JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 12, 707–721 (1973)

Effects of Three Types of Repetition on Cued and Noncued Recall of Words[1]

ROBERT C. MATHEWS

Yale University

AND

ENDEL TULVING[2]

Yale University and University of Toronto

Three list-item memory experiments explored the effects of three types of repetition of a basic unit consisting of two words in an explicitly designated conceptual category. The types were repetition of the complete basic unit, repetition of the category name with two new instances, and repetition of the category name only. All types of repetition showed a large facilitative effect on recall of the basic unit. The effect was considerably larger when repetitions were spaced rather than massed, and it manifested itself mainly in enhanced accessibility of higher-order units (categories) rather than access to elements within these units (words within accessible categories). The results as a whole suggest that the structure of the higher-order memory units is hierarchical, with access to the unit possible only through its control element.

A basic phenomenon of memory that continues to occupy the thoughts of experimenters and theorists concerns the facilitative effect of repetition of the to-be-remembered unit of material on its subsequent recall. In the last decade or so the study of repetition effects has shifted from multitrial list-learning tasks to more tractable single-trial tasks in which the unit of analysis has been a single item rather than a larger collection of items. But the fundamental problem has remained the same, to gain some understanding of the mechanisms and processes that are responsible for the behavioral consequences of repeating an input item. This paper describes three experiments designed to study the problem of repetition in a relatively simple single-trial list-item task.

The starting point of our analysis is the following question: What exactly is it that must be repeated for the facilitative effect on recall to occur? The obvious answer, "the to-be-remembered item," is not entirely satisfactory for several reasons. First, it is now widely accepted that the memory trace of an experimentally designated unit of material cannot be thought of as a fixed entity, but must be conceptualized as a bundle of attributes, elements, or encoding dimensions (e.g., Bower, 1967; Underwood, 1969a; Wickens, 1970). Such a view of the memory trace is usually accompanied by the notion that the properties of the trace do not only depend upon the characteristics of the presented item, but also on the particular encoding operations performed on the item at input (e.g., Bower, 1972; Martin, 1968; Tulving & Thomson, 1973). Since the presentation of an item may thus have variable internal consequences, the

[1] This research has been supported by the National Science Foundation Grant No. 24171X. The paper was written while E. Tulving was a Fellow at the Center for Advanced Study in the Behavioral Sciences in Stanford, California.

[2] Requests for reprints should be sent to E. Tulving, Department of Psychology, Yale University, New Haven, Connecticut 06510.

Copyright © 1973 by Academic Press, Inc.
All rights of reproduction in any form reserved.
Printed in Great Britain

statement to the effect that an "item" was repeated is not particularly useful in the absence of knowledge about encoding operations performed on its different presentations. Moreover, to the extent that encoding of two nominally *different* input items is based on overlapping attributes, the presentation of one of the items may produce a repetition effect on the recall of the other. For example, different words belonging to the same conceptual category, when presented in one and the same list, may facilitate each other's recall very much as repetition of the same word enhances its recall (e.g., Tulving & Patterson, 1968). Thus, the claim that an item must be repeated for the repetition effect to occur implies too narrow a view of the underlying mechanism.

Another reason for the inadequacy of the specification of repetition effects in terms of repeated presentations of the "item" lies in the fact that the typical recall score, frequency or probability of recall, often can be expressed in terms of two independent component measures. When the list of words to be remembered contains subsets of conceptually related words, and the list is presented under conditions where subjects encode the words in terms of their conceptual relatedness, then the number of words recalled from any given list by any given subject can be expressed as a product of the number of categories represented in recall and average number of words recalled from each recalled category (e.g., Tulving & Pearlstone, 1966). Since several variables can be demonstrated to have an effect on one but not on the other component of recall (e.g., Tulving & Patterson, 1968; Tulving & Pearlstone, 1966) the two components must reflect different underlying processes, only some of which may be affected by the repetition of an "item." Tulving and Pearlstone (1966) interpreted the two components as measures of accessibility of higher-order units in the memory store and accessibility of elements within higher-order units. If a list item is stored in memory as a member of a higher-order unit, then its repetition does not involve only the repetition of the "item" but also its corresponding higher-order unit, and the repetition effect may be attributable to one or the other, or both, types of repetition. Thus, to claim that it is the repetition of an "item" that is responsible for the repetition effect overlooks the possibility that the repetition of the item's higher-order unit may play an important role as well.

From these considerations it follows that repetition effects may occur in lists in which no items are repeated. Such repetition effects in absence of repeated items cannot be studied in experimental situations in which the experimenter does not know which words share encoded attributes for which subjects. But repetition effects in absence of item repetition can be explored in tasks where conceptually related words are used and subjects are encouraged to encode the words as members of conceptual categories.

In the experiments to be described, we studied the effects of three types of repetition on recall of words presented on a single trial. The basic unit of experimental design was always a pair of familiar words belonging to one and the same conceptual category. The two elements (words) of the basic unit (category) were always presented adjacent in the list and were preceded by their category name. Thus, for instance, the subject would be presented with a sequence, *precious stones*: RUBY, OPAL, as part of an input list, under instructions to remember the two exemplars (Experiments 1 and 2) or both the category name and the two instances (Experiment 3). The experimental question of interest in all three experiments concerned the recall of the basic unit given that (*a*) the unit or any part thereof was not nominally repeated anywhere in the list, and (*b*) given that the unit and some of its components did occur again later in the list, or had occurred earlier in the list. Two different measures of recall were obtained: probability of recall of a unit, defined in terms of recall of at least one exemplar from the

category, and probability of recall of words within recalled categories, as described by Tulving and Pearlstone (1966).

