Journal of Experimental Psychology: Human Perception and Performance
1976, Vol. 2, No. 2, 151–161

When Preparation Fails: Disruptive Effects of Prior Information on Perceptual Recognition

Edward E. Smith
Stanford University

Susan E. Haviland
University of California, Irvine

Lynne M. Reder
University of Michigan

Hiram Brownell
Johns Hopkins University

Nancy Adams
Stanford University

There is a conflict in the literature on selective attention. Suppose a subject is briefly presented an item followed by a multiple-alternative recognition test. If the items are pictures, the subject's performance is facilitated by presenting the alternatives beforehand (a before facilitation), but when the items are letters the subject's performance is disrupted by presenting the alternatives beforehand (a before disruption). Five experiments were conducted to resolve this conflict, and all involved a comparison of tachistoscopic recognition when alternatives were either presented beforehand or not. The first two studies showed that the before disruption with letters was not due to certain task parameters. Experiments 3–5 demonstrated that this effect was due to masking conditions. Experiments 3 and 4 revealed that the disruption occurred only when a mask is used, while the last experiment indicated the disruption effect was sensitive to the type of mask employed. Presumably, the before disruption arises because a subject erroneously considers the features of a mask along with those of the test item in arriving at a perceptual decision.

Since the turn of the century, psychologists have tried to demonstrate that an observer can facilitate his perception of a pattern by selectively preparing for it. Much of the modern evidence for such selective perception comes from tachistoscopic experiments that have employed a before–after design. In such an experiment, a pattern (say a square) is presented briefly to a subject, who then indicates which of n alternative patterns (say a square and a rectangle) was actually presented. In the before condition, the subject is allowed to study the alternatives prior to tachistoscopic presentation, and sometimes after presentation as well; while in the after condition the alternatives are available only subsequent to the tachistoscopic flash.

If perception can be facilitated by prior information, then recognition accuracy should be greater in the before than the after condition (a before facilitation); but if perception cannot be selectively

This research was supported by U.S. Public Health Service Grant MH-19705. The authors are extremely indebted to Jerry Balzano for the extensive advice he has given us on many aspects of the present research.

Requests for reprints should be sent to Edward E. Smith, Department of Psychology, Stanford University, Stanford, California 94305.

tuned, then performance should be equal in the two conditions. Many such studies have been performed (see Egeth, 1967, for a review) with almost all of them using relatively unfamiliar, nonverbal forms (such as pictures of objects clipped from magazines). The typical result has been a before facilitation (e.g., Egeth & Smith, 1967; Gummerman, 1971; Pachella, 1970), though some studies have reported no difference between before and after conditions (Gummerman, 1970).

In such studies no one had ever obtained poorer performance in the before condition, no one, that is, until Reicher (1969). As a part of a larger study, Reicher compared recognition accuracy of single letters as a function of whether the two alternative letters were presented both before and after the tachistoscopic flash or only after the flash. He found a disruptive effect of presenting the alternatives beforehand. Reicher obtained this same before disruption when words or strings of unrelated letters were used as the forms, rather than individual letters. These results are a clear contradiction to the findings obtained with nonverbal forms. Moreover, they play havoc with any current conception of the perception of verbal items.

To appreciate this, consider a summary statement of what theories of letter and word perception have in common (this summary is drawn from Smith & Spoehr, 1974). All such theories agree that the major task of the perceiver is to decide which of several alternatives the test form best corresponds to. More explicitly, if a subject is given no prior information and is simply instructed to report the identity of a briefly flashed test letter, then his main chore is to decide which of the 26 alternative letters the test item best matches. To accomplish this matching, the subject must extract features from the test item and perform certain operations on them, where the requisite number of features and operations generally increases with the number of alternatives. And it is always assumed that perceptual performance is limited by the time or capacity needed to extract these features or to perform these operations. It follows, then, that as the number of alternatives is decreased, perceptual accuracy should improve, or at worst remain unchanged when the decrease in alternatives does not result in a reduction in the required features and operations. It is this basic prediction that is blatantly contradicted by Reicher's findings of a before disruption.

While Reicher (1969, p. 279) himself was keenly aware of the contradiction between his results and those of previous before–after studies, his surprising findings seem to have been almost completely ignored. The present article picks up where the Reicher report left off and seeks an understanding of the before disruption for alphabetic material. In all, five experiments are reported in detail. The first two replicate and extend Reicher's before disruption and show it cannot be attributed to a number of seemingly plausible factors. The last three experiments deal with the nature of the masks used in before–after studies and indicate that the before disruption is due to the subject confusing the features of the test letter with those of the mask.

