 20) = 17.34, p < .001, and of SOA, F(4, 40) = 35.64, p < .001. Post hoc comparisons showed that the happy face produced significantly higher ratings of subjective brightness than either of the other two targets, t(11) = 4.36, p < .001, which did not differ from each other.

Detection scores were also analyzed and yielded a significant effect of stimulus type, F(2, 20) = 7.24, p < .01, and of SOA, F(4, 40) = 13.42, p < .001. The analysis also yielded a significant interaction of stimulus type and SOA, F(8, 80) = 4.28, p < .01. A post hoc analysis indicated that across SOAs, the rate of detection was significantly greater for the happy face (89%) than for both the inverted face (69%) and the scrambled face (63%), t(9) = 3.17, p < .01; the detection rates for the scrambled and inverted faces did not differ. Because the false-positive response rate was extremely low (6%), the results cannot be attributed to guessing.

Discussion

Experiment 1 provides the best evidence thus far that under strictly controlled conditions the happy face is significantly more resistant to MM than its scrambled or inverted variants. Because our earlier work (Mack & Rock, 1998) clearly documented the capacity of the happy-face icon to capture attention, the most obvious explanation of these results is that the captured attention modulates masking. If the mask masks because it takes attention away from the prior stimulus, then increasing the attention-grabbing characteristic of the prior stimulus inevitably must reduce masking, and it seems to.

EXPERIMENT 2

The second experiment was designed to extend the examination of the influence of attention on MM by assessing the degree to which another stimulus known to capture attention and resist IB reacts to MM. The stimulus we chose was the subject’s own name. Here again, our basic hypothesis was that the name would show more resistance to metac-contrast than a scrambled variant despite the fact that the number of letters, luminance, and spatial frequency of the two stimuli were identical. We also added the high-frequency word “TIME” as a control target, because IB research had revealed that despite its frequency of occurrence and therefore its familiarity, it tended not to be seen under conditions of inattention. Moreover, we wanted to determine whether any difference in the maskability of the name and its scrambled variant was due to a word-nonword difference (Reicher, 1969), although,

---

5. Subjects were instructed not to rush their responses because subjects who are forced to respond as rapidly as possible in MM experiments are less affected by changes in SOA. One account of this finding presumes that MM occurs only if there is sufficient time for the development of the iconic representation of both target and mask (Lachter & Durgin, in press).
of course, if it were, this still would implicate meaning in MM. Because we deemed it important to keep the degree of visual angle subtended by the stimuli roughly equivalent, we limited testing to subjects with four-letter names. Our subjects’ names varied by ±6 pixels in extent from the word “TIME.”

**Method**

**Subjects**
Ten subjects were tested and paid for their participation.

**Stimuli**
The target stimuli were all four letters long and consisted of the subject’s name, a scrambled variant (letters arranged randomly), and the word “TIME.” All words were written in capital letters in 12-point, New York font in bold script. The letter stimuli were white, against a black background, and measured 0.67 cm (0.5°) at a viewing distance of 76 cm. The mask consisted of two white horizontal bars, one above the other, forming a flanking mask. The bars were both 1.5 cm wide and 0.2 cm tall, with a 0.22-cm black gap between them. The visual angle of the mask was 1.13°. The background luminance was approximately 1.52 lumen/m². The luminance of the fixation dot was approximately 1.28 lumen/m², and the luminance of the target and mask were identical, also about 1.28 lumen/m².

Prior to the main experiment, we established the luminance level that yielded a 50% detection threshold for the word “TIME” with a target-mask SOA of 20 ms. This luminance, approximately 1.52 lumen/m², became the target luminance in the experiment.

**Procedure**
The procedure was identical to that in Experiment 1, except that subjects were shown five mask-only trials at the beginning of the experiment and again halfway through the testing session, in order to give them a clearer idea of what a trial with no target preceding the mask looked like. Another difference in procedure was that in this experiment, and in Experiment 3, we asked subjects only whether or not they detected the stimulus. We did not ask them to rate its brightness as well, because the results of Experiment 1 indicated that these latter reports were redundant and the detection reports allowed for a simple yes/no response. Subjects were given a 5-min break at the end of the third block of trials, as well as short breaks between other blocks.

