Encoding Variability: A Role in Immediate and Long-term Memory?

Ruth Hipple Maki and Lynn Hasher
North Dakota State University and Temple University

Two experiments tested the effect of variable encoding on recall. In the first, critical items occupied the same or different contexts, and presumably memory chunks, across successive study trials. There was no advantage either early or later in learning for items occurring in two contexts. In the second experiment, a list of homographs was learned and retention was tested one week later. During acquisition, independent groups of subjects were provided with none, one, or two meanings for each homograph. There was no advantage, either during acquisition or at retention, for lists having two versus one encoding per word.

The theory of encoding variability has recently been invoked to explain a number of memory phenomena. Martin used it to interpret transfer (1968) and retroaction (1971) effects. It has also been invoked to explain higher recall when repetitions are distributed as opposed to massed, the MP-DP effect; and when the spacing between distributed repetitions is increased, the lag effect (Melton, 1970; D'Agostino and DeRemer, 1973). In spite of the widespread use of the theory as an explanatory concept, whether variable encoding actually does lead to superior recall can be questioned. Goggin and Martin (1970) and Williams and Underwood (1970) attempted direct tests and found no support, whereas Bevan, Dukes, and Avant (1966) did find some support, for the notion that variable encoding leads to improved recall.

Part of the difficulty in testing this theory is that few precise statements describing it are available. Furthermore, there seem to be two rather different ways of conceptualizing encoding variability.[1] Melton (1970) described one type of encoding variability that will here be referred to as *contextual* variability. He noted that items repeated under distributed conditions, particularly at long lags, occur in the context of two unique sets of words, while items repeated under massed conditions occur in only one context. The distributed items then have an advantage at recall, because verbal contexts can function as retrieval cues and the distributed items have two sets of cues while the massed items have only one. The

American Journal of Psychology
1975, Vol. 88, No. 2, pp. 217–231

second type of encoding variability, here called *referential* variability, is similar to the concept proposed by Martin (1968, 1971). Referential variability occurs when two different semantic interpretations of a word are stored after successive presentations. The word 'fan' serves as an example; it may refer to a cooling device or to an enthusiast. When such a word is variably encoded, both meanings are stored.

The major difference, then, between these types of encoding variability is that with contextual variability, one semantic representation of a word is stored in two different memory chunks, whereas with referential variability, the semantic representation itself varies and the item may or may not occur in more than one memory chunk. The present experiments were attempts to operationalize these two types of encoding variability in order to study their unique contributions to memory.

EXPERIMENT I

Experiment I was aimed at investigating contextual variability by testing whether a word that is repeated in different contexts is actually remembered better than one repeated in the same context. Two previous experiments (Wood, 1972; Gartman and Johnson, 1972) seem to have operationalized Melton's (1970) description of encoding variability in MP-DP and lag experiments; that is, words were repeated in either the same context or in varied contexts. Wood (1972) presented homographs (words with two distinct meanings, e.g., 'cardinal') along with other members of their categories (e.g., 'birds') and repeated them either with members of that same category or with members of a different category (e.g., 'clergymen').

While Wood found no difference in recall as a function of constant or varied context, it might be argued that the repetition must occur within a trial, as in the typical MP-DP experiment, in order for a change of context to be beneficial. Indeed, Gartman and Johnson (1972) also repeated homographs in either the same or different contexts within a single trial and found that repetition in different contexts resulted in much better recall than repetition in the same context. There are two important problems with these studies. First, it is difficult to generalize the results of these studies to standard MP-DP experiments in free recall, since homographs may have properties different from those of ordinary words with a single meaning. Second, there was an unassessed confounding of contextual and referential variability in both studies.

In order to study the effects of contextual variability, referential variability must be minimized. Consequently, words with only one meaning are the most appropriate stimulus materials. Bower, Lesgold, and Tieman (1969) varied context with nonhomographs as materials, but they did so for all items rather than for a critical few. Such a manipulation does not conform to Melton's description of contextual variability but is reminiscent of the difference between random and constant orders of presentation in multiple-trial free recall, where, as in the study by Bower et al., there is an advantage for constant orders of presentation (Postman, Burns, and Hasher, 1972).

Experiment I was designed to further investigate the role of contextual variability in free recall. Because of this, words with only one meaning were used. These were repeated either within a single study trial or across successive study trials, with the context of a specific word remaining constant or changing.

