Journal of Experimental Psychology:
Learning, Memory, and Cognition
1984, Vol. 10, No. 3, 435–441

Copyright 1984 by the
American Psychological Association, Inc.

Perceptual and Conceptual Masking of Pictures

Geoffrey R. Loftus and Michael Ginn
University of Washington

We report an experiment in which target pictures, presented for 50 ms, were followed by masks. Two mask variables were implemented: mask luminance and amount of attention demanded by the mask. Luminance but not attention demand affected subsequent picture-memory performance when the mask followed the picture immediately; however, attention demand but not luminance affected performance when the mask was delayed by 300 ms following the offset of the picture. We conclude that qualitatively different processes are being carried out at 0 versus 300 ms following the offset of a 50-ms picture. We argue that these processes can profitably be viewed as perceptual processes, which operate on raw stimulus input, and conceptual processes, which operate on the output of perceptual processes.

A visual mask is a stimulus that is presented close in time and space to some other, target, stimulus. In experiments designed to investigate visual perception and memory, masks are often used for one or both of two purposes. First, the experiment may be specifically concerned with the mechanisms that cause masking (e.g., Breitmeyer & Ganz, 1976; Turvey, 1973). Second, the experimenter may need to control the processing time of a visual stimulus as closely as possible; a mask would thus be used to eliminate the iconic image of a target stimulus, thereby allowing extraction of information from the stimulus only during its physical duration (e.g., Sperling, 1963).

Perceptual and Conceptual Masking

In recent picture-memory experiments, particularly those of Potter (1975, 1976; Potter & Levy, 1969) and Intraub (1980, 1981, 1984), briefly presented, complex, naturalistic, target pictures have been followed by masks of various sorts. The effects of these masks were assessed by subsequently presenting the target pictures in a memory test. The poorer the memory performance, the more effective the mask was assumed to be.

The research reported in this paper was supported by National Science Foundation Grant No. BN82-09617 to Geoffrey R. Loftus. We thank Rose Ahearn and Cheryl Kulas for collecting the data. Helene Intraub, Ellen Markman, and Molly Potter provided valuable comments on an earlier version of the manuscript.

Requests for reprints should be sent to Geoffrey R. Loftus, Department of Psychology, NI-25, University of Washington, Seattle, Washington, 98195.

Potter (1976) postulated that a mask following the offset of a picture can have two distinct effects, which she termed *perceptual masking* and *conceptual masking*. Perceptual masking, according to Potter's account, interrupts initial visual processing. Conceptual masking, in contrast, occurs after initial identification of the picture has occurred, and interrupts the higher level processing that is required for long-term storage of the information corresponding to the picture. Potter's general framework may be captured more specifically by the following two assumptions.

1. When a picture is initially presented to an observer, raw sensory information is extracted either from the picture itself or from an icon that follows it. The perceptual processes that extract this information require the presence of the stimulus or the icon as input. Based on this perceptually acquired information, the picture is identified.
2. Once the picture has been identified, further, conceptual, encoding is necessary if the picture is to be transferred to a more permanent memory. The processes by which conceptual encoding is carried out (a) operate on the output of the perceptual processes and (b) do not require the presence of the picture itself or its icon in order to operate.

Given these two assumptions, a perceptual mask may be defined as one that inhibits the operation of the perceptual processes, whereas a conceptual mask may be defined as one that inhibits the operation of conceptual processes.

At present, there do not exist rules by which one can determine a priori whether a physical stimulus will act as a perceptual or a conceptual mask. Indeed, most masks could and probably do act as both. However, one can postulate variables that may reasonably be expected to selectively affect one type of masking versus the other. One basis for such postulation is the assumption that target and mask compete for the resources required for any type of processing; hence, variation in a given type of mask processing leads to the opposite variation in the corresponding target processing. On this assumption, mask luminance would affect perceptual masking, whereas the amount of attention demanded by the mask would affect conceptual masking. As we discuss presently, there is empirical evidence for at least this latter contention.