Three basic types of repetition were studied in the experiments to be reported: (*a*) repetition of the category name and the two presented exemplars, (e.g., *precious stones*: RUBY, OPAL ... *precious stones*: RUBY, OPAL), (*b*) repetition of the category name together with a new pair of exemplars (e.g., *type of fruit*: GRAPE, BANANA ... *type of fruit*: ORANGE, PLUM), and (*c*) repetition of the category name alone (e.g., *some furniture*: SOFA, DRESSER ... *some furniture*).

Another variable that was manipulated experimentally, although not for all types of repetition, was the spacing of repetition. The repetition of the basic unit occurred either immediately after the presentation of the basic unit (massed repetition) or with inputs from other categories intervening (spaced repetition). Several previous studies have found that related items are recalled at a higher level when their presentation is blocked rather than distributed throughout the list (e.g., Cofer, Bruce, & Reicher, 1966; Glanzer, 1969; Tulving & Patterson, 1968; Weist & Powell, 1972). These data seem to suggest that repetition by conceptually related items does not obey the same laws as does repetition by lexically identical items for which recall is higher under spaced than massed conditions (e.g., Madigan, 1969; Melton, 1970; Underwood, 1969b).

It is possible that the discrepancy in spacing effects for repeated and related words is a consequence of differences in encoding. If a word occurs in isolation from its related member in the list, the subject may encode it in relation to a word in the list other than the "related" word the experimenter has in mind and uses in scoring, with the result that the "repetition" effects considered by the experimenter do not reflect those that in fact have taken place in the system. This interpretation gains some support from the data and analyses recently reported by Borges and Mandler (1972), Weist and Powell (1972), and Schwartz (1973). In the experiments to be reported in this paper, encoding of the conceptually related items by subjects was under reasonably tight experimental control. As the data reported below show, spacing of related words under these conditions resulted in higher recall than massed presentation.

In summary, the purpose of the experiments was to investigate the effects of intralist repetition as a function of type and spacing of repetition. We hoped that such an investigation and its outcome might serve some useful purpose in relating an old problem of memory, the effects of repetition, to what is a relatively new conception of the repeated "item" as a complex multidimensional bundle of information in the memory store. The adoption of a pair of conceptually related words as the basic unit of design and the consequent separability of recall of higher-order units and elements within the units provided an experimental means for the analysis of the effects of intralist repetition in a form not possible in conventional paradigms.

Experiment 1

Only one kind of repetition was studied in Experiment 1: repetition of the category with different exemplars. The primary interest was focused on the effects of such repetition by related words under two conditions, massed and spaced repetition.

Method

Design. Each subject was tested with three lists, each list representing one replication of the experiment. In each list, basic two-word units were presented for study under three conditions: (*a*) massed repetition by related exemplars, (*b*) distributed repetition by related exemplars, and (*c*) no repetition.

Subjects were shown each list under instructions to remember as many exemplars of categories as they could, and at the end of the presentation of the list were given a free-recall test, followed by a cued recall test in which category names of list items served as retrieval cues.

The following shorthand notation will be used to facilitate description of the experiment and its results. We refer to the basic unit as the C(A) unit, and to the

corresponding experimental condition as the C(A) condition. The C represents a category name, and the (A) stands for a pair of exemplars, for example, *earth formations*: CLIFF, RIVER. The unit repeated under massed conditions is referred to as a C(AB) unit. Here C and (A) have the same meaning as before, and (B) represents the immediately following second pair of exemplars, for example, *type of fruit*: ORANGE, PLUM, GRAPE, BANANA. The category name was not repeated between pairs of examplars in the case of massed repetition. The spaced repetition condition is represented as C(A)–C(B). In this case the basic C(A) unit appeared in the first half of the list, and two related instances, together with the same category name, C(B), appeared in the second half of the list, for example, *reading material*: TEXT, PAMPHLET ... (other words) ... *reading material*: BOOK, MAGAZINE. The lag between distributed units varied nonsystematically within and between lists, ranging from 14 to 22 exemplars of other categories.

Lists. A total of six lists were constructed, representing three unique sets of words, each having two forms in which the units in the first and second half of the lists were transposed.

Each list contained 26 categories and 48 exemplars. Sixteen category names or descriptive phrases, each accompanied by a *single* to-be-remembered exemplar, served as buffer units. Four buffer units were placed at the beginning of each list, four in the middle (separating the first half of the critical units from the second half), and eight buffer units were at the end of each list. The remaining 10 categories and 32 exemplars in each list constituted the critical units providing experimental data of interest. They included four C(A) units, two in each half of a list; four C(AB) units, again two in each half of the list; and two C(A)–C(B) units, the C(A) part appearing in the first half of the list, and the C(B) part in the second half.

The categories and the exemplars comprising the critical units were selected from the Battig and Montague (1969) norms. In each case the exemplars chosen from the norms were the second, fourth, sixth, and eighth most frequently given response to the category in the norms. When only two exemplars were needed for a particular unit, they were selected randomly from among the four. Since there were 10 categories of critical units in each list, a total of 30 categories was selected for critical units from the norms. The remaining categories in the norms, together with some others made up for the purpose, were used for the buffer units. The order of occurrence of different types of critical units within each half of the list was determined randomly.