METHOD

Materials and procedure

Two lists of 48 words were chosen. The words ranged from 2.17 to 6.63 on concreteness, from 4.16 to 7.36 on meaningfulness (Paivio, Yuille, and Madigan, 1968), and their frequencies ranged from 0 to 1,599 according to the Kučera and Francis (1967) norms. The words in each list were randomly grouped into 4 three-word sets, 4 four-word sets, and 4 five-word sets: a total of 12 groups of words. The size of these groups was varied because subjects were urged to use the groups for their organization and it is probable that the size of chunks formed by subjects is variable. The words of a group were presented in a column on a slide.

In each group of words, two critical words were randomly chosen, one of which served as a critical word for a given subject. This word was not marked in any way on the slide presented to the subjects. For the subjects given a '*same context*,' the same groups were repeated, with the order of the words within each group constant across successive presentations. For the subjects given a '*different context*,' each critical word randomly moved to a different group on repetition and took the serial position of the critical word which had been in that group on the first presentation: for example, if words designated as ABCD were presented as one group and FGH as another group, words ABGD might be presented on repetition, with G as a critical word. Note that the noncritical words of a group (ABD) remain in the same serial position within the group. The design was balanced so that across subjects a critical word was a member of a particular group on the first and second presentation equally often. With two lists, two critical words selected for each word group, and two orders of appearance of a particular critical word in a particular word group, eight subjects per condition were necessary to balance the design.

Slides were presented via a Kodak 860H Carousel projector so that each word was shown for 2 sec; slides with three-word groups were shown for 6 sec, four-word groups were shown for 8 sec, and five-word groups for 10 sec. The slide durations were controlled by pulses from a Wollensak 2550 tape recorder. The subjects had 4 min for written recall.

Table 1 shows the sequences of study and test trials. For the *'double sequences,'* the 12 groups of words were presented and then immediately repeated with no intertrial interval. In the double/same condition, the words within the groups retained their positions but the groups occurred in a different order. In the double/different condition, the groups were repeated in a different order and each critical word was moved to occupy a different group. This double study trial was followed by a test trial and then by two more double study and test trials, for a total of three test trials. For the *'single sequences,'* that is, the single/same and single/different conditions, the 12 groups were presented once and were followed by a test trial. Two more study/test trials were then given, for a total of three test trials. These were followed immediately by a double study and test trial. If study and test trials are each treated as an event, then recall on events 6 and 9 will have been preceded by the same number of events and can be compared for the four conditions.

Subjects

A total of 64 University of California undergraduate volunteers were randomly assigned to the four conditions, 16 subjects per condition. Subjects participated as a part of a course requirement. Assignment to conditions was in blocks of four with one subject per block assigned to each condition.

RESULTS AND DISCUSSION

Clustering

The logic of the comparison between the 'same' and 'different' contexts depends on the subjects having used the input groups as a basis for their organization. Consequently, a clustering analysis was performed on the first test trial of the conditions involving a 'single sequence.'

A pronounced recency effect was observed in recall; that is, in many cases, subjects first recalled the final group of words as a unit. Because this recency effect would have artificially inflated the clustering measure,

Table 1. Sequences of study and test trials; Experiment I

	Event								
	1	2	3	4	5	6	7	8	9
Double sequences									
Same and different contexts	study	study	test	study	study	test	study	study	test
Single sequences									
Same and different contexts	study	test	study	test	study	test	study	study	test

a correction for recency was performed before the clustering analysis. Tulving and Colotla (1970) suggested that memory span depends both on the number of items presented and on the number recalled between the presentation and recall of a given item. Following this suggestion, a word that was recalled within a total of seven input and output words was identified as a recency item. Because some groups of four and five words were recalled as a cluster, this estimate was extended to include an entire cluster if any part of it was within seven input and output words. Any words identified as recency items were dropped from total recall for the clustering analysis. Intrusions and repetitions were counted in total recall.

