

# Recognition test trials without distractors: A comparison of test trials and study trials on recognition and recall

WILLIAM P. WALLACE and KAREN A. PAGE  
*University of Nevada, Reno, Nevada 89557*

Different groups of subjects were presented with mixed lists consisting of 25 low-frequency words and 25 high-frequency words. Each subject received a five-trial sequence in which the first and fourth trials were study trials and the fifth trial was a test. This critical test trial involved free recall, recognition with an equal number of target words and distractor words randomly mixed, or recognition with all target words tested prior to testing distractor words. Trials 2 and 3 in the sequence were study trials, paced oral recall trials, or paced oral recognition trials without distractors. The total time required to complete the first four trials was constant regardless of whether Trials 2 and 3 involved study presentations, free recall tests, or distractor-free recognition tests. Typical word-frequency effects were replicated, but neither recognition nor recall on Trial 5 varied as a function of the different study or test requirements imposed during Trials 2 and 3.

Comparisons of the effects of study trials and test trials on subsequent performance has had relevance for the total time hypothesis in verbal learning, for the understanding of recall and recognition processes, and for descriptions of coding control processes. Hogan and Kintsch (1971) demonstrated that recall and recognition test trials were better for delayed recall compared to study trials, but study trials and recognition test trials were better for delayed recognition compared to earlier recall trials. Wenger, Thompson, and Bartling (1980) provided evidence indicating that recall trials could also benefit delayed recognition performance. Others have shown only small effects or no differences when contrasting effects of early recall and study trials (Birnbaum & Eichner, 1971; Lachman & Laughery, 1968). When testing expectancies are developed from prior lists rather than prior trials, the results have also been somewhat mixed. Some have shown better performance in the direction of subjects' expectations (Bernbach, 1973), some have shown that recall expectations lead to better recall and better recognition (Balota & Neely, 1980; Hall, Grossman, & Elwood, 1976; Neely & Balota, 1981), and some have found no differences (Jacoby, 1973).

Important methodological considerations have been raised in connection with assessing effects of prior test trials and developing test expectations from prior lists. Mandler and Rabinowitz (1981) reported that prior recognition tests increased the number of correct responses on both delayed recall and recognition tests. However, the increases in correct responding were balanced by similar increases in intrusions during recall and in false positives during recognition (identifying a new item as old). Whitten and Leonard (1980) demonstrated that delayed free recall was related to charac-

teristics of distractors on prior recognition tests. Neely and Balota (1981) have argued that unless testing expectations are developed from prior lists for which subjects have had equal numbers of recall and recognition expectations, greater proactive interference effects (from the inclusion of distractors on recognition tests) could confound the desired distinction between recall and recognition expectations. Their procedure involves giving subjects an equal balance of prior recall and recognition lists, informing them appropriately about what testing to expect prior to each list. Prior to the critical list, however, subjects are either correctly or incorrectly informed of the ultimate testing mode.

In the present experiment, an alternative procedure was used to contrast study trials and test trials for recognition and recall testing. Elsewhere it has been argued that distractors (words that did not appear on a study list) are not always a necessary procedural or conceptual requirement for conducting recognition tests (Wallace, 1980, in press). Actually, the distractor-free recognition test procedure may have some unique advantages in contrasting study and test trials relevant for concerns that have been raised by Mandler and Rabinowitz (1981), Whitten and Leonard (1980), and Neely and Balota (1981). New distractor words are not presented during early test trials with this procedure, hence there should be no priming during test trials of possible intrusions or false positives, and no increased potential for proactive interference. An additional important procedural advantage with the distractor-free recognition method is that it enables the researcher to equate the total time required for study trials, recall test trials, and recognition test trials. The purpose of the present experiment was to contrast the effects of study trials, recall trials, and recognition trials on subsequent recognition and recall,

with total time held constant for all contrasts. The additional variable of word frequency (common vs. rare words) was included since recall and recognition are known to be sensitive to this variable, and to provide additional data regarding the effects of word frequency in mixed lists designs (see Balota & Neely, 1980; Gregg, 1976).