---

**Fig. 4.** Mean number of target detections and clarity ratings for the three target stimuli and no-target trials in Experiment 1. Clarity was rated on a scale from 1 to 5, with 5 representing maximum clarity. SOA = stimulus onset asynchrony.
Results

Maximum masking again occurred at 20 ms. The average frequencies of detection of the various targets are shown in Figure 5. A repeated measures ANOVA indicated a significant main effect of SOA, verifying the effectiveness of the mask, and a significant interaction between SOA and stimulus type, $F(2, 9) = 11.88, p < .01$. Most important, a post hoc analysis revealed that at the 20-ms SOA, subjects detected their names significantly more often than both the scrambled variants and the word “TIME,” $t(9) = 3.14, p < .01$; detection frequencies for the latter two targets did not differ from each other. Once again, we found that a stimulus that captures attention and resists IB resists MM relative to control stimuli that share its size, spatial frequency, and luminance. It should be noted that of the subjects who detected the masked target, almost none were able to correctly identify it when informally questioned. This was also true in Experiment 1.

EXPERIMENT 3

In this experiment, we explored the effect of meaningfulness on the power of the mask, so we changed the role of the subject’s name from target to mask. We reasoned that if attention modulates masking and is responsible for the name’s resistance to masking, then the name also might be a more effective mask than a scrambled variant; a mask capable of more readily diverting attention away from the target should be more effective than a variant with the same luminance and spatial-frequency profile but without the same capacity to capture attention.

Method

The methods of Experiment 2 were repeated, with a few exceptions. In the second experiment, the mask always consisted of two flanking bars, and the target varied from trial to trial. In Experiment 3, the target was constant but the mask varied from trial to trial. There were three types of masks: the subject’s own name, a scrambled variant of the name, and the word “TIME.” Following each trial, subjects reported whether they had seen anything other than the mask, and subsequently tried to identify as many of the letters in the mask as they could.

Subjects

Ten subjects were each paid $15 for their participation.

Stimuli

On each trial, the mask consisted of two copies of the subject’s own name, its scrambled variant (two scrambled versions were used: first and fourth letters exchanged and second and third letters exchanged), or the word “TIME.” The two copies were stacked one above the other in the center of the screen. The masks were presented in uppercase, 12-point New York font. An example is presented in Figure 6. As in Experiment 2, all subjects had four-letter names. The target was the word “CORK.”

Procedure

As in Experiment 2, there were five mask-only trials prior to and again halfway through testing. During the main testing, each subject viewed 10 blocks of trials. Each block consisted of 20 trials that were composed of randomly ordered presentations of each of the three masks plus mask-only trials at each of the five SOAs used earlier.
Fig. 6. Example of a mask used in Experiment 3.

Fig. 7. Mean number of detections for target-present trials with three different masks and for no-target trials in Experiment 3. SOA = stimulus onset asynchrony.
Metacontrast Masking and Attention

There were thus 50 presentations of each of the mask types and an equal number of mask-only trials over the 200 trials per subject.

Results

Figure 7 shows the mean number of detections of the target for each of the three different masks at each SOA. It is clear that we again succeeded in obtaining a characteristic U-shaped masking function. The results were subjected to an ANOVA to determine the effect of the different masks on target detection; the ANOVA proved significant, \( F(2, 9) = 4.84, p < .05 \). Post hoc comparisons of target-detection differences among the three different masks at the 20-ms SOA revealed, as we had speculated, that the name was a significantly stronger mask than either the scrambled name \((p < .05)\) or “TIME” \((p < .01)\), and there was no difference between the latter two masks. Thus, these data also support the proposition that attention drawn either to the mask or to the target modulates MM.

GENERAL DISCUSSION

The pilot results and the results of Experiments 1 and 2 support the view that stimuli that capture attention under conditions of inattention and that are therefore less vulnerable to IB than stimuli that do not capture attention are also resistant to MM; the results of Experiment 3 indicate that when an attention-capturing, IB-resistant stimulus serves as a mask, it has greater power than a mask that is not IB resistant. This is strong evidence for the role of meaning and attention in MM and, of course, more generally in perception, given the prior research cited in the introduction.\(^6\) The importance of meaningfulness is made especially clear by the finding that just as rearranging local features of a highly salient stimulus without altering luminance, spatial frequency, or visual extent significantly reduces the resistance of the stimulus to IB, it also significantly alters the resistance of the stimulus to masking as well as its power to mask. These results cannot be explained by any low-level processes.