Procedure. Some subjects were tested individually, while most were tested in small groups of from two to five persons. They were told that they would hear words belonging to certain explicitly labeled categories, and that their task was to remember as many of the exemplars from the categories as they could. They were also told that they need not remember the category names, but that these names would be given as cues for recall of the exemplars on a second test that would follow the first free-recall test.

The list was presented auditorily, by means of a tape recorder. Each list was presented just once, with each category name or exemplar occurring every 2 sec.

Immediately after the presentation of a list, subjects were allowed 3.5 min for written free recall of the category exemplars. As soon as the free-recall sheets from this test had been collected, the cued test was begun. Subjects were given a sheet listing all the category names, and again allowed 3.5 min for written recall of the exemplars.

After the cued-recall test of the first list was completed, the subjects were alerted for the presentation of the next list, and the entire procedure described above was repeated until all three lists had been presented.

Subjects. Forty Yale undergraduates of both sexes participated in the experiment for either course credit or a monetary reward of $2.00. Subjects were arbitrarily divided into four groups for the sole purpose of counterbalancing of list halves and test sequences. Otherwise, all subjects contributed data to all conditions equally.

Results

Background data. There was a total of 5760 subject-words in the experiment (40 subjects × 3 lists × 48 words). The proportion of these words that were recalled was .41 on the noncued test and .72 on the cued test. Across the three lists learned by a subject there was a tendency for recall to be slightly higher on the second list relative to the first and third. The mean proportion of buffer items recalled was .36 on the noncued test and .76 on the cued test.

On the noncued test, the mean proportion of exemplars recalled from C(A) categories presented in the first half of a list was .27, and the second half, .28. The mean proportions of noncued recall for C(AB) categories were .39 for categories presented in the first half of a list and .40 in the second half. In the C(A)–C(B) categories, .69 of the (A) exemplars from the first half of the list were recalled, and .64 of the (B) exemplars from the second half. In

all subsequent analyses the data were combined across list halves for C(A) and C(AB) units which had all their input within a single half of the list.

To check on our assumption that subjects' higher-order memory units corresponded to experimentally defined categories, the percentage of times the exemplars of various types of units were recalled in adjacent output positions, given that they were recalled at all, was calculated. Out of 104 cases where both (A) exemplars of the C(A) categories were recalled, 97% of the time they were recalled adjacently. Out of 177 cases from the C(AB) categories and 154 cases from the C(A)–C(B) categories in which *at least three* exemplars were recalled, the percentage of times all recalled exemplars occupied adjacent output positions were 84% and 82%, respectively. This is taken as evidence that higher-order memory units were formed to a large extent on the basis of category membership in all three conditions of repetition.

Analysis of recall. The data of primary interest are provided by the noncued recall of higher-order units and words within recalled units, from the nominally nonrepeated C(A) condition, and from the massed, C(AB), and spaced, C(A)–C(B), repetition conditions. Since the cued recall data reflect the combined effects of the earlier presentation and of the free recall of the items, they will be mentioned only in passing.

A higher-order unit was regarded as having been recalled by a subject if the subject recalled at least one word from the corresponding category. Data from other experiments (e.g., Borges & Mandler, 1972), as well as those from Experiment 3 reported in this paper, show that this convention introduces only a negligible bias. The criterion adopted for the recall of the unit was identical for the noncued and cued tests. Recall of words within recalled units was obtained for each subject by dividing the number of words recalled in a given experimental condition by the number of units recalled in that condition.

In the analysis of the data, the observations were first reduced to means of each of the two recall components (unit recall, and words-within-units recall) within each condition for each subject. For this purpose, each subject's data were pooled over all three lists with which he had been tested. These subject-means were converted into proportions and means of these proportions, as well as standard deviations, calculated as summary statistics for each condition. This procedure ensures equal weighting of data from all subjects. Occasionally some subjects did not recall any words within a particular condition. Whenever this happened, the subjects received a score of zero for the unit recall measure, and no score for the words-within-units measure. The reduced N will be indicated in a footnote to the appropriate table in which the data are summarized.

Recall of higher-order units. Table 1 presents the means and standard deviations of the component recall measures derived for the three types of units in this experiment. Repetition effects produced by the inclusion of a second pair of exemplars belonging to the same category increased category recall, R_C, under both massed, C(AB), and spaced, C(A)–C(B), conditions, relative to the nominally nonrepeated units, C(A). As shown in Table 1, the mean proportion of C(AB) categories recalled (.53) was considerably greater than the proportion of C(A) categories recalled (.33), $t(39) = 6.22$, $p < .01$, and the proportion of C(A)–C(B) categories recalled (.84) was much higher than that of C(AB) categories, $t(39) = 9.12$, $p < .01$.