Although no experiments have been designed to expressly distinguish between conceptual and perceptual masking of pictures, conceptual masking per se has been demonstrated rather convincingly. Consider, for example, a picture-recognition experiment reported by Intraub (1980, Experiment 1). In three conditions of this experiment, pictures were presented during an initial study phase for 110 ms. The interstimulus interval (ISI) between the pictures (a) consisted of 4,890 ms of darkness, (b) consisted of a 4,890-ms masking photograph, or (c) was zero, that is, the next target picture followed immediately. Performance declined monotonically over these three conditions. Note that Conditions b and c were perceptually identical in the sense that the duration of the target stimuli was the same, and in both conditions, target pictures were immediately followed by another photograph. Thus, the performance difference between the two conditions must be attributed to factors that affect conceptual rather than perceptual processes.

What about the difference between Conditions a and b? It might be argued that these two conditions were conceptually identical since conceptual processing of a masking picture was not required in either. On this argument, the performance difference would be attributable to perceptual masking. However, that conceptual processing of the mask was not required in Condition b does not necessarily mean that such processing was not carried out. If, as seems quite plausible, the appearance of a picture automatically demands some conceptual processing, then the difference between Conditions a and b could be due to conceptual masking.

A picture-memory experiment reported by Intraub (1981; see also Intraub, 1984) incorporated three conditions across which masks were perceptually identical but demanded progressively increasing amounts of attention. In the first condition, target photographs were masked by another photograph that was the same on every trial. In the second condition, a different masking photograph followed each target. The third condition was identical to the second except that the subject was required to remember the changing, masking photograph. Subsequent recognition performance for the target picture declined monotonically over the three conditions. Because the masks in the three conditions were perceptually identical, perceptual masking presumably did not differ, and the decreasing performance is attributable to increasing degrees of conceptual masking. Conceptual masking is, therefore, determined at least in part by the amount of attention demanded by the mask.

In the present experiment, we attempted to separate conceptual from perceptual masking in a picture-memory paradigm. In this experiment, each target picture was presented for 50 ms, and the subject's task was to recall as many details as possible from the picture. There were eight masking conditions, made up as follows: A mask was presented either immediately or 300 ms after the target picture's offset. We henceforth refer to these two delay conditions as the immediate and delayed masking conditions. Within each delay condition, there were four types of masks, defined by a 2 × 2 (Luminance × Attention Demand) factorial combination of luminance (bright or dim) and attention demanded by the mask (high or low). The latter variable was implemented in a manner similar to that of Intraub (1981) and Intraub (1984, Experiment 1): The mask was either random noise (low attention demand) or a picture—similar to the target pictures—that changed on every trial (high attention demand).

Within the framework that we have described, masking that occurs 50 ms after the onset of the picture must be primarily perceptual; in pilot work we found it to be rare that a picture could be identified within this short a time. Conversely, masking that occurs 350 ms after the onset—and 300 ms after the offset, by which time the picture's icon will have decayed away—should be completely conceptual, because there is no longer available any raw stimulus information on which perceptual processes could operate. Thus, we expect an effect on memory performance of luminance but not of attention demand when the mask is presented immediately. Conversely, we expect an effect of attention demand but not luminance when the mask is delayed by 300 ms.

Method

Subjects. Forty-five University of Washington undergraduates served as subjects. They were run in nine groups of 5 subjects per group.

Stimuli. The target stimuli were 54 naturalistic, color slides, chosen to be rich in detail. They included street scenes, indoor scenes, pictures of people, and landscapes. The pictures were quite dissimilar from one another. There were 54 masking photographs, 1 randomly assigned to each target, chosen from the same pool as the targets. In addition, there was 1 random noise mask, consisting of a jumble of black and grey slashes on a white background. The visual angle subtended by the pictures ranged from 18° horizontal × 12° vertical to 27° horizontal × 18° vertical, depending on where the subject sat.