It seems that masking occurs, at least in part, because the appearance of a mask co-opts attention, diverting it away from the target stimulus. A target with high attentional valence (such as the happy-face icon or name) is able to hold on to attentional processes, thus lowering the effectiveness of the mask. Similarly, if the mask has high attentional valence, its power to mask is increased because it is better able to divert attention away from the target. This high-level, attentional effect is consistent with the results reported by Ramachandran and Cobb (1995) but at odds with the more widely held explanations of MM, which rest heavily upon low-level sensory processes and therefore at the very least require a revision to take account of top-down interactions.

The link between MM and IB is obvious: Both are a function of attention, and both are consistent with the claim that attention is necessary for conscious perception.\(^7\) In one case, the absence of attention results in blindness; in the other, it results in masking. Only a very few highly meaningful and familiar stimuli have been found capable of capturing attention when an assigned task directs it elsewhere, and these same stimuli have a powerful effect on masking. In the case of masking, augmenting attention to the target or mask by manipulating its meaningfulness increases or decreases, respectively, the probability of target detection. To extend the analogy between IB and masking further, it would be necessary to demonstrate that masked targets are deeply encoded despite their unavailability to conscious perception, which is the case for stimuli suffering IB, as they are capable of semantic priming (Mack & Rock, 1998).\(^8\) If in fact masked targets are deeply encoded although unavailable to consciousness, and there is some reason to think they are (see, e.g., Greenwald, 1992), then the masking story must be seriously recast, as this would mean that both masked and undetected targets are deeply processed and not obliterated by low-level, inhibitory interactions (Breitmeyer, 1980; Breitmeyer & Ganz, 1976). On this account, both masked stimuli and stimuli suffering IB are processed and encoded but not consciously perceived because they are not objects of attention.

One recently presented account of MM (Di Lollo & Dixon, 1988; Di Lollo et al., 1999) offers a possible explanation for how attention might modulate MM through a process of iterative reentrant processing. On this account, any variable that increases the number of iterations necessary to identify a target increases the strength of the mask. One such variable is selective attention. If, as seems reasonable, a highly salient stimulus like one’s name or a happy face requires fewer iterations for identification than a less salient stimulus, this might explain its resistance to masking. Unfortunately, this account, as currently presented, seems to depend on target identification and so cannot fully account for our results, because almost no subjects were able to identify the targets whose presence was detected. Moreover, it is not clear how this account could explain the effect of manipulating meaningfulness of the mask.

REFERENCES


---

6. The alternative explanation that meaningful stimuli may be easier to report and that therefore these results are unrelated to attention can be rejected on the grounds that in the current experiments subjects detected but rarely identified the meaningful target.

7. As pointed out previously, there is additional evidence of the centrality of attention for conscious perception in research on visual neglect and the attentional blink.

8. Stimuli that are unseen because of either visual neglect or the attentional blink are likewise deeply encoded but unavailable to consciousness.
detection and spatial judgments: An apparent automatic and attentive response
mechanisms effect. *Journal of Experimental Psychology: Human Perception and
Performance, 9*, 278–287.
Cambridge, MA: MIT Press.
(Ed.), *The cognitive neurosciences* (pp. 625–648). Cambridge, MA: MIT
Press.
processing in an RSVP task: An attentional blink effect. *Journal of Experimental
of Experimental Psychology, 81*, 275–280.
Rensink, R.A., O’Regan, J.K., & Clark, J.J. (1997). To see or not to see: The need for
attention to perceive changes in scenes. *Psychological Science, 8*, 368–373.
Shapiro, K., Caldwell, J., & Sorenson, R. (1997). Personal names and the attentional blink:
A visual “cocktail party” effect. *Journal of Experimental Psychology: Human Per-
ception and Performance, 23*, 504–514.
Shapiro, K.L. (1994). The attentional blink: The brain’s “eyeblink.” *Current Directions in
Psychological Science, 3*, 86–89.
Simons, D.J. (1996). In sight, out of mind: When object representations fail. *Psychologi-
cal Science, 7*, 301–305.
261–267.

(Received 10/14/98; Accepted 2/16/99)