Recall of words within recalled units. Data in Table 1 reveal that the presence of the second, (B), pair of exemplars in the input list had no effect on recall of the original, (A), pair of exemplars within recalled units. Recall of the (A) words given that the category was recalled, $R_{(A)/C}$, was virtually the same for all three types of units. A comparison of $R_{(A)/C}$ to $R_{(B)/C}$ in both types of categories in which the (B) pair occurred shows that the first presented, (A), pair of exemplars was recalled more frequently

TABLE 1

RECALL PROPORTIONS OF CATEGORIES AND WORDS WITHIN RECALLED CATEGORIES IN NONCUED AND CUED TESTS AS A FUNCTION OF CATEGORY TYPE, EXPERIMENT 1

Category type	Noncued test			Cued test		
	R_C	$R_{(A)/C}$	$R_{(B)/C}$	R_C	$R_{(A)/C}$	$R_{(B)/C}$
C(A)	.33 (.15)	.82 (.14)	—	.81 (.15)	.81 (.09)	—
C(AB)	.53 (.16)	.78 (.13)	.69 (.17)	.92 (.09)	.76 (.12)	.65 (.18)
C(A)–C(B)	.84 (.14)	.82 (.15)	.77 (.17)	.98 (.05)	.84 (.16)	.72 (.20)

Note: Designation of category types is explained in the text. R_C refers to recall of categories, as defined in the text, $R_{(A)/C}$ and $R_{(B)/C}$ represent proportions of words recalled within recalled categories, the former measure being based on the first two instances presented in a category and the latter on the last two. Each proportion is based on 40 subject means. Standard deviations are shown in parentheses.

than the second pair. This difference was significant for only the massed C(AB) categories on the noncued test, $t(39) = 2.86$, $p < .01$, but was also significant for the spaced C(A)–C(B) categories on the cued test, $t(39) = 2.96$, $p < .01$. The difference in $R_{(B)/C}$ between C(AB) and C(A)–C(B) categories was not significant.

Cued recall. The results of the cued recall test, also shown in Table 1, reflect the same pattern described by the noncued test results. The main difference between the cued and noncued recall was higher category recall for all category types on the cued test, attesting to the power of the category names as cues for their exemplars. The relative stability of the recall of words within recalled units across the two tests replicates similar findings in other experiments (e.g., Tulving & Pearlstone, 1966).

Summary of results. This experiment replicated earlier findings that a "repetition" effect can be produced by the inclusion of conceptually related words in a list. Recall of an original pair of exemplars of a category was increased greatly when a second pair of exemplars from the same category also occurred in the list. A new finding was derived from the component recall analysis: The effect of repetition by related exemplars was entirely on category recall, since recall of the exemplars within recalled categories was not affected. Furthermore, the effect of repetition by related words was found to be greater under spaced than massed conditions, the difference again being attributable solely to the enhanced recall of the category component. Finally, recall of exemplars presented first in a repeated category was higher than recall of exemplars presented second.

EXPERIMENT 2

The purpose of Experiment 2 was to compare the repetition effects produced by inclusion of related words in a list with repetition effects produced by repetition of identical words. As in Experiment 1, the nominally nonrepeated unit consisted of a category name followed by to-be-remembered exemplars, C(A). Repetition by related words occurred whenever a second pair of exemplars belonging to the same category was included in the list, C(AB) or C(A)–C(B). In the contrasting condition the original pair of exemplars of a category occurred a second time in the list C(AA) or C(A)–C(A).

Method

Design. The design of this experiment was essentially the same as that of Experiment 1, with the addition of two new types of repeated units. (*a*) Massed repetition of exemplars: The original pair of exemplars was repeated immediately after its initial presentation in the same order (e.g., *some furniture*: SOFA, DRESSER, SOFA, DRESSER). Using the shorthand notation developed for Experiment 1, where C stands for the category name and (A) represents a pair of exemplars, these types of units will be referred to as C(AA), units. (*b*) Spaced repetition of exemplars: These units designated C(A)–C(A) units, had both the category name and the original pair of exemplars repeated later in the list (e.g., *article of clothing*: SOCK, SKIRT ... (other words) ... *article of clothing*: SOCK, SKIRT). The nominally nonrepeated unit, C(A), and the massed and spaced units for which related words were repeated, C(AB) and C(A)–C(B), as in Experiment 1, were also included in this experiment. All five types of units occurred in each of three lists presented to each subject.

Lists. A total of six lists were constructed, as in Experiment 1. Each list contained 20 categories and 34 to-be-remembered exemplars. Twelve category names, each accompanied by a single to-be-remembered exemplar, served as buffer units: four at the beginning, two in the middle, and six units at the end of each list.

The remaining eight categories and 22 different exemplars (28 tokens) constituted the critical units of each list. These consisted of two each of the nominally nonrepeated units, C(A); two types of massed repetition, C(AA) and C(AB), with one unit of each type occurring in each half of the list; and one each of the spaced repeated units, C(A)–C(A) and C(A)–C(B), with the (A) pair of exemplars presented in the first half, and either the same pair or the (B) pair presented in the second half of the list.

The materials were essentially the same as those used in Experiment 1 with appropriate modifications to accommodate the types of units utilized in this experiment.

Procedure. The procedure was very similar to that of Experiment 1. Each list was presented once by means of a tape recorder, at the rate of 2.4 sec per category name or exemplar. A noncued free recall test was given immediately after each list, which was followed by a cued test. Subjects were informed about the categorized nature of the lists and that some words would be repeated in a list. In Experiment 2 they were also instructed to write twice on the recall sheet any words they remembered as having been repeated in the list.

Subjects. Forty subjects of both sexes served in a 1-hr experimental session for $2.00. They were solicited by notices posted in the vicinity of Yale University, Southern Connecticut State College, and the New Haven branch of the Connecticut Army National Guard.

Results

Background data. The subjects in Experiment 2 recalled 42% of the total of 3360 subject-words on the noncued test and 64% on the cued test. As in the first experiment, subjects tended to do best on the second list. The mean proportion of buffer units recalled was .33 on the noncued test and .71 on the cued test.

