Journal of Experimental Psychology
1972, Vol. 94, No. 1, 18–24

SELECTIVE SEARCH IN DIRECTED FORGETTING [1]

WILLIAM EPSTEIN,[2] DOMINC W. MASSARO, AND LUCINDA WILDER

University of Wisconsin

Introduction of a cue that directs *S* to forget part of the presented material enhances recall of the to-be-remembered material. Selective search and selective rehearsal accounts of this effect were tested in two paired-associate probe experiments. Selective rehearsal was tested by varying the time available for rehearsal. Selective search was examined by comparing the effectiveness of the forget cue in recall and on a matching test that insured that the search set was identical with and without a forget cue. The selective search hypothesis was able to account for the findings in both experiments. The results indicate that the forget cue provides *S* with a list tag that can be used to eliminate the to-be-forgotten (incorrect) items as acceptable responses at test.

The major objective of the present study is to assess the contribution of selective search in a directed forgetting task. Directed forgetting refers to the finding that introduction of a forget cue that relieves *S* of the responsibility for retaining part of the material enhances recall of the to-be-remembered (TBR) material (e.g., Bjork, 1970; Block, 1971; Elmes, Adams, & Roediger, 1970). As an example, consider two sample trials from a study by Shebilske, Wilder, and Epstein (1971). Each trial began with presentation of two CVC-word pairs, then an interval, two additional pairs, an instruction cue, another interval, and a probe test for one of the four pairs. The cue FIRST or SECOND informed *S* that the tested pair would come from the designated set, so that the other set could be forgotten. The cue EITHER informed *S* that the tested pair would be selected from either the first or second set so that all pairs had to be retained. Recall was greater when *S* was responsible for only one set of pairs ("only" trials) than when he was responsible for both sets of pairs ("either" trials). The difference between recall on "only" and "either" trials was called the *Only effect.* Two explanations of the Only effect were examined in the present experiments.

One explanation stresses selective rehearsal. On an "only" trial the TBR set may be rehearsed selectively during the postcue interval, while on an "either" trial rehearsal time must be distributed among twice the number of pairs. Selective rehearsal implies that TBR items should be better recalled than to-be-forgotten (TBF) items. Superior recall of TBR over TBF items has been demonstrated by Reitman, Malin, Bjork, and Higman (1971) and Woodward and Bjork (1971). However, selective rehearsal alone cannot account for the results of Shebilske et al. (1971), who failed to find an increase in the magnitude of the Only effect with increases in the opportunity for rehearsal. Their results indicated that preventing rehearsal with a subtraction task during the postcue interval did not decrease the Only effect. However, in their critical comparison, increasing the opportunity for selective rehearsal did not facilitate overall performance significantly. Accordingly, it might be unreasonable to expect the Only effect to be larger since the extra rehearsal time did not lead to better performance. These results warranted another test of the selective rehearsal hypothesis.

A second account of the Only effect is that the forget cue directs *S* to search selectively among the members of the TBR

[1] This research was supported by Research Grant MH 20584-01 from the United States Public Health Service to the first author. Authors are listed in alphabetic order. All three authors contributed equally to this research. We wish to thank Wayne Shebilske and John Barresi for valuable discussions of this work.

[2] Requests for reprints should be sent to William Epstein, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

set for the correct response. The difference between recall on an "only" and an "either" trial is attributable to the difference in the size and composition of the search set on the two trials. The selective search hypothesis implies that decreasing the difference between the search sets on an "only" and "either" trial will decrease the magnitude of the Only effect. In the limiting case, when the search-set size on "only" and "either" trials is identical, the Only effect should be absent. The present experiments were also designed to test this implication.

Experiment I

The general plan of Exp. I was to compare the Only effect under two conditions of testing and two rehearsal conditions. The testing conditions were recall and matching. The matching test established the size and composition of the search set and insured that the search set would be identical in size and composition on both "only" and "either" trials. Therefore, the selective search hypothesis predicts equal "only" and "either" performance on the matching test, but superior "only" performance on the recall test. In one rehearsal condition, the postcue interval was filled with an interpolated activity; in the other condition, the inteval was blank. If selective rehearsal contributes to the Only effect, a larger effect should be observed following the blank interval under both testing conditions.