Experiments 1–2: Extensions of the Before Disruption

Experiment 1

We first thought Reicher's disruption effect was due to a specific aspect of his procedure. While in his after condition the two alternative letters were presented visually, in his before condition the two alternatives were presented auditorily before the test item and visually after it. Perhaps this change in modality of the alternatives, or the use of the auditory modality per se, had somehow disrupted visual recognition in the before condition. Accordingly, our first experiment attempted to replicate Reicher's finding when the alternatives in the before condition were presented visually both before and after the tachistoscopic flash of a test letter.

Method. In each of these sessions, every subject was tested in both the before and after conditions. Separate blocks of trials were used for the two conditions and all 26 letters occurred in both blocks. Consider first the procedure for the before condition. At the start of a trial, the subject was presented with two cards outside of the tachistoscope. On each there was a single letter of the same size and type font as the test letter. The subject was instructed to inspect the two letters for a few seconds as one of them would soon be presented tachistoscopically. The incorrect alternative (i.e., the one that would not be presented) was randomly selected from a set of three foils, where different trios of foils were used with different test letters.

After the subject looked at the alternatives for about 3 seconds, he fixated a small black dot in one field of the tachistoscope. This dot marked the horizontal and vertical center of the area where the test item would be presented. When he was ready for the stimulus flash, he pressed a foot pedal that produced the immediate onset of the test item. The offset of the test item was followed by the onset of a pattern mask that consisted of rows of randomly intermixed Xs and Os that covered the area previously occupied by the test letter. Immediately after the stimulus flash, the subject again looked at the two alternative letters outside of the tachistoscope and this time indicated which one he thought had been presented.

The procedure for the after condition was identical to that above except that the subject did not inspect any alternatives before fixating the dot. However, there were two blank cards in front of the subject at the start of each trial (these were the alternative cards for that trial turned face down), and after the presentation of the test item the subject turned these cards over and made his response.

The subjects were Stanford students (one male and two females), who were paid for their participation. The order of conditions was partially balanced across subjects and sessions,[1] while the order of the 26 test items within a block was randomly determined for each condition and session. To construct the test items and alternatives, block, capital, press-on letters were separately mounted on white cards. Each letter occupied the center of the visual field and subtended a horizontal angle of 0°31′ and a vertical angle of 0°32′ when viewed in a three-field Iconix tachistoscope as used in the experiment. The luminance of the presentation field was 67.8 c/m^2.

As an aid to establishing a suitable exposure duration for a subject, each subject was given 10 practice trials prior to any of the experimental trials. On these practice trials, a single digit was presented and the subject attempted to name it. The exposure durations for two of the subjects had to be changed from the first to the second session, but for each subject the duration remained constant for the last two sessions. So the results from the latter two sessions are the critical ones.[2] The exposure durations for the three subjects on the latter sessions were 22, 34, and 40 msec.

Results and discussion. In the second session, the percentages of correct recognitions for the before and after conditions were 71 and 89, respectively; in the third session, these percentages were 67 (before) and 81 (after). All subjects showed a before disruption in both critical sessions, and averaging over both subjects and sessions the effect proved significant, $t(2) = 4.43$, $p < .05$.[3]

Thus, Reicher's (1969) curious disruption effect cannot be attributed to the modality of the alternatives, and this effect remains in conflict with the before facilitation typically obtained with nonverbal forms. At this point, we considered another possible explanation for the conflicting results. In some of the studies showing a before facilitation (Egeth & Smith, 1967), this facilitation was greater when the alternatives were similar than when they were dissimilar. It is thus possible that for any type of material, the use of extremely dissimilar alternatives leads to a disruption rather than a facilitation and that the random selection of alternatives employed in Reicher's (1969) and our own experiment inadvertently resulted in such extremely dissimilar alternatives. That this was not the case, was demonstrated by a follow-up study. In essence, we replicated Experiment 1, except that in the critical

[1] The order of the before and after conditions was not perfectly balanced because this experiment also included a third condition and the study was designed to be balanced with respect to all three conditions. In this third condition, no alternatives were presented and the subject simply reported the letter he thought he had seen. The results of this report condition are irrelevant to the main concerns of the present article, and consequently no mention is made of them.