Apparatus. The target pictures were displayed by a Kodak Carousel projector. The masking slides were displayed by a Kodak Random Access projector. A filter wheel in front of the latter projector allowed trial-to-trial variation in mask luminance. A dim fixation point was displayed by a third projector. Gerbrands tachistoscopic shutters with rise and fall times of approximately 1 ms were used to control timing. Response boxes containing the digits 0–9 were used to collect data. Picture presentation and response collection were under the control of an Apple II computer system. All projection apparatus was enclosed in a soundproof box.

When no picture (mask or target) was being presented, the luminance of the projection field was 0.34 cd/m^2.

Design. Each target picture was presented for 50 ms. The design of the experiment was 2 × 2 × 2, plus a no-mask control condition. As described earlier, the three factors constituting the 2 × 2 × 2 portion of the design were the following:

1. Mask temporal position. A mask was presented either immediately following the offset of the target picture or was delayed by 300 ms following the offset of the target picture.

2. Attention demand. The mask was either a random noise pattern (low attention demand) or was another photograph that changed from trial to trial (high attention demand).

3. Mask luminance. The mask was either unattenuated (bright) or attenuated by two log units (dim). The unattenuated luminances of the various portions of the noise mask were 29.12 cd/m^2 (white background), 3.43 cd/m^2 (grey slashes), and 3.08 cd/m^2 (black slashes). The geometric mean luminance of the 54 photo masks was equal to the overall luminance of the noise mask.

In the control condition, no mask was presented. Thus, there were nine conditions, all within subjects. Each target picture was rotated through the nine conditions across the nine groups of subjects.

Procedure. An experimental session consisted of two phases. In the initial test phase, the 54 pictures were presented. The nine experimental conditions were shown in a randomized order over the 54 trials with the restriction that within each block of 18 trials, each condition must occur exactly twice. Each test trial consisted of the following sequence of events.

1. A 1.5-s warning tone signaled the subjects to look at the fixation point that simultaneously appeared on the wall in front of them.

2. The fixation point disappeared, and the target picture appeared for 50 ms.

3. Except in the control condition, the mask appeared following its appropriate delay. The mask remained on for 500 ms.

4. There was a delay of 20 s, during which subjects did two things. First, in order to ensure that they had paid attention to the mask (some pilot subjects reported closing their eyes following picture presentation), they jotted down whether the mask was a photo, random noise, or nothing. Next, they carried out the primary response task, which was to write down as many details as they remembered from the target picture that they had just seen. They were instructed to "write down as many details as you can so that a person looking at your list would be able to reproduce the picture as accurately as possible." Following the 20-s delay were the warning tone and fixation point for the next trial.

In the scoring phase, which immediately followed the test phase, the 54 pictures were re-presented, 1 by 1, in the same order in which they had been presented originally. For each picture, the subjects wrote down the number of correct details that they had originally listed. They were told that "a detail" should correspond to a single object listed from the picture. For example, the response "two people" should count as two details. In practice, there were very few cases in which responses were ambiguous. In such cases, subjects were told to use their own judgment.

Subjects were cautioned not to add any details during the scoring phase but just to indicate how many they had previously written down. After each subject had computed the number of details written down for that trial, he or she entered that number into his or her response box. When all subjects had responded, the next test trial began.

The dependent variable that we used—number of reported details—is less common than recognition performance in picture-memory research (although, see Intraub, 1980, Experiment 4). It was used for two reasons. First, it is methodologically simpler than recognition memory in the sense that one does not have to counterbalance target/distractor. Second, test position effects, which are very strong in recognition memory experiments, do not come into play when each picture is tested immediately after it is presented. This dependent variable has been used a good deal in our laboratory. In two experimental paradigms, it yielded results that were qualitatively identical to results obtained in corresponding recognition tests.[1]

Results

Subjects had no difficulty carrying out the response instructions. Nonexistent details were written down in less than 5% of the trials. The responses almost invariably consisted of the names of objects (e.g., "a person in the middle" or "a boat in the upper left") rather than mere mention of some physical characteristic. An informal survey of the data indicated no qualitative differences in the sorts of details that were reported as a function of the various experimental conditions.