Table 2. Raw mean *z* scores corrected (and uncorrected) for recency for single/same and single/different conditions on the first two recall trials; Experiment I

	Trial 1	Trial 2
Single/same	2.84 (3.56)	5.34 (5.86)
Single/different	2.27 (2.64)	3.40 (4.42)

Table 2 shows the raw mean *z* scores obtained by comparing the number of category runs with that expected by chance for each subject (Frankel and Cole, 1971). Scores uncorrected for recency are shown in parentheses. For trial 2, clustering by the groups of words actually presented on that trial is shown in the second column of the table. Since a z of 1.96 shows significant clustering [$p = .05$], all z scores in Table 2 are significant. A two-way analysis of variance on the recency-corrected clustering scores for the two contexts on trial 1 and trial 2 showed that only the effect of trials was significant [$F(1, 30) = 12.22, p < .01$]. Subjects apparently did use the input groups as a basis for organization in both conditions involving a single sequence and across two trials. Presumably, subjects in the conditions involving a 'double sequence' also used the experimenter-defined groups. It is, then, likely that changing words in the double/different condition increased the probability of contextual encoding variability in that condition relative to the double/same condition. Therefore, by comparing recall for the double/same and double/different conditions, we should be able to see the role of contextual variability in memory when a repetition of each word occurs before a test.

Recall

The mean number of critical words recalled in each condition is shown in Table 3. The event numbers refer to the cumulative number of study and test trials. Those events that were not recall trials for any condition

Table 3. Mean critical and total word recall as a function of condition and event; Experiment I

	Event				
	2	3	4	6	9
Critical-word recall					
Double/same condition		3.88		7.00	8.06
Double/different condition		3.81		7.38	8.31
Single/same condition	3.12		5.50	7.56	8.56
Single/different condition	3.25		5.12	7.38	8.31
Total word recall					
Double/same condition		15.00		25.81	33.00
Double/different condition		14.06		26.00	33.00
Single/same condition	12.62		22.69	30.69	35.69
Single/different condition	11.31		20.44	27.19	32.50

were eliminated. The crucial comparison for the test of contextual encoding variability is between the double/same and double/different conditions on event 3, the first recall trial for both conditions. There was no difference between the groups [$F < 1$]. Whether a word occurred twice with the same set of other words or with a different set of words on each occurrence did not appear to influence the recall of that word.

At higher levels of learning, that is, at events 6 and 9, the effects of same versus different context and study/test sequence can be analyzed. Here, too, recall did not differ as a function of context [$F < 1$]. Neither was there a difference produced by the sequence of study and test events [$F < 1$]. Thus, there was no evidence in this study, either early in learning (event 3) or later in learning (events 6 and 9), of an advantage in recall as a result of changing the verbal contexts between successive occurrences of a word.

Total recall scores can also be seen in Table 3. On event 3, the effect of same versus different context in the double sequence was not significant [$F < 1$]. At events 6 and 9, neither the main effect of context [Fs < 1] nor the main effect of study/test sequence [$F\text{s}(1, 60) \leq 3.33, p > .05$] was significant. The interaction between context and study/test sequence was also nonsignificant at events 6 and 9 [$F\text{s}(1, 60) \leq 1.23, p > .10$]. Earlier research (e.g., Lachman and Laughery, 1968) showed that an equal number of preceding events yielded an equal level of recall independent of the study and test nature of these events; this conclusion was not modified by the presence or absence of contextual variability.

Increasing the probability of contextual encoding variability by varying the group in which a critical word appeared did not increase recall either in a simulation of an MP-DP experiment (the double sequence) or in multiple-trial free recall (the single sequence). The assumption that words presented in more than one context are stored in more than one memory chunk and have a higher probability of being recalled underlies the contextual version of the theory of encoding variability. Because recall was not increased by varying context, the validity of this assumption is questioned. However, in another experiment that manipulated contextual variability, Nelson and Hill (1974) found that long-term retention of words was improved by forcing subjects to learn two serial organizations for them, as opposed to only one serial organization. This experiment, which produced results apparently inconsistent with those of the present experiment, will be discussed in detail later.

EXPERIMENT II

While there was no evidence in Experiment I for an increase in recall with an increase in the opportunity for contextual encoding variability, the second type of encoding variability, namely, referential variability, may be a more potent determinant of retention. Referential variability refers to the storage of different attributes of words on successive presentations. It was assumed to have been minimal in Experiment I, since all the critical words had only one distinct meaning. In Experiment II, we attempted to manipulate the probability of referential variability while holding contextual variability at a minimum.