[2] In all experiments reported, the first session was used to find a suitable exposure duration for subsequent testing, necessitating some changes in a subject's duration during the session. Hence this session never offers critical data. Thus, we report only the data obtained in critical sessions, that is, sessions in which no changes were made in exposure duration.

[3] This statistical test was performed on untransformed percentage scores. The same is true of all statistical tests presented in this article. When these tests are redone on arc sine transformations of the percentage scores, there is not a single important change in results.

session each incorrect alternative was selected so as to be maximally similar to the correct letter (where similarity was determined by referring to published norms, such as those in Townsend, 1971). Again, performance was significantly poorer in the before (75%) than the after (85%) condition, $t(8) = 2.01$, $p < .05$, one-tailed. So the before disruption does not seem to depend on the similarity of the alternatives.

Experiment 2

Since the before disruption has only been demonstrated in two-choice recognition tasks, it is possible the effect obtains when the subject has to inspect multiple alternatives prior to the presentation of one of them. That is, the effect may really be triggered by a conscious comparison of alternatives, rather than by the presentation of any prior information. The purpose of the present experiment was to check whether the before disruption could be obtained in a paradigm that did not involve a comparison of prior alternatives, namely, a same–different task.

Method. In each of three sessions, every subject was tested on four blocks of trials, two for the before and two for the after condition. All 26 letters were presented in each block. The procedure in the before condition differed from that of Experiment 1 in several respects. First, the subject inspected a single alternative (a lowercase letter) for 7 seconds. He then saw a briefly flashed, lowercase letter, next the mask, and then the alternative again. He had to decide whether the alternative and test items were the same or different letters. The after condition differed from this only in that the subject did not inspect the alternative prior to the presentation of the test form, but rather stared at a blank card for 7 seconds.

On that half of the trials where the alternative differed from the test form, the alternative was chosen so as to be maximally similar to the test form. Similarity was determined by the extent to which the alternative and test item shared similar values on the following four dimensions: (a) ascending–descending (e.g., f and h are both ascending), (b) open–closed (e.g., a and o both have closed curves), (c) angular–curved (e.g., w and x are both angular), and (d) general shape (e.g., h and k were judged to be similar in general shape).

The subjects (four males and four females) were Stanford undergraduates and served for three sessions, only the last of which was critical. Within a session, the order of conditions was completely counterbalanced across subjects, but a given subject experienced the same order of conditions in all three sessions. (For any subject, this order involved either two before blocks followed by two after blocks or the reverse.) The items in this study were lowercase letters, constructed from press-on letters, and subtended a horizontal angle of 0°31′ and a vertical angle that ranged from 0°48′ (e.g., s) to 1°12′ (e.g., k) when viewed tachistoscopically. Other procedural details were as in Experiment 1, except that the exposure durations for the critical session ranged from 19 to 44 msec.

TABLE 1

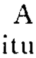
PERCENTAGES OF CORRECT SAME AND DIFFERENT RESPONSES FOR THE BEFORE AND AFTER CONDITIONS (EXPERIMENT 2)

Response	Condition	
	Before	After
Same	93	84
Different	61	88
Average	77	86

Results and discussion. Table 1 contains the percentages of correct same and different responses for the before and after conditions. The third row of this table gives the average performance levels, and here we find something of a before disruption, $F(1, 7) = 4.19$, $p = .09$. But when the data are partitioned by the type of response required (see Rows 1 and 2), a striking interaction emerges, $F(1, 7) = 22.6$, $p < .01$. For same responses, performance was actually somewhat better in the before (93%) than the after (84%) condition, but for different responses accuracy was much lower in the before (61%) than the after (88%) condition. What is happening, then, is that subjects in the before condition used the same response with far greater frequency (66%) than they should have, while subjects in the after condition used both responses approximately equally often. Consequently, overall performance shows a before disruption. Thus, while we have been able to extend the before disruption to a same–different task, we have also uncovered a constraint on this effect (i.e., the above interaction) that must be accounted for in any explanation of the effect.

So switching to a new paradigm, and using lowercase letters and similar alternatives, did not eliminate the conflict between the before disruption with letters and the before facilitation with pictures. In thinking about this conflict, one obvious difference between letters and pictures comes to mind. Letters are more familiar than pictures, and it is possible that the before disruption occurs only with extremely familiar forms. But this idea proved to be a blind alley, as indicated by two follow-up experiments.