Method

Subjects.—The 192 *S*s were college students; of these, 144 participated to satisfy a course requirement and 48 were paid volunteers.

Materials and apparatus.—One hundred and eighty paired associates were prepared. The stimuli were CVC nonsense syllables selected from the Archer (1960) norms (Range 37–61). An attempt was made to include an equal number of syllables beginning with each consonant and to minimize visual and acoustic similarity. The responses were chosen from categorized lists of common English words prepared by Battig and Montague (1969). Six words were selected from each of 24 categories for use in the experimental lists. Three words were selected from each of 12 additional categories for use in the practice lists. The highest frequency words within a category which were not given as responses for another category and were between three and six letters long were chosen. Syllables and words were randomly paired with the restrictions that the words within a given category were paired with syllables having different first letters and that visual and acoustic similarity within the pairs were avoided. Subtraction problems, two digits minus one digit, served as the filler task.

All materials were presented on slides by means of a Kodak Carousel projector. A Psionix Series 1,600 logic system controlled the presentation rate.

Design.—Thirty, six-pair lists (6 practice and 24 experimental lists) were presented to each *S*. All lists were categorized; i.e., the first set of three paired associates (PAs) in each list contained words that were members of one category and the second set contained words that were members of a different category; for example, BAF-DOOR, NUJ-ROOF, TEV-WALL, and CEB-SPIDER, VAZ-BEETLE, QYT-ROACH. One PA in each list was arbitrarily chosen as the test item. The lists were presented as follows: after a 1-sec. ready signal, three pairs were presented one at a time, a subtraction problem appeared for 3 sec., three more pairs were presented one at a time, followed by a .10-sec. exposure of the instruction cue, a 3-sec. postcue interval, and finally a test slide. The pairs were presented for 2 sec. each, the test slide was presented for 10 sec., and a 2-sec. intertrial interval preceded the next ready signal.

Composition of the 3-sec. postcue interval was a between-*S*s variable; in the filled condition a subtraction problem was presented, and in the unfilled condition a blank slide was shown. When a subtraction problem appeared, *S* was required to write the answer to the problem and "O" if the answer was an odd number or "E" if the answer was an even number.

There were three within-*S*s variables: instruction cue, serial position of the test item, and type of test. For a given list the instruction cue was the word FIRST, SECOND, or EITHER. FIRST and SECOND were "only" cues in that the word FIRST indicated that the tested pair would be one of the first three pairs, and the word SECOND indicated that the tested pair would be one of the second three pairs. The cue EITHER indicated that any one of the six pairs would be tested. Each of the six serial positions was tested with an "only" cue and an "either" cue. In the experimental lists, each of the 12 cue-position combinations was tested once with a recall test and once with a matching test. On recall tests, a stimulus from the set indicated by the cue was presented and *S* was required to supply the correct word. On matching tests, a stimulus and the three words from the set indicated by the cue were presented and *S* was required to choose the word that had been paired with that stimulus. On "either" trials, the two incorrect words were from the same set as the tested PA. The words from a given set appeared on the test slide in a random order. The 24 test PAs were counterbalanced across *S*s so that each PA was tested with a recall and matching test in every cue-position combination in both the filled and unfilled conditions. This counterbalancing required 24

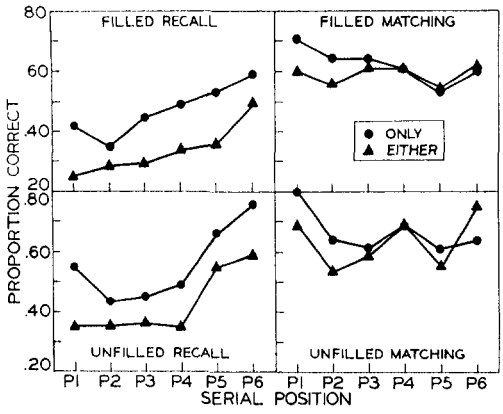

FIG. 1. Correct response proportions for recall and matching in the filled and unfilled conditions as a function of cue and serial position.

versions of the basic set of lists. The presentation sequence of each version was random with the restriction that there were no runs greater than three of the same cue, the same serial position of the test item, or the same type of test. Four *Ss* were assigned to each of the 24 versions in each of the filler conditions for a total of 192 *Ss*.