Variations in referential encoding should be easily produced by using homographs, words that have more than one distinct meaning, as stimulus materials. It is this type of encoding variability that Madigan (1969) proposed as a factor in the MP-DP and lag effects. If such variability does improve recall in the MP-DP task, it should do the same in a standard free-recall task. In addition, it might well be expected that multiple encodings would allow for superior long-term retention as compared to unitary, stable encodings.

In Experiment II, subjects learned and were tested for their retention of a single list of homographs. Encoding was manipulated in the manner used in earlier experiments on the MP-DP and lag effects (e.g., Johnston, Coots, and Flickinger, 1972). That is, for each target word in a list, a cue word associated with that target was presented. The extent of referential encoding variability presumed to occur during learning was manipulated in the following manner: a subject was given none, one, or two of the meanings of each homograph. If two cues were provided, a further attempt to influence variability was made by either presenting both cues on every trial or by alternating the two different cues on successive trials. This latter condition was assumed to provide the greatest opportunity for referential encoding variability.

METHOD

Design

The subjects learned 20 homographs under one of four conditions and their retention was tested either immediately after learning or one week later. Three of the four conditions provided subjects with cues that biased a particular meaning of the homograph. In one condition, *'stable/1,'* the subject was given the same associate for each word on every trial. In a second condition, *'stable/2,'* the subject was given two associates for each word, one biasing each of the two meanings of the word. In the third cued condition, *'unstable/2,'* the subject was given two cues, each biasing a different meaning; these cues, however, did not occur together but alternated on successive trials. A final condition, *'uncued control,'* was also included in the design, which was thus a 4 (cue conditions) × 2 (retention intervals) factorial.

Materials and procedure

Two lists of 20 homographs were selected from Cramer's (1970) norms. For each homograph two associates were selected to serve as cues. One of these was a strong associate designating one meaning of the homograph and the other was a strong associate designating the second meaning. The homographs and their associates were then divided into two sets such that in the first set, half the homographs had the stronger of the two associates as a cue and half had the weaker associate. The remaining 20 associates then served as cues in the second set. In the total experiment there were two lists of homographs, each list with two sets of cues, for a total of four unique combinations. Each of these combinations was used equally often in the stable/1 condition. In the unstable/2 condition, the order of these sets of cues alternated on *successive* trials, so that there were only two unique lists for this condition.[2] There were also only two unique lists for the stable/2 condition, where both sets of cues were presented on *every* trial, so that each homograph was paired with two associates which designated its two dominant meanings.

The homographs were presented centered on the screen, in capital letters. All cues were shown in lowercase letters at the top of the screen. Words were presented on the same order on every trial, with four unique serial orders used for each list. This was done to maximize the opportunity for the subject to use the cue(s).

First, each of the subjects learned 20 homographs, under one of four cue conditions, by the method of free recall to a criterion of 16/20 correct plus one trial. A word was presented by a slide projector for 3 sec on the study trial; recall on the test trial was written and the subject was allowed 1½ min. Then, retention was tested, either immediately after learning or a week later, by allowing the subject 4 min to write down as many words as he could. No cues were given during recall or at retention.

Subjects

There were 16 subjects in each group, for a total of 128 in the entire experiment. Subjects were recruited from introductory psychology courses at Temple University, or else they were comparable undergraduates who received $2.00 for their participation. Subjects were assigned to groups randomly such that each group was run once before any other was repeated.

RESULTS AND DISCUSSION

Acquisition

The mean trials to the 16/20 criterion of the eight groups are shown in Table 4. A two-way analysis of variance showed that neither the main effect of cue condition or of retention interval nor their interaction was significant [all Fs ≤ 2.31, $p > .05$]. Table 4 also shows the mean numbers recalled on the postcriterial trial. Analysis of variance again revealed no differences among cue conditions [Fs < 1]. Thus, despite the differential availability of cues during learning — and presumably, therefore, the variability of referential encoding — there was no difference either in speed of acquisition or in performance at the end of learning.