In the first follow-up, we used a recognition task similar to Experiment 1, except that on half of the trials the test letter and alternatives were rotated 180° (on the remaining trials, they were upright). If the before disruption was due to familiarity, it should be substantially reduced with the less familiar inverted letters. In fact, the before-disruption effect was as large with inverted letters (64% recognition for before vs. 80% for after) as it was for upright ones (66% for before vs. 77% for after).

In our second follow-up study we tried a more extreme variation of familiarity. We essentially redid Experiment 2, but substituted for each letter a totally novel pattern that was constructed from the the same features as the letters. To illustrate, for A we substituted  and for B we substituted  . With such novel patterns, we expected to eliminate the before disruption, or at least markedly reduce it. But to our dismay, the results paralleled those obtained in Experiment 2. Thus, with these novel patterns: (a) there was an overall before disruption, 81% recognition for before versus 87% for after, $F(1, 10) = 6.86$, $p < .05$, and (b) same decisions were somewhat more accurate in the before than the after condition, 98% versus 88%, while different decisions were far less accurate in the before than the after condition, 65% versus 86%, with $F(1, 10) = 15.07$, $p < .01$, for the interaction. In fact, when we combined these results with those of Experiment 2 and performed a new analysis of variance, there was no effect of the type of forms used, $F < 1$.

Thus we found no evidence at all that familiarity per se mediates the before disruption. Rather, it seemed there must be something about the features common to upright letters, inverted letters, and our novel patterns that was responsible for this curious and robust effect. We suggest that this "something" is that all three types of forms contained features that were also shared by the mask.

Experiments 3–5: The Role of Masking in the Before Disruption

The Masking Hypothesis

Our proposed explanation of the before disruption is termed the masking hypothesis. Let us first illustrate it in the context of a same–different task with letters. We assume that when a subject in the before condition is given the prior alternative, he determines a set of critical features for it, looks for these features in the tachistoscopically presented item, and responds "same" if he finds them. On some same trials, however, the subject fails to find the critical features in the test item, detects them in the mask instead, and erroneously believes they were in the test item. That is, there is a failure in the subject's ability to resolve temporally the test item from the mask when they contain similar material, as was the case in all of our studies. When the features of the mask are erroneously combined with those of the test item, some of the mask's features match and some mismatch the critical features for which the subject is looking. But we assume the matching features play the greater role in the subject's decision, which leads to a spuriously high probability the subject would detect any critical feature he is looking for, which in turn leads to a bias toward same decisiosn in the before condition.

With regard to same trials in the after condition, there is no opportunity to determine a set of critical features. The subject's strategy is presumably to extract as many features as possible from the input, match these to stored descriptions of letters, compare the closest matching description to the alternative presented at the end of the trial, and respond "same" only if they match. Again, the subject would erroneously extract some features from the mask when it is similar to the test item and consider them in the matching process. Though these extra features may bias the selection of the closest matching description, there is no reason to expect them to induce a bias toward same decisions. Hence, for same trials, there should be fewer correct decisions in the after than the before condition.

Now consider different trials. By the above reasoning, the before condition should again lead to a same bias, which would now hurt performance, while the after condition again induces no special bias. Thus there should be more correct different decisions in the after than the before condition. So the masking hypothesis can account for the interaction between conditions and response type that we found in Experiment 2.

We now have to extend this hypothesis to account for the results in the recognition task of Experiment 1. The major new assumption is that the subject transforms the recognition task into a same–different one. That is, when presented with two alternatives, say A and B, the subject focuses on one of the alternatives, say A, and determines a set of features present in A but not in B. These constitute the critical features, and from here on the process is identical to that posited earlier. The subject searches the input for critical features, responding with the A alternative if he finds them and with the B alternative otherwise. (Note that the A and B responses correspond to implicit same and different decisions, respectively.) Again, some of the features detected are those of the mask, and this leads to a bias toward the A alternative (i.e., toward an implicit same response). Thus, as before, we expect this bias to lead to a before facilitation when A occurs as the test item and a before disruption when B occurs. As the alternative the subject chooses to focus on (A in our example) may vary from subject to subject and trial to trial, overall performance should be a mixture of about half before facilitation and half before disruption; since Experiment 2 indicates that the magnitude of the before disruption (for different decisions) exceeds that of the before facilitation (for sames), the final prediction is that overall performance in the recognition task shows a before disruption.