## METHOD

### Materials

The word lists used in this experiment were derived from materials used in a previous study by Wallace, Sawyer, and Robertson (1978). Two separate lists of 50 words were prepared, each list consisting of 25 high-frequency (HF) and 25 low-frequency (LF) words. Half of the subjects in each condition were presented with one list form and half received the other list form. When it was necessary to include distractor words on Trial 5, words from the alternate list provided the distractors.

### Conditions

There were nine basic groups in the experiment, determined by the 3 by 3 factorial combination of the activity required on Trials 2 and 3 and the type of test on Trial 5. The experiment involved five separate trials with a given 50-word list. For all conditions, the first trial was a study trial. The second and third trials involved free recall (FR) tests for one-third of the subjects, distractor-free recognition tests (RND) for one-third of the subjects, and study trials (S) for the remaining one-third of the subjects. Trial 4 was a study trial for all subjects. The fifth trial was a critical test trial, with an equal number of subjects in each of the three preceding groups receiving a free recall test (FR-FR, RND-FR, S-FR), a recognition test with 50 target words and 50 distractor words randomly mixed (FR-RD, RND-RD, S-RD), or a recognition test in which the first 50 words tested were target words (this is equivalent to a distractor-free recognition test), followed by testing of the 50 distractor words (FR-RND, RND-RND, S-RND).

### Procedures

The experiment involved five separate trials. The running time for each of the first four trials was 100 sec. The list of words was presented on a Stowe memory drum, with each word presented for 2 sec during study trials. On FR test Trials 2, 3, and 5, the memory drum made an audible click at a 2-sec rate while a blank appeared in the memory drum window. Subjects were instructed to try to call out a different list word each time the memory drum clicked. For RND test Trials 2 and 3, the list presentation was identical to the presentation for the groups that received a second and third study trial. The list words were presented at a 2-sec rate; however, the RND test subjects were instructed to call out "yes" if they recognized the word as having been on the study list and "no" if they did not recognize the word. Subjects in the S groups were silent during Trials 2 and 3. The recognition test procedures for Trial 5 were identical to those for Trials 2 and 3, with the addition that 50 distractor words were also presented, either randomly mixed with the target words (RD) or after all the target words had been tested (RND). It should be noted that the total time equivalence breaks down slightly on Trial 5 for the RD test, as only half (approximately) of the target words are presented during the first 100 sec of the test. There was a 30-sec interval between each successive trial, during which subjects were given instructions appropriate for the next trial.

### Subjects

A total of 144 introductory psychology students at the University of Nevada, Reno, volunteered as subjects for this experiment. Sixteen subjects were randomly assigned, based

on order of appearance at the laboratory, to each of the nine groups in the experiment.

## RESULTS

### Performance on Trials 2 and 3

Six groups in the experiment had either recall or distractor-free recognition tests on Trials 2 and 3. Performance on these recall and recognition tests was compared in separate analyses to assess the effects of word frequency, trials, and the type of test subsequently given on Trial 5. The mean numbers correct are presented in Table 1. The last three rows in this table indicate that recognition performance on Trials 2 and 3 was consistently higher for LF words than for HF words. This word-frequency effect was the only significant effect in the analysis of variance on the recognition scores [ $F(1,45) = 22.88, p < .01, MSe = 26.95$ ]. The results were a little more complicated for recall scores. As can be seen in Table 1, recall scores for HF words were higher than LF words [ $F(1,45) = 31.85, p < .01, MSe = 6.54$ ]. The magnitude of this frequency effect appeared to be slightly greater on Trial 3 than on Trial 2, resulting in a significant Frequency by Trials interaction [ $F(1,45) = 6.89, p < .05, MSe = 1.21$ ]. Although HF words were always recalled better than LF words, the magnitude of the difference varied among conditions sufficiently to produce an interaction between Word Frequency and Trial 5 Test Conditions [ $F(2,45) = 4.35, p < .05, MSe = 6.54$ ]. The only other significant effect was for trials, as Trial 3 performance was poorer than Trial 2 performance [ $F(1,45) = 17.60, p < .01, MSe = 1.62$ ].