The lack of differences in speed of acquisition may indicate that increasing the probability of multiple encodings has no effect on learning, or perhaps that the effect occurs early in learning, or perhaps that subjects simply ignored the cues, which would mean that the probability of multiple encodings was not manipulated after all. In order to evaluate these alternatives, a comparison of recall among the cue conditions was done on trial 1. The mean trial 1 recall for the uncued condition was 9.03, while the means for the two conditions that provided only one cue on each trial (stable/1 and unstable/2) and for the condition that provided two cues on every trial (stable/2) were 7.58 and 7.94, respectively. Planned comparisons showed that these three cued conditions had poorer recall than the uncued control condition [$F(1, 125) = 5.66$, $p < .05$] and that the stable/1 and unstable/2 conditions did not differ from the stable/2 condition [$F < 1$]. The difference between the conditions with cued and uncued input indicates that the subjects did not simply ignore the cues. Since cues during input actually reduced recall, some attention must have been directed to them. A similar negative effect of input cues was found by Freund and Underwood (1970) and is consistent with the idea that cues during acquisition function as 'excess baggage.' In any event, at no point in the acquisition of these lists, early or later, did the increased probability of multiple representations appear to facilitate performance.

Retention

Mean retention for the four cue conditions is shown in Table 4. Loss scores (i.e., the difference between the number correct on the last trial during acquisition and that on the test of retention) are also shown in Table 4. Analyses of variance on both dependent variables revealed the same effects, which will here be reported for the more sensitive loss scores. There were no differences among the cue conditions, nor did they interact with the retention interval [all Fs < 1]. Significant forgetting was observed across the week's interval [$F(1, 120) = 214.16$].

Subjects' reports

While it is not possible to positively determine whether subjects actually stored two meanings of the homographs, the subjects' reports suggest that they were more likely to do so in the stable/2 and unstable/2 conditions than in the stable/1 condition. After the delayed test of retention, those subjects were given an alphabetical listing of the homographs and asked to write down the meanings they had used in learning the words. The numbers of words learned with one meaning (including cases in which the subject wrote two words with similar meanings such as 'good' and 'well' for FINE) and of those learned with two meanings (such as 'good' and 'tax' for FINE) were calculated for each subject. Out of 20 words, subjects in the stable/1 condition reported having learned a mean of 2.00 words with two meanings, while subjects in the stable/2 and unstable/2 conditions reported learning 4.56 and 6.44 words by using two meanings. Although there were no differences in recall between the cued conditions, it appears that subjects in the stable/2 and unstable/2 conditions were more likely to make use of the referential variability than were subjects in

Table 4. Performance during acquisition and retention; Experiment II

	Trials to criterion	Postcriterial trial	Retention	Loss
Immediate retention				
Uncued control condition	2.62	17.00	17.06	−.06
Stable/1 condition	3.62	16.37	16.56	−.19
Stable/2 condition	3.12	16.68	16.25	.43
Unstable/2 condition	2.93	15.69	16.12	−.43
Delayed retention				
Uncued control condition	2.94	16.81	9.81	7.00
Stable/1 condition	3.00	16.50	7.31	9.19
Stable/2 condition	3.12	16.00	8.50	7.50
Unstable/2 condition	4.12	16.56	8.88	7.68

the stable/1 condition. It might also be expected that recall should be a function of the number of words assigned more than one meaning regardless of the cuing manipulation. Across the cued conditions, subjects who assigned single meanings to 16 or more words ($N = 26$) retained 8.27 homographs, while subjects who assigned single meanings to fewer than 16 words ($N = 22$) retained 7.82 homographs. This difference is in the wrong direction to support the referential version of the theory of encoding variability.

Assessed in terms of both immediate and long-term retention, varying the conditions under which referential encoding variability was more or less likely to occur did not influence performance. That is, there was no evidence of superior retention for those subjects provided with two associative cues and two potential encodings as compared to others provided with one or none. Pilot research done by Esrov (Underwood, 1972) also failed to find differences in recall between conditions given different associates across learning trials compared to those given the same cues.[3] Similarly, Tulving and Osler (1968) did not find better recall when they provided two associates as cues for target words as compared to one. In contrast to these findings, Bevan, Dukes, and Avant (1966) found superior immediate and long-term recall when adjectives were used to vary encoding. As Bevan and Dukes (1967) suggested, it may be that the cues and to-be-remembered words must form a coherent unit (e.g., nouns and appropriate descriptive adjectives) in order for cue variation to aid free recall. However, if the conditions under which encoding variability is beneficial are that specific, it is difficult to see how such a theory can explain the effects of distribution in MP-DP experiments.