Finally, the masking hypothesis must explain why there is a before facilitation when pictures are used in the recognition task. Our explanation is that a featural description of the pictures used in these tasks might include quite different entities from the salient features in the mask, and therefore a subject should be unlikely to confuse the two sets of features. Thus the mechanism underlying the before disruption is no longer operative, and the selectivity permitted by the before alternatives leads to a facilitation.

So the masking hypothesis can account for all the relevant data, though it requires many assumptions to do so. These assumptions, however, lead to three sets of testable predictions. First, in a recognition task, the before disruption should decrease if no mask is used, because there are no extra confusing features. We tested this in the next experiment. The second set of predictions concerns the same–different task. Again we expect that if we remove the mask, we should reduce the before disruption for overall performance. But removal of the mask should also lead to more specific effects: it should reduce both the before facilitation for sames and the before disruption for differents, because both effects arise partly out of the same bias created by the mask. We tested these proposals in Experiment 4. Our final set of predictions also deals with a same–different task, and in particular with the similarity of the mask to the alternative. When this similarity is high, we expect the kind of results we obtained in Experiment 2 (where the features of the mask were similar to those of the alternative). But when the mask is dissimilar to the alternative, there should be decreases in the before facilitation for sames and the before disruption for differents. This final set of predictions was tested in Experiment 5.

Experiment 3

Two groups of subjects were tested in our recognition task, one with a mask and one without. We expected the before disruption to be less in the latter group.

Method. In each of two sessions, every subject was tested on four blocks of recognition trials. Two of the blocks were for the before and two for the after condition and all 26 uppercase letters occurred as a test item once in each block. For half of the subjects, the test item was followed by a mask of superimposed Xs, Ks, and Os (call this the mask group), while for the remaining subjects, the test item was followed by a blank postexposure field (the no-mask group). The illumination of the test field was far less in the no-mask group (5.4 c/m^2) than in the mask group (67.8 c/m^2) in order to insure that recognition accuracy was comparable in the two groups. The general procedure for the before and after conditions was the same as in Experiment 1, except that the time allotted to study the alternatives in the before conditions was 7 seconds.

The subjects (10 males and 10 females) were Stanford students. Within each group (10 subjects each), the order of the before and after conditions was balanced across subjects. There were four orders—ABAB, BABA, ABBA, and BAAB (where A and B designate the after and before conditions, respectively)—and at least 2 subjects in each group received each order in the critical session. The practice trials consisted of 30 report trials with letters, and the exposure durations in the critical second session ranged from 13 to 40 msec for the mask group and from 21 to 58 msec in the no-mask group.

Results. Table 2 gives the percentages of correct recognitions for the before and after conditions, separately for the mask and no-mask groups. The results conform perfectly to the predictions of our masking hypothesis. For the mask group, we replicated our before disruption, as accuracy was less in the before (68%) than the after (80%) condition. For the no-mask group, we finally found a before facilitation with letters, as performance in the before condition (81%) exceeded that in the after condition (76%). The interaction between conditions and groups was significant, $F(1, 18) = 13.18$, $p < .005$, while neither groups nor conditions produced a significant main effect, $F(1, 18) = 1.10$, $p > .1$ and $F(1, 18) = 2.15$, $p > .1$, respectively.

Experiment 4

In a same–different task we again used mask and no-mask groups of subjects. We expected that removal of the mask would lead to a reduction in both the before facilitation for sames and the before disruption for differents, yielding a decrease in the overall before disruption. The critical predictions are therefore in the form of interactions: a two-way interaction between groups (mask vs. no mask) and conditions (before vs. after), which reflects the prediction for overall performance, as well as a three-way interaction between conditions, groups, and response type.

Method. Each subject was tested for two sessions, in either the mask or no-mask group. The nature of these groups was identical to that in the previous experiment. The procedure for the before and after conditions was the same as in Experiment 2, except that: (a) a subject in the before condition inspected the prior alternative for 7 seconds, but had no comparable visual task in the after condition and (b) on each different trial, the alternative was randomly selected. The subjects (6 males and 10 females) were Stanford undergraduates, where 8 of them served in the mask group and 8 in the no-mask group. The balancing of the order of the before and after conditions and the nature of the practice trials

TABLE 2

Percentages of Correct Recognitions for the Before and After Conditions in the Mask and No-Mask Groups (Experiment 3)

Group	Condition: Before	After
Mask	68	80
No mask	81	76
Average	74	78

TABLE 3

Percentages of Correct Same and Different Responses for the Before and After Conditions (Experiment 4)

Response	Mask: Before	Mask: After	No mask: Before	No mask: After
Same	85	75	90	62
Different	74	91	67	85
Average	80	83	78	74

were as in the previous experiment. The range of exposure durations during the critical second session was 11–33 msec for the mask group and 24–55 msec for the no-mask subjects.