Mean proportions of noncued recall from units presented in the first half and in the second half of a list were, respectively: C(A) units, .23 and .21; C(AA) units, .58 and .37;[3] and C(AB) units, .45 and .41. In the C(A)–C(B) units, .53 of the (A) pairs of exemplars presented in the first half of the list, and .50 of the (B) exemplars from the second half of the list were recalled. In subsequent analyses the data were combined across list halves for the C(A), C(AA), and C(AB) units.

Recall of higher-order units. Component recall measures were calculated for each type of unit in the same way as described in Experiment 1. The means and standard deviations of these measures are presented in Table 2.

Units repeated by massed related exemplars, C(AB), were better recalled (.63) than nominally nonrepeated units, C(A) categories (.27), replicating the findings of Experiment 1. Units repeated by massed identical exemplars, C(AA), were also recalled better (.56) than C(A) categories, $t(39) = 7.17$, $p < .01$. The two types of massed repeated units, C(AB) and C(AA), did not differ significantly in category recall in either the noncued or the cued test.

As in Experiment 1, repetition by related exemplars resulted in higher category recall under spaced, C(A)–C(B), than massed,

[3] The large discrepancy between the first and second half of the list may reflect the "isolation" effects accruing to the first C(AA) unit that appears in the list, since the design of the experiment was such that item-specific sampling effects were ruled out.

TABLE 2

RECALL PROPORTIONS OF CATEGORIES AND WORDS WITHIN RECALLED CATEGORIES IN NONCUED AND CUED TESTS AS A FUNCTION OF CATEGORY TYPE, EXPERIMENT 2

Category type	Noncued test			Cued test		
	R_C	$R_{(A)/C}$	$R_{(B)/C}$	R_C	$R_{(A)/C}$	$R_{(B)/C}$
C(A)	.27 (.20)	.84[a] (.21)	—	.78 (.18)	.73 (.12)	—
C(AB)	.63 (.16)	.71 (.17)	.64 (.21)	.91 (.14)	.70 (.17)	.60 (.18)
C(AA)	.56 (.19)	.83 (.16)	—	.87 (.14)	.79 (.14)	—
C(A)–C(B)	.75 (.31)	.72[b] (.20)	.67[b] (.23)	.97 (.10)	.63 (.22)	.57 (.23)
C(A)–C(A)	.87 (.19)	.93 (.11)	—	.95 (.14)	.94 (.10)	—

Note: Designation of category types is explained in the text. R_C refers to recall of categories, as defined in the text, $R_{(A)/C}$ and $R_{(B)/C}$ represent proportions of words recalled within recalled categories, the former measure being based on the first two instances presented in a category and the latter on the last two. Each proportion is based on 40 subject means except as noted below. Standard deviations are shown in parentheses.

[a] $N = 35$.

[b] $N = 37$.

C(AB), conditions, .75 vs .63; however, the difference between the two probability values was only .12 in this experiment compared to a difference of .31 between the same unit types in Experiment 1. Spacing of repeated exemplars had a much larger effect on category recall. The difference in R_C between C(AA) and C(A)–C(A) units was .31, .56 vs .87. Spaced categories with repeated exemplars, C(A)–C(A), were recalled better (.87) than spaced categories with related exemplars, C(A)–C(B), (.73), $t(36) = 2.24$, $p < .05$.

Recall of words within recalled units. Table 2 shows that a higher proportion of (A) than (B) exemplars was recalled in all types of units that had the (B) pair, replicating a similar finding from Experiment 1. The recall of words within recalled units, in the C(AB) and C(A)–C(B) conditions did not differ significantly in either the noncued or the cued tests.

Massed repetition of exemplars had no effect on recall of words within units. $R_{(A)/C}$ of C(AA) categories did not differ from that of C(A) units. However, spaced repetition of exemplars in C(A)–C(A) categories did increase $R_{(A)/C}$ above that found in the C(A) condition. This difference approached significance on the noncued test, $t(39) = 1.96$, $p < .06$, and was significant on the cued test, $t(39) = 8.19$, $p < .01$.

Token scoring. Since subjects were asked to write repeated exemplars twice on the recall tests, the total number of tokens recalled from repeated units could be measured. In token scoring, a repeated word written twice on the recall test counts as two tokens. For categories without any repeated exemplars, the number of tokens is the same as the number of words recalled and scoring is not changed. Our primary interest in token data was to see if the difference in the recall of words within recalled categories between units with repeated identical exemplars and units with repeated related

exemplars would disappear under token scoring. Such a disappearance would imply that subjects recalled the same number of tokens or events in both conditions.

For C(AA) units the token $R_{(A)/C}$ means were .76 and .63 on the noncued and cued tests, respectively. For C(A)–C(A) units token $R_{(A)/C}$ was .87 on the noncued test and .86 on the cued test. These scores may be compared to the mean of $R_{(A)/C}$ and $R_{(B)/C}$ measures for units repeated by related words. For C(A)–C(B) units the corresponding values were .70 and .60 for the noncued and cued tests. Thus, the advantage of spaced units repeated by identical words over spaced units repeated by related words remained under token scoring, but the superiority of massed units with repeated identical words over C(A)–C(B) units did vanish.