Procedure.—The *Ss* were tested in groups of one to four and were alternately assigned to the filled and unfilled conditions. A tape recording of the instructions was played which explained the cues as follows:

> The word FIRST indicates that the tested pair will come from the first set of three pairs. Therefore, you can forget the second set because it will not be tested. The word SECOND indicates that the tested pair will come from the second set of three pairs. Therefore, you can forget the first set because it will not be tested. The word EITHER indicates that the tested pair will come from either the first or second set of pairs.

It was emphasized that the cues always would correctly represent the test. The *Ss* were told that they would not know if the test would be recall or matching until after the list had been presented and were instructed to learn the pairs in such a way as to perform well on either type of test. They were encouraged to give a response for each test and were informed that the lists were categorized and that each pair would be presented only once. The *Ss* were further instructed to answer and classify all subtraction problems correctly, but not to spend any of the learning or test time on subtraction problems.

The practice and experimental lists were then presented. The six categorized practice lists were composed of pairs from categories not used in the experimental lists and the postcue filler slides corresponded to the experimental condition to which *S* was assigned. The *Ss* wrote their responses for practice and experimental lists in a booklet. After each response, *Ss* turned the page so that the previous responses were not visible when a new list was presented. The *Ss* were not allowed to change or make additions to previous responses. At the conclusion of the experiment, *Ss* were given a short questionnaire.

Results

The correct response proportions are shown in Fig. 1 as a function of cue and serial position for the filled and unfilled conditions. The proportion correct was significantly greater when a matching test was given (.630) than when a recall test was given (.447), F (1, 190) = 161.36, $p < .001$. Proportion correct was greater on "only" trials (.578) than on "either" trials (.500), F (1, 190) = 42.84, $p < .001$. However, the significant Cue × Test interaction, F (1, 190) = 16.44, $p < .01$, indicated that there was a significant 12.9% Only effect when a recall test was given, F = 52.00, $p < .001$, and a nonsignificant 2.6% Only effect when a matching test was given, F (1, 190) = 2.11, $p > .25$.

The postcue filler task lowered overall performance significantly, F (1, 190) = 8.78, $p < .01$; proportion correct was .507 in the filled condition and .570 in the unfilled condition. The reduction (8.7%) in recall performance was greater than the reduction (3.8%) in matching performance, but this difference was not significant, F (1, 190) = 2.96, $p > .10$. Although the filler task lowered overall performance, it did not reduce the magnitude of the Only effect, F (1, 190) < 1. When a recall test was given, the Only effect was 13.3% for the filled condition and 12.5% for the unfilled condition. When a matching test was given, the Only effect was 2.1% and 3.1% for the filled and unfilled conditions.

The proportion correct was .547, .474, .501, .530, .544, and .637 for Positions 1–6, respectively. The differences among the positions were significant, F (5, 950) = 9.39, $p < .01$. Multiple comparisons revealed that performance in Position 6 was significantly ($p < .01$) better than in the other five positions and that performance in Positions 1 and 5 was significantly ($p < .01$) better than performance in Position 2. The significant Test × Posi-

tion interaction, F (5, 950) = 10.84, $p < .01$, is shown in Fig. 2. Recall as a function of serial position showed a marked recency effect, while matching performance was highest in Positions 1, 4, and 6. A trend analysis revealed that there was a significant difference in the linear trend across serial position for recall and matching, F (1, 950) = 44.20, $p < .001$. This accounted for 82% of the total variation in the Test × Position interaction. Twelve percent of the remaining variance was accounted for by a significant difference in quartic trend, F (1, 950) = 6.55, $p < .05$.

Although the Filler × Position interaction was not significant, F (5, 950) = 1.44, $p > .25$, it may suggest the rehearsal strategy employed by *S*s in the unfilled condition during the 3-sec postcue interval. In the filled condition, the proportion correct as a function of serial position hovered around .50 for Positions 1–5 and increased to .60 in Position 6, as would be expected with little opportunity for rehearsal. In the unfilled condition, on the other hand, recall was high (.60) in Position 1, fell to .50 in Positions 2 and 3, and increased to .55, .60, and .68 in Positions 4, 5, and 6. This significant difference in quadratic trend, F (1, 190) = 4.31, $p < .05$, suggests that in the unfilled condition, *S*s rehearsed the first and last few items during the blank interval. The similarity of the curves for "only" trials and "either" trials suggests that *S*s employed the same strategy regardless of the cue.