Results. Table 3 presents the percentages of correct same and different decisions for the before and after conditions, tabulated separately for the mask and no-mask groups. Note first that with regard to overall performance, there is a before disruption in the mask group, but a before facilitation in the no-mask group, $F(1, 14) = 5.48$, $p < .05$, for the interaction. This is one of the two critical interactions predicted by the masking hypothesis. However, the other prediction—an interaction between conditions, groups, and response type—was not confirmed, $F(1, 14) = 2.34$, $p > .1$. Instead, both groups showed our usual interaction of condition and response type, $F(1, 14) = 34.96$, $p < .01$. The data, then, offer at best mixed support for our masking hypothesis. We are unsure why one of our predictions failed so utterly, but this prediction was given a second chance in our next study, which offered an even more demanding test of our hypothesis.

Experiment 5

According to the masking hypothesis we should be able to reduce the before disruption for letters by using a mask whose features are dissimilar to those of the alternatives. To test this, we employed a same–different task, with two sets of alternatives and two sets of masks. One set of alternatives contained angular letters, and the other curved ones; similarly, one set of masks was constructed from angular features, and the other from curvilinear ones. It was therefore possible to form alternative–mask combinations that were either similar or dissimilar. Where they were similar, the masking hypothesis predicts that subjects should erroneously combine the features of the mask with those of the test item and produce a before facilitation for sames and a before disruption for differents. When the mask and alternative letter were dissimilar, however, the hypothesis predicts that subjects should be less likely to combine the features of the test item and mask, thereby reducing the before facilitation for sames and the before disruption for differents. Our critical prediction is that of a four-way interaction involving type of alternatives, type of mask, conditions, and response type.

Method. We first selected two sets of capital letters to serve as alternatives. One included the angular letters A, K, M, W, X, and Z, while the other contained the curvilinear letters C, J, O, Q, S, and U. We next constructed three pairs of masks (see Figure 1), such that one member of each pair contained angular features, the other curvilinear ones. Different mask pairs were used with different subjects, and every subject experienced four types of trials: angular alternatives with either an angular or a curvilinear mask and curvilinear alternatives with either an angular or a curvilinear mask.

The design and procedure were similar to those of the preceding experiment in most respects. There were, however, some changes. Each of the four blocks contained 30 trials (half same and half different). This gives a total of 120 trials per session, and each of the 12 critical letters listed above served as an alternative exactly 10 times. For the 60 different trials (where the alternative does not match the test item), we selected the test items as follows. On 48 of these trials, the test item was one of the 12 critical letters, where each critical letter appeared twice with a similar letter as the alternative and twice with a dissimilar one; on the remaining 12 trials the test item was drawn from a new letterset, including B, D, G, P, and R.

The subjects (nine males and three females drawn from our usual population) were divided into three groups of four, each group being tested with a different mask pair. Within each group, two subjects received the conditions in the order ABAB (A for after, B for before), and the other two had

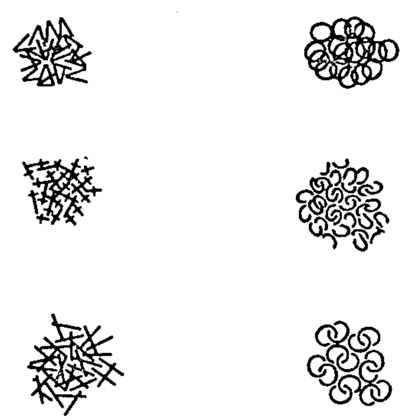


FIGURE 1. The three pairs of masks used in Experiment 5.