Possible selection effects in recall. While the general pattern of results in Experiment 2 was very similar to that found in Experiment 1, spacing of related words did have a considerably smaller effect on category recall in Experiment 2 than in Experiment 1. The difference between C(A)–C(B) units and C(AB) units on the R_C measure was .31 in Experiment 1 and only .12 in Experiment 2. Apparently the discrepancy in the magnitude of the spacing effect for units repeated by related exemplars has something to do with the presence of units repeated by identical exemplars in Experiment 2, and absence of such possibly more conspicuous units in lists used in Experiment 1. The units repeated by identical exemplars may have stood out from among the other types of units in Experiment 2, and such an "isolation" of these units may have replaced the similar status of spaced units repeated by related exemplars.

To study the effects of repetition by identical exemplars and by related exemplars under conditions where possible selection artifacts would be absent, Experiment 3 was designed to compare the spacing effects for the two types of repeated units in separate lists. In Experiment 3, subjects were additionally asked to recall category names as well as exemplars to reduce any possible selection effect that may have arisen, because of special instructions, in Experiment 2, to remember which *exemplars* were repeated in the list. Finally, in Experiment 3, a new type of repetition was studied in which only the category name of the basic unit was repeated without any exemplars.

Experiment 3

Method

Design. This experiment compared three types of *spaced* repetition: (*a*) repetition by identical exemplars, C(A)–C(A) units; (*b*) repetition by related exemplars, C(A)–C(B) units; and (*c*) repetition by the category name alone, C(A)–C and C–C(A) units. This latter type of repetition consisted of the presentation, either before or after the basic C(A) unit, in the other half of the list, of the category name alone, accompanied by a silent interval corresponding to the duration of the presentation of two exemplars (e.g., *type of metal*: ZINC, COPPER ... (other words) ... *type of metal*: ———, ———). The fourth experimental condition was represented by a nonrepeated C(A) unit. In this condition, two words of a different category were inserted in a list where the second input for a repeated unit would have appeared in the lists representing the three types of repetition.

Each experimental condition was represented by a separate list. In each list, there were two critical units, either repeated in one of three ways or not repeated, embedded in a context of other two-word and four-word C(A) and C(AB), units. Since repeated units were always spaced in Experiment 3, repetition of a basic unit occurring in the first half of the list always occurred in the second half. In the condition where repetition was by category name alone, the single category name occurred in the first half of the list for one of the critical units, and in the second half for the other critical unit. (The former case, category name alone presented in the first half of the list, was the only case in which a subject could have singled out the critical unit from among context units, and in which an "isolation" effect may have occurred. In all other cases the subjects could not identify the to-be-repeated unit until its actual repetition in the second half of the list.)

Each subject was tested with four lists, each representing a different experimental condition. Thus all subjects contributed data equally to all conditions. The order of presentation of lists and conditions was counterbalanced within groups of subjects.

Lists. Four unique lists were constructed, each contained 15 context and buffer categories with 30

exemplars, plus the two critical categories and their exemplars. The noncritical categories in each list were 11 two-word, C(A), categories and four four-word, C(AB), categories. Seven of the C(A) categories served as buffers, being arranged with two at the beginning, two in the middle, and three at the end of each list. The remaining four C(A) categories and four C(AB) categories constituted the intralist context material. These were arranged so that two of each type occurred in each list half.

Four different list structures were devised, defined in terms of the order of occurrence of the context categories and the locations at which the critical units were to be inserted. The mean lag for the critical units across all lists was 12 intervening exemplars from other categories.

In all there were four different sets of context and buffer categories, four list structures and four sets of critical categories. The four sets of context units and the four sets of critical units were counterbalanced within subjects by a Latin square arrangement. List structures were treated as a between-subjects variable. Thus, within a group, the four lists learned by a subject had an identical structure.

Materials for the lists were selected from those used in the two previous experiments.

Subjects. Subjects were 48 Yale undergraduates of both sexes, each serving in a 1-hr experimental session for course credit.

Procedure. Subjects were tested in small groups of from 1 to 5. Each list was played once on a tape recorder at a rate of 2.4 sec per category or exemplar, with an extra 4.8 sec of silence after a category name when it occurred alone (in the C–C(A) and C(A)–C conditions). As in the two earlier experiments, subjects were informed of the categorized nature of the lists, but unlike in the previous experiments, subjects were asked to remember category names as well as exemplars. At the time of recall they were instructed to write twice on the recall tests any words (category names or exemplars) they remembered as having been repeated in a list.

A noncued test was followed by a cued test as in Experiments 1 and 2. Each test was terminated when 3.5 min had elapsed. The next list was presented immediately after the cued test.

Results

Background data. Recall of category names was not entered into the analysis of the data except in determining category recall as explained below.

Out of the 8448 subject-words in the experiment, subjects recalled 40% on the noncued test and 70% on the cued test. As in the previous experiments, recall tended to be highest in the second list. A mean proportion of .43 of the exemplars was recalled from buffer categories on the noncued test and .76 on the cued test.

As in the previous experiments, list half had some effect on recall. For the context categories, the following proportions of exemplars were recalled from units occurring in the first and second half of the list, respectively, on the noncued test: C(A) units, .30 and .28; C(AB) units, .42 and .33. From the C(A)–C(B) critical units, .65 of the first presented (A) exemplars and .62 of the later occurring (B) exemplars were recalled. In subsequent analyses the data were pooled across list halves for all C(A) and C(AB) units.

Subjective unitization by experimental categories was again checked. Out of the 181 cases where both exemplars were recalled, 99% of the exemplars of C(A) context categories were recalled in adjacent output positions. Out of 276 cases where *at least three* members of the C(AB) units were recalled and 64 cases where at least three members of C(A)–C(B) units were recalled, the percentages of times that all recalled exemplars occupied adjacent output positions were 96% and 89%, respectively. As in the previous experiments, the present data show a high degree of clustering by category membership.