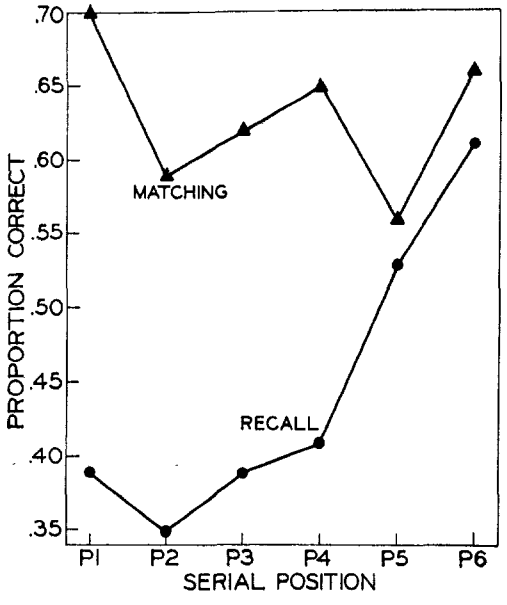

FIG. 2. Proportion correct at each serial position for recall and matching.

TABLE 1

PROPORTION OF INTRALIST INTRUSIONS FROM TO-BE-REMEMBERED (TBR) AND TO-BE-FORGOTTEN (TBF) SETS

Cue[a]	Unfilled		Filled	
	TBR	TBF	TBR	TBF
Only	.93	.07	.86	.14
Either[b]	.63	.37	.55	.45

[a] Averaged over the six positions.
[b] In the case of the "either" cue, the TBR-TBF breakdown refers to intrusions from within the probed set (TBR) as compared to intrusions from the unprobed set (TBF).

Error data.—Virtually all of the errors on matching tests were intraset intrusions; there were very few omissions. On recall tests, 20% of the errors were omissions and 80% were intrusions. A large proportion (.88) of the intrusions were intralist intrusions. Table 1 shows the proportions of intralist intrusions which were members of the TBR and TBF sets for the filled and unfilled conditions. With an "either" cue, *S* was responsible for both sets and the labels TBR and TBF merely designate the tested and not-tested sets, respectively. There was a greater opportunity for TBF intrusions than TBR intrusions because there were three untested items in the TBF set and two untested items in the TBR set. In order to correct for differential opportunity, mean numbers of intrusions were substituted for total numbers and proportions were then calculated. The proportions shown in Table 1 thus reflect the intrusion rate per item. In the unfilled condition, the proportion of TBF intrusions was .07 after an "only" cue and .37 after an "either" cue. In the filled condition, the proportion of TBF intrusions was .14 on "only" trials and .45 on "either" trials.

Questionnaire data.—The answers of the *S*s in the filled and unfilled conditions were

similar and will be described together. Eighty-six percent of the *S*s reported that they could not rehearse during intervals filled by subtraction problems. Of the remaining *S*s, 6% said they could rehearse during easy problems and 8% said they could rehearse during all problems. Ninety-two percent of the *S*s indicated that matching tests were easier than recall tests. Sixty-two percent of the *S*s reported that they did not try to anticipate whether a recall or matching test would be given. However, only 7% of the *S*s who anticipated the type of test changed their learning strategies as a function of the test they expected. When asked to rank the cues with respect to which cue(s) made the task easiest, 46% ranked them SECOND, FIRST, and EITHER, 17% reported that SECOND was easiest and that FIRST and EITHER were equal, 14% ranked the cues SECOND, EITHER, and FIRST, and 9% reported that FIRST and SECOND were equal and made the task easier than did EITHER. The remaining *S*s were approximately equally distributed among other orderings of the cues.

EXPERIMENT II

Before discussing the processes responsible for the Only effect, we must settle an interpretive ambiguity in Exp. I. It can be argued that the recognition process, rather than restriction of the search set, was responsible for the elimination of the Only effect on the matching test of Exp. I. This argument can be tested by designing a matching test that does not restrict the search set. If selective search is responsible for the Only effect, the effect should appear on the new matching test. On the other hand, if elimination of the Only effect is due to the recognition process, then the Only effect should also fail to appear on the new matching test. A matching test having these characteristics is readily achieved by including all six response items as the alternatives for matching, the TBR as well as the TBF set. Under these conditions the "only" cue can again exert a selective influence on the search process. If the cue is maximally effective, the search-set on "only" trials will include only the three items in the cued set, while on "either" trials all six items will be included. The data from this new matching test may also help sharpen our description of the operation of selective search.