### Performance on Trial 5

The results for Trial 5 are presented in the last two columns of Table 1. As was true with Trials 2 and 3, recognition performance on Trial 5 was higher for LF words than for HF words [ $F(1,90) = 11.29, p < .01, MSe = 5.48$ ]. Although there was a slight advantage in number correct with the RND test on Trial 5, compared to the RD test, the difference was just short of statistical

Table 1  
Mean Number Correct

Groups	Trial 2		Trial 3		Trial 5	
	HF	LF	HF	LF	HF	LF
S-FR					8.50	5.94
S-RD					18.62	20.88
S-RND					21.88	22.69
FR-FR	6.50	3.50	5.69	2.00	9.94	6.75
FR-RD	5.06	3.44	4.88	2.06	20.38	21.69
FR-RND	4.06	3.69	4.00	3.00	21.12	22.19
RND-FR	14.62	18.69	14.38	18.00	8.75	7.44
RND-RD	16.00	18.50	16.69	19.25	21.31	22.25
RND-RND	16.25	19.75	15.56	20.81	22.50	22.94

Note—S = study trials, FR = free recall, RND = distractor-free recognition, HF = high frequency, LF = low frequency.

significance [ $F(1,90) = 3.83$ ]. No other effects were significant, and of particular interest, the type of study or test activity on Trials 2 and 3 did not affect Trial 5 recognitions. Similar results were obtained on the recall test. High-frequency words were better recalled than LF words [ $F(1,45) = 27.71$ ,  $p < .01$ ,  $MSe = 4.80$ ]; however, the type of study or test activity on Trials 2 and 3 did not affect recall.

The two recognition groups did not differ significantly in terms of number of correct recognitions, although subjects in the RND groups recorded an average of just over one correct recognition more than their counterparts in the RD groups. However, testing all target words before testing distractors, as was done for the RND groups, resulted in significantly fewer false positives to distractors. The RD groups averaged 4.17 false alarms, compared to 1.94 for the RND groups [ $F(1,90) = 6.40$ ,  $p < .05$ ,  $MSe = 9.32$ ]. There were no other significant effects in the analysis of false positives to distractors. Intrusion errors in recall were too infrequent to subject to statistical analysis. Ignoring repetition errors, there was only a total of five outside-list intrusion errors across all subjects.

Table 2 reveals an additional similarity between the two recognition tests. This table presents the proportions correct on Trial 5 for those items correct on either Trial 2 or Trial 3 (or both) and for those items incorrect on both Trials 2 and 3. Data from Hogan and Kintsch (1971) are also presented for comparison, although it should be noted that the comparison tests were separated by a study trial in the present experiment (Trial 4), but not in the Hogan and Kintsch experiment. Nonetheless, the proportions are very similar between the studies, with the exception of the higher proportion of recall for HF words not recognized on Trials 2 and 3 in the present study. Table 2 indicates that items recalled on Trials 2 and 3 are better recalled and better recognized on Trial 5 than items that were not recalled on Trials 2

and 3. However, correct or incorrect recognition on Trials 2 and 3 did not appear to be related to correct recall or recognition on Trial 5, except perhaps for LF words.

## DISCUSSION

The purpose of the present experiment was to compare the effects of study trials, recall test trials, and recognition test trials on subsequent recall and recognition. The procedures were designed to equate early study trials and test trials in terms of total time. In addition, a distractor-free recognition test procedure was used for recognition testing to minimize differences in potential proactive interference effects among study, recall, and recognition trials. In the five-trial acquisition sequence used in the present experiment, study and test trials were equally effective in determining performance level on the final test trial. Performance on the final recall test and performance on the final recognition test did not differ as a function of whether the second and third trials were study trials, recall trials, or recognition trials. Word frequency was the only variable that influenced performance consistently throughout the experiment. Recall was consistently higher for HF words than for LF words, and recognition was consistently higher for LF words than for HF words. The present results are in agreement with Balota and Neely (1980), and contrary to Gregg's (1976) conclusion that word-frequency effects in free recall are restricted to an unmixed list design.