GENERAL DISCUSSION

Two experiments tested two versions of the theory of encoding variability. In Experiment I, contextual variability was introduced by having critical words occur in either one or two groups of other words on a presentation trial. Recall was no better for the words that occurred in two contexts than for those that occurred in one. A similar effect was seen when context was varied across study trials rather than within a single trial. While it is impossible to determine whether critical items actually occupied one or two memory chunks, the conditions under which this should occur were established. If the contextual version of the theory of encoding variability were correct, subjects given different contexts within or across study trials should have shown a higher level of recall than subjects given the same context on all study trials. They did not.

The present results, however, stand in marked contrast to a recent demonstration of facilitated long-term retention with contextual variability (Nelson and Hill, 1974). In that study, subjects who learned a list of words in two different serial orders showed better retention over a seven-week interval than subjects who learned the list in one serial order. Subjects with two serial orders had two contexts for each word, while subjects with one serial order had only one context per word. These orders, or contexts, were presumed to provide two retrieval routes for the list items and so to boost retention.

While this appears to be strong evidence of contextual encoding variability, an alternative explanation should be offered. This is based on a consideration of processes operating during the acquisition of the second serial order in Nelson and Hill's experiment. First, if one serial order is learned, and then another order of the same items, negative transfer results (Young, 1968) along with, presumably, some unlearning. Second, Nelson and Hill used a transfer procedure in which the acquisition of the second order alternated with continued trials on the first order. Their subjects thus experienced successive cycles of unlearning and relearning. There is evidence from earlier studies that superior retention results under comparable conditions; that is, when the original learning is subjected to unlearning and is subsequently relearned (Abra and Roberts, 1968; Hasher, 1971). The superior retention shown for the two orders in Nelson and Hill's study may thus have been the product of interference effects and not of variability in contextual encoding.

In Experiment II, referential variability was introduced by using homographs. Giving the subjects one versus two cues to the homographs' meaning(s) made no difference in either acquisition or retention. Whether the subjects actually stored two representations of the homographs in memory cannot, of course, be determined, yet the conditions under which this should occur were established. Under these ideal conditions for referential encoding variability, there was no evidence that recall was increased. How, then, can such an explanation be appropriate for other memory phenomena?

A review of the literature presents us with substantial amounts of controversial evidence on the theory of encoding variability. The evidence against the operation of encoding variability in transfer and retroaction is fairly strong (see Postman and Underwood, 1973). Within the context of MP-DP and lag effect studies, the results are notable chiefly for their own variability. In several studies, attempts to eliminate the MP-DP and lag effects by forced encodings have met with failure in free recall (e.g., Madigan, 1969; Johnston, Coots, and Flickinger, 1972); while in other, similar instances, such attempts have met with success (e.g., Gartman and Johnson, 1972; D'Agostino and DeRemer, 1973).

In the present experiments, neither providing two contexts for a target word nor providing two associative cues during learning seemed to boost the subjects' memory of a target word. These are conditions which, according to the theory of encoding variability, should have increased the number of retrieval routes to the target words and thereby improved memory. That there was no evidence for a beneficial effect of either contextual or referential variability must serve as a caveat against the extension of this model of information storage to the explanation of other memory phenomena, such as the MP-DP and lag effects, transfer, and retroactive interference.

Notes

Experiment I was conducted by the first author as one of a series of experiments submitted to the University of California, Berkeley, as partial fulfillment of the requirements for the Ph.D. Appreciation is extended to Dr. Leo Postman, chairman of the dissertation committee, and to Dr. Geoffrey Keppel and Dr. William Rohwer for their helpful suggestions. The research was supported by NIH Training Grant USPHS 5 TO1 GM 1207 from the National Institutes of Mental Health. It was conducted at the Institute of Human Learning, Berkeley, California. Parts of Experiment I were presented at the meetings of the Midwestern Psychological Association, Chicago, 1974. Experiment II was a collaborative effort and was supported with funds from the Temple University Faculty Senate Committee on Research. Reprint requests should be sent to Ruth Hipple Maki, Department of Psychology, North Dakota State University, Fargo, North Dakota 58102. Received for publication September 20, 1974.

1. A similar distinction has been made by Edwin Martin (personal communication, January, 1973; December, 1973) and by Melton (1973).

2. Because learning was to criterion plus one trial, whenever this last, postcriterial, trial occurred on an odd-numbered trial for a subject in the unstable/2 condition, the two cue sets did not occur equally often. An attempt was made to minimize the impact of any cue-set preference by having half the subjects in these conditions start learning with cue set 1, half with cue set 2.