TABLE 4

PERCENTAGES OF CORRECT SAME AND DIFFERENT RESPONSES FOR THE BEFORE AND AFTER CONDITIONS AS A FUNCTION OF WHETHER THE ALTERNATIVE AND MASK USED WERE CURVILINEAR OR ANGULAR (EXPERIMENT 5)

	Similar mask and alternative			Dissimilar mask and alternative		
	Curvilinear alternative/ curvilinear mask			Curvilinear alternative/ angular mask		
Response	Before	After	Before–After	Before	After	Before–After
Same	79	53	26	84	69	15
Different	61	84	−23	77	94	−17
	Angular alternative/ angular mask			Angular alternative/ curvilinear mask		
	Before	After	Before–After	Before	After	Before–After
Same	82	69	13	68	74	−6
Different	69	90	−21	75	89	−14
	Average			Average		
	Before	After	Before–After	Before	After	Before–After
Same	80	61	19	76	71	5
Different	65	87	−22	76	92	−16

the reverse order. Also, within a group, one of the subjects receiving a particular order had an angular mask on the first two conditions and a curvilinear one on the last two; this assignment of masks to conditions was reversed for the other subject in that group who received the same ordering of conditions. In this way, both the order of conditions and the assignment of masks were balanced over subjects. As for practice, in addition to the 30 report trials with letters, we also included 24 trials with letters on the same-different task. All of these practice trials were tested under the same condition as that used in the first block of test trials. The exposure durations for the critical second session ranged from 18 to 46 msec.

Results and discussion. Table 4 contains the percentages of correct same and different decisions for before and after conditions. The left half gives percentages for cases where the mask and alternative were similar and the right half for the cases where the mask and the alternative were dissimilar. To understand the table, start with the top third of it, that is, with the data for curvilinear alternatives. When the mask is also curvilinear, the alternative and mask are similar, and we expect our usual before facilitation with sames and before disruption with differents. The data bear this out, except that the before facilitation is unexpectedly greater than the before disruption. When the mask is angular, the alternative and mask are dissimilar and the masking hypothesis predicts decreases in the before facilitation for sames and the before disruption for differents. The data are in line with this, as the before facilitation is reduced from 26% to 15%, while the before disruption is decreased from 23% to 17%. A comparable picture emerges for the data for angular alternatives, presented in the middle of the table. When the mask is similar to the alternative, we find a before facilitation with sames of 13%, and a before disruption with differents of 21%; when the mask is dissimilar to the alternative, the before facilitation is reduced to −6%, while the before disruption decreases to 14%. Finally, the last third of Table 4 summarizes the above results and shows that when similar alternative–mask pairings are replaced by dissimilar ones, the before facilitation is reduced from 19% to 5%, while the before disruption is reduced from 22% to 16%.

When these data are subjected to an analysis of variance, there is a significant four-way interaction between type of alternative, type of mask, conditions, and response type, $F(1, 11) = 15.22$, $p < .01$. The other important results were that angular masks led to better performance than curvilinear ones, $F(1, 11) = 7.43$, $p < .05$, and that, in general, same decisions showed a before facilitation, while different decisions manifested a before disruption, $F(1, 11) = 11.76$, $p < .01$, for the interaction.[4]

General Discussion

While the masking hypothesis does a reasonable job of accounting for the above results, there is one serious problem. When the alternative and mask were both curvilinear, the before facilitation was greater than the before disruption; but in extending the masking hypothesis to the recognition task, we assumed the before facilitation must be less than the before disruption (see p. 156). Since this theoretical extension was a bit strained to begin with, the problem is a serious one. There is also another reason to be suspicious of our masking hypothesis: Almost all subjects in all five experiments reported that their task in the before condition was qualitatively different from that in the after condition, and the masking hypothesis does not adequately capture this difference. That is, subjects reported they consciously searched for features in the before condition, but tried to perceive whole letters in the after condition. This suggests that the erroneous consideration of the mask's features, as specified by the masking hypothesis, was confined to the before condition.

What all of this leads us to is a revision of the masking hypothesis (suggested to us by Balzano, Note 1). Starting with an account of the same–different task, the revision offers exactly the same explanation of performance in the before condition as the original masking hypothesis did. In accounting for performance in the after condition, however, the revision assumes a subject uses a completely different strategy in the after condition, one that is relatively unaffected by the features of the mask. Consequently, we are led to focus only on variations in before performance, as it is only on before trials that the mask plays its disruptive role. Such trials should always manifest a same bias, and hence same decisions should be more accurate than different decisions in all before conditions that employed a mask. This was true in all tests of it: Experiment 2 (see Table 1), the mask group of Experiment 4 (see Table 3), and those conditions in Experiment 5 where the alternative and mask were similar (see the left-hand side of Table 4). Further, the revised hypothesis predicts that the superiority of same over different decisions should be reduced when the mask and alternative are dissimilar (since this reduces the source of the same bias), and this was found to be the case in Experiment 5 (see Table 4).