The present results provide one instance in which the nature of the activity required on early trials (study, recall test, or recognition test) was not related to performance on a later test. Although previous results have been somewhat mixed, there have been several studies that have shown that different types of test trials and study trials have different effects on later performance. We do not have a compelling rationale to reconcile discrepancies in the literature. We would merely note at this time that the null results in the present experiment were observed following procedures that equated trial times and avoided the introduction of new items on early recognition test trials.

## REFERENCES

- BALOTA, D. A., & NEELY, J. H. Test-expectancy and word-frequency effects in recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory*, 1980, 6, 576-587.
- BERNBACH, H. A. Processing strategies for recognition and recall. *Journal of Experimental Psychology*, 1973, 99, 409-412.
- BIRNBAUM, I. M., & EICHNER, J. T. Study versus test trials and long-term retention in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 516-521.
- GREGG, V. Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition*. New York: Wiley, 1976.
- HALL, J. W., GROSSMAN, L. R., & ELWOOD, K. D. Differences in encoding for free recall vs. recognition. *Memory & Cognition*, 1976, 4, 507-513.
- HOGAN, R. M., & KINTSCH, W. Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 562-567.
- JACOBY, L. L. Test appropriate strategies in retention of categorized lists. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 675-682.
- LACHMAN, R., & LAUGHERY, K. R. Is a test trial a training trial in free recall learning? *Journal of Experimental Psychology*, 1968, 76, 40-50.
- MANDLER, G., & RABINOWITZ, J. C. Appearance and reality: Does a recognition test really improve subsequent recall and

Table 2  
Probabilities of Recall (FR) or Recognition (RD or RND)  
on Trial 5 Given Correct performance on Trials 2 or 3  
or Incorrect Performance on Trials 2 and 3

	HF Words	LF Words	Hogan & Kintsch
P (RD/FR)	.933	.982	.974
P (RND/FR)	.990	1.000	
P (RD/FR)	.787	.850	.808
P (RND/FR)	.816	.861	
P (RD/RND)	.846	.854	
P (RND/RND)	.889	.864	
P (RD/RND)	.865	.948	
P (RND/RND)	.910	.898	
P (FR/RND)	.336	.286	.296*
P (FR/RND)	.406	.102	.046*
P (FR/FR)	.614	.602	
P (FR/FR)	.288	.201	

\*Hogan and Kintsch (1971) used a distractor recognition (RD) test.

- recognition. *Journal of Experimental Psychology: Human Learning and Memory*, 1981, 7, 79-90.
- NEELY, J. H., & BALOTA, D. A. Test-expectancy and semantic-organization effects in recall and recognition. *Memory & Cognition*, 1981, 9, 283-300.
- WALLACE, W. P. On the use of distractors for testing recognition memory. *Psychological Bulletin*, 1980, 88, 696-704.
- WALLACE, W. P. Distractor-free recognition tests of memory. *American Journal of Psychology*, in press.
- WALLACE, W. P., SAWYER, T. J., & ROBERTSON, L. C. Distractors in recall, distractor-free recognition, and the word-frequency effect. *American Journal of Psychology*, 1978, 91, 295-304.
- WENGER, S. K., THOMPSON, C. P., & BARTLING, C. A. Recall facilitates subsequent recognition. *Journal of Experimental Psychology: Human Learning and Memory*, 1980, 6, 135-144.
- WHITTEN, W. B. II., & LEONARD, J. M. Learning from tests: Facilitation of delayed recall by initial recognition alternatives. *Journal of Experimental Psychology: Human Learning and Memory*, 1980, 6, 127-134.

(Received for publication October 8, 1982.)