The analysis of data into recall components differed slightly from the method used in the first two experiments. Since subjects were asked to recall category names on the noncued test, category name recall was used as an index of category accessibility. This method of scoring category recall allows recall of words within recalled units to vary from zero to one, rather than being constrained as it is when category recall is inferred from recall of at least one exemplar.

Using category name recall as the criterion for category recall also makes possible the detection of cases where either exemplars are recalled but the category name is forgotten, or

where the category name is recalled with no exemplars. The number of such cases was quite low in the present experiment. There were 49 cases of exemplars recalled without the correct category name, and 16 cases where the category name was recalled with no exemplars. These numbers may be compared to 1613 cases where the category name and at least one exemplar were recalled. In the light of these and other similar data reported by Buschke and Lazar (1973) and Borges and Mandler (1972), it is clear that the measure of category recall inferred from recall of exemplars is highly correlated with the direct measure.

Context categories. The means and standard deviations of the component measures derived for both the context and critical units are presented in Table 3 for the noncued and cued tests. The recall components of the context units are quite similar to those derived for the same types of units in the earlier experiments. These data serve as a baseline against which the recall of the critical units can be compared. The recall components of the C(A) context units were quite comparable to the components of the C(A) critical units.

Recall of critical higher-order units. All three types of spaced repetition had large effects on category recall relative to nominally non-repeated C(A) units. On the noncued test R_C was .81 for C(A)–C(B) units, .85 for C(A)–C(A) units, and .91 for C–C(A) and C(A)–C units. The differences were not statistically reliable. When the two orders of category name

TABLE 3

RECALL PROPORTIONS OF CATEGORIES AND WORDS WITHIN RECALLED CATEGORIES IN NONCUED AND CUED TESTS AS A FUNCTION OF CATEGORY TYPE, EXPERIMENT 3

Category type	Noncued test			Cued test		
	R_C	$R_{(A)/C}$	$R_{(B)/C}$	R_C	$R_{(A)/C}$	$R_{(B)/C}$
			Context items			
C(A)	.37 (.15)	.81[a] (.14)	—	.84 (.14)	.79 (.11)	—
C(AB)	.52 (.12)	.78 (.13)	.67 (.16)	.91 (.09)	.75 (.12)	.67 (.12)
			Critical items			
C(A)	.32 (.24)	.82 (.24)	—	.79 (.23)	.83 (.14)	—
C–C(A) & C(A)–C	.91 (.19)	.77 (.26)	—	.95 (.18)	.80[c] (.21)	—
C(A)–C(B)	.81 (.32)	.76[b] (.30)	.77 (.25)	.98 (.10)	.76 (.26)	.72 (.28)
C(A)–C(A)	.85 (.25)	.95[c] (.12)	—	.97 (.16)	.93[c] (.12)	—

Note: Designation of category types is explained in the text. R_C refers to recall of categories, as defined in the text, $R_{(A)/C}$ and $R_{(B)/C}$ represent proportions of words recalled within recalled categories, the former measure being based on the first two instances presented in a category and the latter on the last two. Each proportion is based on 48 subject means except as noted below. Standard deviations are shown in parentheses.

[a] $N = 38$.

[b] $N = 44$.

[c] $N = 47$.

repetition were scored separately (each based on only 48 observations) the R_C scores were .96 for C–C(A) and .85 for C(A)–C on the noncued test, and .96 and .94 on the cued test.

Recall of words within recalled units. Of the three types of repetition, only spaced repetition by identical exemplars, C(A)–C(A), produced a higher level of recall of words within recalled categories, in comparison with the nonrepeated C(A) condition. On the noncued test, the mean $R_{(A)/C}$ score was .81 for the critical C(A) units and .95 for the C(A)–C(A) units, $t(37) = 3.30$, $p < .01$. The $R_{(A)/C}$ scores for these two types of units also differed on the cued test, $t(46) = 3.93$, $p < .01$.

When scored separately, the two orders of category name repetition had virtually identical $R_{(A)/C}$ means: .77 for C–C(A) and .79 for C(A)–C units on the noncued test, and .78 and .81 on the cued test. Since rehearsal of the not yet presented exemplars would be impossible in the C–C(A) units this equality contraindicates implicit retrieval of the earlier presented exemplars when the category name was repeated in C(A)–C units.

Token analysis. A token analysis was performed similar to the one in Experiment 2. The token $R_{(A)/C}$ scores for C(A)–C(A) units were .91 and .75 for the noncued and cued tests. These scores can be compared to .77 and .74, the corresponding averages for C(A)–C(B) units on the noncued and cued tests. Thus, more tokens were remembered for units with repeated identical than related exemplars on the noncued test, but there was no difference on the cued test.

General Discussion

The purpose of these experiments was to study the effects of three types of repetition in a single-trial memory task in which a basic unit of analysis was a pair of conceptually related words and where subjects were induced to encode the list words as members of identifiable higher-order units. The adoption of these units of analysis made it possible to consider the effect of repetition separately on the access to higher-order units as well as to individual elements within accessible units.