Method

Subjects.—The *S*s were 24 paid volunteers from the university community.

Materials and apparatus.—The materials and apparatus were the same as those used in Exp. I.

Design.—Thirty categorized lists (6 practice and 24 experimental) were presented to each *S*. The composition of the lists and the exposure time of the slides were the same as those used in Exp. I. All *S*s received lists in which the postcue interval was filled by a subtraction classification problem. The tested item in each list was the same item that had been tested in Exp. I. The test slide contained the stimulus and the six words from both sets. The three words from the first set appeared above a dotted line and the three words from the second set appeared below this line. The order in which the words from each set appeared on the test slide was randomized, but the two sets were always segregated.

There were three within-*S* variables: instruction cue, serial position of the test item, and replication. Each of the six serial positions was tested twice with an "only" cue and twice with an "either" cue. Twelve of the test items were counterbalanced across *S*s so that each item was tested once in every cue-position combination. These items constituted the first replication. The remaining 12 items were similarly counterbalanced across *S*s and constituted the second replication. This counterbalancing required 12 versions of the basic set of lists used in Exp. I. The presentation sequence of each version was that used in Exp. I. Two *S*s were assigned to each of the 12 versions for a total of 24 *S*s.

Procedure.—The *S*s were run in groups of two. The basic procedure was the same as that used in Exp. I. The *S*s were told that the words from the first set would appear above the dotted line and that the words from the second set would appear below the line. It was again emphasized that the cue would always be accurate.

Results

The results support the selective search hypothesis. Performance on "only" trials (.691) was significantly greater than performance on "either" trials (.562), $F(1, 23) = 14.27$, $p < .01$. Replication, serial position, and all interactions were not significant. The Only effect of 12.9% is almost identical to the Only effect of 13.3% obtained in the filled recall condition of Exp. I.

Discussion

The recall test in Exp. I showed a significant Only effect, confirming the earlier findings of Shebilske et al. (1971). The matching test in Exp. I was designed to restrict *S*'s consideration of alternatives exclusively to the set containing the correct response. In these circumstances the superiority of "only" trials over "either" trials was not observed. This outcome conforms to the selective search account of the Only effect. Temporal separation and the categorical structure of the list partitions the list into two sets. The recall cue on an "only" trial identifies one of the sets as the TBR set and directs *S* to confine his search to that set. It should be clear that it is the cue that *directs* the search; the stimulus merely *initiates* the search. Given the stimulus, *S* searches the TBR set and emits the word that has the strongest association with the stimulus. On an "either" trial, the search set is not circumscribed; consequently, *S* must search the complete list, a set twice the size of the "only" set.

The distribution of errors also is evidence that a selective search can be maintained. On "only" trials, 11% of the intralist errors stemmed from the TBF set, while on "either" trials 41% of the errors originated in the TBF set. Also consistent with the selective search hypothesis is the finding reported by Shebilske et al. (1971) that the Only effect is greater for categorized lists, like those used in our experiment, than for lists whose sets contain responses that are unrelated categorically. It is reasonable that selective search will be facilitated by a variable that distinguishes the TBR and TBF sets.

However, on the matching test designed for Exp. I, the directing influence of the "only" cue is superfluous, since the search set is explicitly and exclusively designated by the set of alternatives presented on the test. Consequently, there is no advantage to an "only" cue and no Only effect. Parenthetically, we should note that the absence of an effect on the matching test in Exp. I is not due to a ceiling effect since the proportion correct on the matching test was only .63. This may seem suprisingly low, but keep in mind that the three alternatives were members of the same category and had been presented together on the learning trial. Finally, it should be noted that the presence of an Only effect in Exp. II shows that it is the identity of the search set in the matching test of Exp. I that is responsible for elimination of the Only effect and not other processes associated with the change from recall to matching. When TBF as well as TBR items were presented on the matching test, the "only" cue regained its usefulness as a direction to limit the search to the appropriate set, and consequently the Only effect reappeared.