The one stumbling block for our revision occurs in Experiment 4, where the superiority of the same over different decisions actually increased, rather than decreased, when the mask was removed. (These results were also inexplicable under the masking hypothesis.) But the main virtue of our revision becomes apparent when we consider performance in the recognition task. There is no longer any need to assume subjects transform this task into a same–different one. Rather, the only prediction is that recognition performance should be poorer in the before than the after condition when a mask is used, because the feature

---

[4] There is one other finding that deserves mention. Recall that two-fifths of the different trials involved similar alternatives and test items, while another two-fifths included alternatives and test items that were dissimilar (see p. 158). Accuracy was better for the dissimilar trials in the after condition, $F(1, 22) = 4.45$, $p < .05$. This accounts for why after performance was better for different than same decisions in both the present study and the previous one (both contained mainly dissimilar different trials), while after performance was roughly equal for different and same decisions in Experiment 2 (which used only similar different trials).

Throughout this article we have presented the results of our same–different tasks only in terms of percentage of correct responses. It is also possible to perform Signal Detection Analyses on these data and to combine the same and different responses into a single measure of recognition accuracy, namely d'. We performed such analyses on the data from Experiments 2, 4, and 5 and found only that d' reflects the overall difference in accuracy between before and after conditions.

search strategy used in the before condition considers the mask's features, while the strategy used in the after condition does not. This accounts for the before disruption found in Experiment 1 and the mask group of Experiment 3. It is also consistent with the elimination of the disruption effect in the no-mask group of Experiment 3, since the strategy used in this before condition no longer considers erroneous information. Thus the revised masking hypothesis gives an adequate account of the relevant data and avoids some of the pitfalls that plagued its predecessor. It does this, however, at the expense of assuming a strategy difference between before and after conditions. We think the price may be worth it.

Finally, let us consider some broader implications of the last three experiments. First and foremost these studies remove the conflict between Reicher's (1969) curious before disruption and current theories of letter and word perception (see p. 152). That is, we can still maintain that when one reduces the number of possible test items, there is a concomitant reduction in both the number of features that must be extracted and the number of operations that must be performed on them. Within this theoretical context, the before disruption arises because feature extraction is not confined to the test item, and the extra features taken from the mask lead to a particular kind of bias when alternatives are presented beforehand.

A second general implication of our studies is that masking conditions, usually thought to be of only procedural interest, may sometimes be responsible for supposedly critical effects obtained in studies of letter and word perception. The present results offer one example of this. Another can be found in studies of perceptual differences between words and single letters. Johnston and McClelland (1973) found that words were more perceptible than single letters when a mask was used, but the reverse obtained when there was no mask. Again we are led to suspect that the advantage of words over letters may be partly due to the subject erroneously extracting features from the mask. Indeed, Estes (1975) has recently proposed just such an explanation of Johnston and McClelland's results. The general point is that some of what now seems most mysterious about letter and word perception may disappear when more careful consideration is given to the nature of the masking conditions.

REFERENCE NOTE

1. Balzano, J. Personal communication, June 1975.

REFERENCES

Egeth, H. E. Selective attention. *Psychological Bulletin*, 1967, *67*, 41–57.

Egeth, H. E., & Smith, E. E. Perceptual selectivity in a visual recognition task. *Journal of Experimental Psychology*, 1967, *74*, 543–549.

Estes, W. K. On the locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General*, 1975, *104*, 127–143.

Gummerman, K. Knowledge of alternatives and perception of tachistoscopic stimuli. *Journal of Experimental Psychology*, 1970, *83*, 385–390.

Gummerman, K. Selective perception and the number of alternatives. *American Journal of Psychology*, 1971, *84*, 173–179.

Johnston, J. C., & McClelland, J. L. Visual factors in word perception. *Perception & Psychophysics*, 1973, *14*, 365–370.

Pachella, R. G. *The nature of the effect of set on tachistoscopic recognition.* Unpublished doctoral dissertation, Johns Hopkins University, 1970.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 1969, *81*, 274–280.

Smith, E. E., & Spoehr, K. T. The perception of printed English: A theoretical perspective. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Potomac, Md.: Erlbaum Associates, 1974.

Townsend, J. T. Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, 1971, *9*, 40–50.

(Received June 25, 1975)