3. Personal communication (Underwood, October, 1973).

References

Abra, J. C., and Roberts, D. 1969. Unlearning and relearning. *Journal of Experimental Psychology* 81:334–339.

Bevan, W., and Dukes, W. F. 1967. Stimulus variation and recall: The role of belongingness. *American Journal of Psychology* 80:309–312.

Bevan, W., Dukes, W. F., and Avant, L. L. 1966. The effect of variation in

specific stimuli on memory for their superordinates. *American Journal of Psychology* 79:250–257.
Bower, G. H., Lesgold, A. M., and Tieman, D. 1969. Grouping operations in free recall. *Journal of Verbal Learning and Verbal Behavior* 8:481–493.
Cramer, P. 1970. A study of homographs. In *Norms of word association,* ed. L. Postman and G. Keppel. New York: Academic Press.
D'Agostino, P. R., and DeRemer, P. 1973. Repetition effects as a function of rehearsal and encoding variability. *Journal of Verbal Learning and Verbal Behavior* 12:108–113.
Frankel, F., and Cole, M. 1971. Measures of category clustering in free recall. *Psychological Bulletin* 76:39–44.
Freund, J. S., and Underwood, B. J. 1970. Restricted associates as cues in free recall. *Journal of Verbal Learning and Verbal Behavior* 9:136–141.
Gartman, L. M., and Johnson, N. F. 1972. Massed versus distributed repetition of homographs: A test of the differential-encoding hypothesis. *Journal of Verbal Learning and Verbal Behavior* 11:801–808.
Goggin, J., and Martin, E. 1970. Forced stimulus encoding and retroactive interference. *Journal of Experimental Psychology* 84:131–136.
Hasher, L. 1971. Retention of free recall learning: The whole-part problem. *Journal of Experimental Psychology* 90:8–17.
Johnston, W. A., Coots, J. H., and Flickinger, R. G. 1972. Controlled semantic encoding and the effect of repetition lag on free recall. *Journal of Verbal Learning and Verbal Behavior* 11:784–788.
Kučera, H., and Francis, W. N. 1967. *Computational analysis of present-day American English.* Providence, R.I.: Brown University Press.
Lachman, R., and Laughery, K. R. 1968. Is a test trial a training trial in free recall learning? *Journal of Experimental Psychology* 76:40–50.
Madigan, S. A. 1969. Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior* 8:828–835.
Martin, E. 1968. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review* 75:421–441.
Martin, E. 1971. Verbal learning theory and independent retrieval phenomena. *Psychological Review* 78:314–332.
Melton, A. W. 1970. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior* 9:596–606.
Melton, A. W. 1973. The concept of coding in learning-memory theory. *Memory and Cognition* 1:508–512.
Nelson, T. O., and Hill, C. C. 1974. Multiple retrieval paths and long-term retention. *Journal of Experimental Psychology* 103:185–187.
Paivio, A., Yuille, J. C., and Madigan, S. A. 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph* 76:(1, part 2).
Postman, L., Burns, S., and Hasher, L. 1970. Studies of learning to learn: 10. Nonspecific transfer effects in free-recall learning. *Journal of Verbal Learning and Verbal Behavior* 9:707–715.
Postman, L., and Underwood, B. J. 1973. Critical issues in interference theory. *Memory and Cognition* 1:19–40.

Tulving, E., and Colotla, V. A. 1970. Free recall of trilingual tests. *Cognitive Psychology* 1:86–98.

Tulving, E., and Osler, S. 1968. Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology* 77:593–601.

Underwood, B. J. 1972. Are we overloading memory? In *Coding processes in human memory,* ed. A. W. Melton and E. Martin, pp. 1–23. Washington, D.C.: Winston and Sons.

Williams, R. F., and Underwood, B. J. 1970. Encoding variability: Tests of the Martin hypothesis. *Journal of Experimental Psychology* 86:317–324.

Wood, G. 1972. Organizational processes and free recall. In *Organization of memory,* ed. E. Tulving and W. Donaldson, pp. 49–91. New York: Academic Press.

Young, R. K. 1968. Serial learning. In *Verbal behavior and general behavior theory,* ed. T. R. Dixon and D. L. Horton, pp. 122–148. Englewood Cliffs, N.J.: Prentice-Hall.