Our expectation was that the two mask variables would be differentially effective at the two delay levels. Accordingly, Table 1 shows the results as two 2 × 2 tables, one for each level of mask delay. The entries in the table are the average number of details reported for that condition. Each mean is based on 270 observations. Note that lower performance indicates more effective masking. The control condition showed the highest performance ($M = 2.35$ details).

The data for the eight conditions defined by mask delay, luminance, and attention demand are quite straightforward. The relevant statistics are included in a note to Table 1.[2] Essentially, when the mask is immediate, there is a large effect of luminance, no effect of attention demand, and no interaction. When the mask is delayed, there is, conversely, an effect of attention demand, no effect of mask luminance, and no interaction.

A changing photograph is clearly effective as a mask after a 300-ms delay. Is the noise mask similarly effective after 300 ms? The answer is yes: The difference between the average of the two delayed noise mask conditions (2.21) and the no-mask control condition (2.35) is statistically significant, $t(352) = 1.78, p < .05$.[3]

Given the data in Table 1, a tentative conclusion is that immediately following the offset of a 50-ms picture, there is only perceptual masking. Three hundred milliseconds later, there is only conceptual masking. This conclusion must, however, be tempered by the

Table 1
Results: Mean Number of Details Recalled

Delay condition	Mask luminance: Bright	Mask luminance: Dim	*M*
Attention demand: 0 ms[a]			
Noise	0.60	2.02	1.31
Photo	0.62	2.00	1.31
M	0.61	2.01	
Attention demand: 300 ms[b]			
Noise	2.22	2.20	2.21
Photo	1.77	1.94	1.86
M	2.00	2.08	

[a] Analyses of variance (ANOVAS) yielded the following: attention demand, $F(1, 352) < 1$, *ns*; luminance, $F(1, 352) = 473.66$, $p < .05$; Attention Demand × Luminance, $F(1, 352) < 1$, *ns*. [b] ANOVAS yielded the following: attention demand, $F(1, 352) = 29.60$, $p < .05$; luminance, $F(1, 352) = 1.55$, *ns*; Attention Demand × Luminance, $F(1, 352) = 2.42$, *ns*.

---

[1] In the first paradigm (Loftus, Johnson, & Shimamura, 1984), picture memory was measured for targets of varying exposure durations followed by either an immediate noise mask or a 300-ms delayed mask. For both detail recall and old/new recognition, performance for a picture shown for d ms with a delayed mask was equal to performance for a picture shown for about $d + 100$ ms with an immediate mask.

In the second paradigm (Loftus, 1984), picture memory was measured for targets of varying duration and luminance. For both detail recall and old/new recognition, performance could be characterized as a monotone function of the product of (a) target duration and (b) some monotone function of target luminance.

[2] The seven error terms stemming from the 2 × 2 × 2 portion of the design did not differ from one another by more than a factor of 2. Accordingly, we pooled all error terms to obtain a single error term, the Subject × Condition interaction (based on 352 *dfs*), equal to 0.186. This error term was used for all statistical analyses.

[3] We performed two pilot experiments that were extremely similar to the present one. These pilots produced data that were qualitatively very similar, both to each other and to the present experiment. These data are worth considering as supporting evidence for marginal effects. The first pilot was identical to the present experiment except that the noise mask was 0.5 log units brighter than the average luminance of the photo masks. In this experiment, the mean of the control condition was 2.72, whereas the mean of the two delayed noise conditions was 2.33. This difference was significant, $t(352) = 4.41$, $p < .05$. The design of the second pilot experiment was identical to that of the first except that the long ISI was 500 ms rather than 300 ms. Here the control condition and average delayed noise condition means were 2.25 and 2.12, respectively. This difference was marginally significant, $t(352) = 1.66$, $p < .05$.

problems inherent in accepting the null hypotheses of no effect of attention demand when the mask is immediate and no effect of luminance when the mask is delayed. Accordingly, we present the following power analyses.