The filler task in Exp. I reduced performance; therefore, it is reasonable to assume that rehearsal was significantly reduced. However, reducing rehearsal did not decrease the Only effect. This result confirms the earlier findings of Shebilske et al. (1971) and is compatible with their conclusion that selective rehearsal is not a necessary condition of the Only effect. Nevertheless, it is curious that the opportunity for rehearsal did not make any contribution to the Only effect. If rehearsal facilitates performance and the cue directs *S*'s processing, then selective rehearsal of the TBR set should have occurred in the unfilled condition, enhancing the Only effect. One answer to this question is that *S*s may simply rehearse responses during the rehearsal interval, which would facilitate overall performance but not necessarily contribute to the Only effect.

The selective search hypothesis can be clarified by a recent theoretical model of forgetting (Massaro, 1970). The model assumes that memory for an item is directly related to the perceptual processing of that item and inversely related to the perceptual processing of other items. Accordingly, additional perceptual processing during the memory test can interfere with memory for the correct item. In recall, *S*s have to retrieve twice as many words on "either" trials than on "only" trials. The matching test in Exp. II eliminated retrieval without reducing the size of the search set. If differences in retrieval interference on "only" and "either" trials exist, this matching test should have reduced the Only effect. Since this was not the case, differential processing required during the test is not responsible for the observed Only effect.

Given no differential interference during search, an Only effect can still occur because of a decision rule provision of the model. Assuming that *S* will select the word with the highest associative strength to the test syllable, the "only" cue circumscribes the number of valid response alternatives. An incorrect alternative may have more associative strength than the correct alternative due to a simple generalization process. When this word is a TBF item, the decision rule allows *S* to elimi-

nate it as a valid alternative on "only" trials but not on "either" trials. The decision rule thus can account for the Only effect on recall and six-alternative matching tests and accounts for its absence in the three-alternative matching task where the number of valid responses is the same on all trials.

The formal model provides a measure of memory strength, d′, that is independent of the number of valid response alternatives. If the decision rule allows six valid alternatives on "either" trials and limits the set to three on "only" trials, there should be no Only effect using this measure of memory strength. Transforming the matching probabilities in Exp. II to d′ values shows lower performance for "only" trials (d′ = 1.15) than "either" trials (d′ = 1.40). This implies that the associations are not remembered better on "only" than on "either" trials. The Only effect employing proportion correct is due to the difference between the number of valid alternatives on "only" and "either" trials. The "only" cue provides *S* with a decision rule which determines whether the association of each alternative should be considered at test. The higher d′ value on "either" trials than on "only" trials indicates that the "only" cue does not always reduce the search set to the correct three alternatives. To equate d′ values on "only" and "either" trials in Exp. II would require "only" trials to have a search set of four instead of three.

REFERENCES

ARCHER, E. J. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs*, 1960, **74**(10, Whole No. 497).

BATTIG, W. F., & MONTAGUE, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, 1969, **80**(3, Pt. 2).

BJORK, R. A. Positive forgetting: The non-interference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 255–268.

BLOCK, R. A. The effects of instructions to forget in short-term memory. *Journal of Experimental Psychology*, 1971, **89**, 1–9.

ELMES, D. G., ADAMS, C., & ROEDIGER, H. L. Cued forgetting in short-term memory. *Journal of Experimental Psychology*, 1970, **86**, 103–107.

MASSARO, D. W. Perceptual processing and forgetting in memory tasks. *Psychological Review*, 1970, **77**, 557–567.

REITMAN, W., MALIN, J. T., BJORK, R. A., & HIGMAN, B. Strategy control and directed forgetting. *Mental Health Research Institute Communication #284*. Ann Arbor: University of Michigan Press, 1971.

SHEBILSKE, W., WILDER, L., & EPSTEIN, W. Forget instructions: The effect of selective rehearsal and categorical distinctiveness. *Journal of Experimental Psychology*, 1971, **89**, 372–378.

WOODWARD, A. E., & BJORK, R. A. Forgetting and remembering in free recall: Intentional and unintentional. *Journal of Experimental Psychology*, 1971, **89**, 109–156.

(Received July 28, 1971)