First, consider the effects of conceptual masking. For the immediate masks, the observed magnitude of the attention demand effect is zero ($M = 1.31$, for the photo and noise conditions). The power of the statistical test (one-tailed, under the assumption that the photo mask is more effective) is 0.95 against the alternative hypothesis that the true magnitude of this effect is greater than 0.07. For the delayed masks, the observed magnitude of the attention demand effect is 0.35. The 95% confidence interval around this value ranges from 0.27 to 0.43. We can conservatively conclude that the attention demand effect for a delayed mask is at least four times bigger than the attention demand effect for an immediate mask.

Now consider the effects of perceptual masking. For the delayed masks, the observed magnitude of the mask luminance effect is 0.08. The power of the statistical test (one-tailed, under the assumption that the bright mask is more effective) is 0.95 against the alternative hypothesis that the true magnitude of this effect is greater than 0.15. For the immediate masks, the observed magnitude of the luminance effect is 1.40. The 95% confidence interval around this value ranges from 1.32 to 1.48. We conservatively conclude that the luminance effect for an immediate mask is at least nine times greater than the luminance effect for a delayed mask.

We cannot quantitatively compare the effects of attention demand with the effects of luminance, because we don't have a comparable scale for the two variables. We can, however, compare them qualitatively. When the mask is immediate, the minimum effect of luminance, 1.32, exceeds the maximum effect of attention demand, 0.07. When the mask is delayed, the minimum effect of attention demand, 0.27, exceeds the maximum effect of luminance, 0.15. Thus, the relative magnitudes of the two effects reverse as the time following picture offset progresses from 0 to 300 ms.

There is one puzzling aspect of our data, which involves the relative effectiveness of the dim photo mask following each of the two delays. It is always to be expected that the masks used in this experiment will be more effective the sooner they follow the target. This is obviously true for the bright masks. The results are hazier, however, for the dim masks. The immediate, dim noise mask is marginally more effective than the delayed, dim noise mask ($M = 2.02$ and 2.20, respectively), $t(352) = 1.98$, $p < .05$. However, the immediate, dim photo mask is slightly less effective than the delayed, dim photo mask ($M = 2.00$ and 1.94, respectively).

There are a number of possible reasons for this anomaly. We note first that the immediate, dim photo mask is a relatively ineffective perceptual mask, whereas the delayed, dim photo mask is a relatively effective conceptual mask. Because the photo is masking in two different ways at the two delay intervals, the relative effectiveness of them is not strictly predictable.

There are, in any event, two considerations which suggest that the anomaly may be due to sampling error, and that in terms of population means, the immediate, dim photo mask is slightly more effective than the delayed, dim photo mask. First, mask luminance is a continuous variable; the levels we chose were arbitrary. If our dim masks had been chosen to be brighter, the only change in the data that could have been reasonably expected was a performance decrease in the immediate, dim conditions. Second, if the delay were to be made longer than 300 ms, performance in the dim photo mask condition would approach that of the control condition, which is higher than that of the immediate, dim photo mask condition.[4]

Discussion

We found that two separate mask variables—luminance and attention demand—

[4] Again, it is worthwhile considering data from the two pilot experiments. In the first pilot, the means of the immediate and delayed, dim photo mask conditions were 1.96 and 2.12, repectively. This difference, although in the expected direction, just missed significance, $t(352) = 1.57$, $.05 > p > .10$. In the second pilot, the means of the immediate and delayed, dim photo mask conditions were 1.78 and 2.09, respectively. This difference was significant, $t(352) = 3.44$, $p < .05$. We reiterate that the long ISI in this experiment was 500 ms rather than 300 ms.

have quite different effects on picture-memory performance when the mask is presented at two different delays following the offset of a 50-ms picture. When the mask is immediate, mask luminance but not attention demand affects memory for the picture. When the mask is delayed by 300 ms, attention demand but not luminance affects memory. At the weakest level, we can thus conclude that two qualitatively different psychological processes are in operation at these two different times.

We have, of course, been making stronger assumptions about what these processes are. Along with Potter (1976), we believe it profitable to label them perceptual and conceptual processes. Thus, we view our results as constituting support for (a) the two-assumption framework of picture perception that we explicated earlier and (b) the assumptions that mask luminance affects only perceptual processing, whereas the attention demanded by the mask affects only conceptual processing.

If one accepts this constellation of assumptions, then one has both a pair of theoretical processes, useful for understanding the nature of picture perception, and a convenient tool for studying how those processes work. By suitable manipulation, for example, one could map out the time courses during which the operation of perceptual and conceptual processes occur, and specific proposals about these time courses (e.g., Potter, 1976, p. 519) could be evaluated.

A noteworthy fact implied by these assumptions and by our data is that a noise mask acts, in part, as a conceptual mask. This follows from (a) our conclusion that perceptual masking no longer occurs after 300 ms and (b) the continuing effectiveness of the noise mask after 300 ms relative to the no-mask control.

The perceptual/conceptual distinction proposed by Potter and elaborated here resembles the peripheral/central distinction proposed by Turvey (1973). Both perceptual and peripheral masking are assumed to operate at an early stage in the nervous system; thus, both are influenced by such factors as the target/mask energy ratio. Both conceptual and central masking are assumed to operate at some more central location in the nervous system and to result from competition between mask and target for central resources.

However, central masking is not the same, either theoretically or empirically, as conceptual masking. Theoretically, central processing results in pattern recognition of the target; thus, a target that is centrally masked is not pattern-recognized. Conceptual processes, in contrast, are assumed to operate on information that has already been pattern-recognized (or, in Potter's terms, on a picture that has already been identified). Turvey's central processes involve specifically tuned feature analyzers that operate on similar visual features in target and mask (see also Sekuler, 1963; Uttal, 1970), whereas Potter's conceptual processes are more general attentional mechanisms than can be allocated to a variety of quite different mental tasks (cf. Kahneman, 1973).

Empirically, central and conceptual processes operate over quite different time courses. Turvey (1973) used target-mask ISIs that were exceedingly short—a typical range was from 0 to 60 ms—and found central masking to occur only with ISIs on the order of less than 50 ms (e.g., see Turvey, 1973, Figure 8). The ISIs used to investigate conceptual masking are an order of magnitude longer. In the present experiment, we found that conceptual masking occurred after an ISI of 300 ms, and Intraub (1980, Experiment 2) found that conceptual masking occurred at ISIs of more than 600 ms.

Thus, conceptual processes fall outside the domain of Turvey's (1973) framework. Perceptual processes, however, may be viewed as incorporating both peripheral and central processes. Thus, perceptual processes could be disrupted either by peripheral or central masking. The perceptual masking that we used is probably peripheral masking; the mask we used consisted of random noise, and both target and mask were presented binocularly. We found, as did Turvey, that given a sufficiently short stimulus onset asynchrony, mask luminance strongly affected performance.

References

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.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 1–12.

Intraub, H. (1981). Identification and processing of briefly-glimpsed visual scenes. In D. E. Fisher, R. A. Monty, & J. W. Senders (Eds.), Eye movements: Cognition and visual perception (pp. 181–190). Hillsdale, NJ: Erlbaum.

Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 115–125.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Loftus, G. R. (1984). *Luminance level and picture perception.* Unpublished manuscript.

Loftus, G. R., Johnson, C. A., & Shimamura, A. P. (1984). *How much is an icon worth?* Manuscript submitted for publication.

Potter, M. C. (1975). Meaning in visual search. *Science, 187,* 965–966.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 509–522.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology, 81,* 10–15.

Sekuler, R. W. (1963). Spatial and temporal determinants of visual backward masking. *Journal of Experimental Psychology, 70,* 401–406.

Sperling, G. (1963). A model for visual memory tasks. *Human Factors, 5,* 19–31.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 80,* 1–52.

Uttal, W. R. (1970). On the physiological basis of masking with dotted visual noise. *Perception & Psychophysics, 7,* 321–327.

Received August 11, 1983
Revision received December 8, 1983 ■