The general picture that emerged from the analysis of the data from the three experiments can now be summarized. All three types of repetition—category name and same exemplars, category name and different exemplars, and category name alone—were quite effective in increasing recall of words in the basic unit. Since the only thing in common to the three types of repetition was the repetition of the category name, the findings suggest that some kind of repeated activation of the higher-order unit, quite independently of repeated activation of its elements, is sufficient to produce repetition effects. Greater effectiveness of spaced than massed repetition was found for units repeated by identical words as well as those repeated by related words. All repetition effects, with a single exception (spaced repetition of identical words), were clearly attributable only to increased accessibility of higher-order units, since, with the sole exception noted, repetition had no facilitating effect on recall of individual elements within accessible categories. The superiority of spaced repetition over massed repetition, again with the single exception noted, was also clearly limited to the category-recall component of the overall measure of recall. These data thus suggest that repetition in general, and the superiority of spaced over massed repetition in particular, facilitates recall very much in the same way as does presentation of specific retrieval cues: It affects accessibility of higher-order units, but not accessibility of elements within higher-order units (Tulving & Pearlstone, 1966).

The fact that the repetition of the basic unit as a whole leads to enhanced recall of members of the unit is, of course, not at all surprising, nor is the finding of mutual facilitation of recall when additional related words are added to a unit. The former finding replicates countless similar results obtained under many different

conditions, and the latter one also represents a well known phenomenon (e.g., Deese, 1960; Schwartz, 1973; Tulving & Patterson, 1968). A finding that has not been previously reported in the literature concerns the third type of repetition investigated in Experiment 3: Repetition of only the category name produced enhanced recall of the words in the corresponding unit. This category name repetition was effective regardless of whether presentation of the category name alone came before or after the corresponding basic input unit. Had it occurred only under conditions where the basic unit had been presented first, and the category name alone later, it would have been possible to interpret the increment in recall by invoking the notion of "rehearsal": Subjects could have rehearsed the words they had seen earlier in the list when the category name occurred alone, particularly since the category name was followed by an empty interval in the list. The fact that a similar repetition effect was obtained when the category name occurred alone before any exemplars had been presented rules out the rehearsal interpretation. A more appropriate interpretation of the effect of repetition by category name is suggested by the finding that all three types of repetition were effective, and the fact that the only thing the three experimental operations had in common was the repetition of the category name. Thus, it looks as if repetition of the higher-order unit to which an item belongs is a sufficient condition for enhanced retrievability of the item regardless of whether the item itself is repeated or not.

The second major observation concerned the greater effectiveness of spaced than massed repetition. In the condition where the identical basic unit was repeated the finding replicates many similar results reported by other investigators in other kinds of experiments (e.g., Madigan, 1969; Melton, 1970; Underwood, 1969b). In the case of repetition by conceptually related items, the superiority of spaced over massed repetitions has not been demonstrated in previous studies. These studies, as we mentioned in the introduction, have shown massed presentation of related words to be more advantageous to their subsequent recall than distributed or spaced presentation of these words (Cofer *et al.*, 1966; Glanzer, 1969; Schwartz, 1973; Tulving & Patterson, 1968). We think that it was the specific category-related encoding of the to-be-remembered words induced through contiguous presentation of at least two words from a category, together with explicit labeling of the category, that was responsible for the large facilitating spacing effects on related words observed in the present experiments. As far as we know there has been only one other experiment showing a facilitative effect of spacing related words, Borges and Mandler (1972), and in that study, too, the category name was presented along with each exemplar of a category, thus ensuring category-related encoding under both massed and spaced conditions.

The third major finding was that the effects of all three types of repetition, as well as the effects of spacing within two of these types, were largely limited to one of the two components of recall, category accessibility. Only spaced repetition by identical exemplars enhanced accessibility of items within recalled categories. None of the other types of repetition examined in the three experiments, with data from seven independent comparisons, showed a similar facilitation of words-within-categories recall.

This pattern of results suggests that access to a higher-order unit, reflected in category recall, is obtained through the control element of the unit, which is different from individual items constituting the unit. Repetition by related items increased the size of the presented category, repetition of the whole unit resulted in more accessible individual constituents of the unit, as reflected by the increase in the words-within-categories measure, and yet repetition by only the category name, which neither increased unit size nor accessibility of elements of accessible units, resulted in as

large an enhancement of category recall as did the first two types of repetition. The increase in category accessibility by category name repetition, without a concomitant increase in the recallability of the unit members, is contrary to conceptualizations of memory structure in which a higher-order unit is retrieved through its members, that is, in which one member of a unit must be retrieved before its connections to other members can be utilized (e.g., Anderson & Bower, 1973; Postman, 1971) or before the search set can be sufficiently reduced to make possible retrieval of other unit members (e.g., Shiffrin, 1970).

A type of memory model that can accommodate the above pattern of results as well as the other findings in these experiments is one which posits a hierarchical structure of the higher-order unit. In such a structure, access to individual elements of the higher-order unit takes place through the code (Johnson, 1972) or control element (Estes, 1972), and there are no direct associative connections between individual elements. Access to the code or control element precedes access to *any* lower level members of the unit and is therefore independent of the recallability of lower level elements. The absence of direct connections between individual elements within a higher-order unit is inferred from the complete lack of any facilitating effect, in any one of three experiments, of the presentation and recall of related words on recall of other words within recalled units. These conclusions about the hierarchical structure of higher-order units are in good agreement with similar conclusions based on findings from other episodic memory experiments, involving the storage and retrieval of conceptually related words (Slamecka, 1972), and experimentally integrated small groups of words (Martin, 1971).

One final observation of a somewhat different sort may be worth mentioning. A number of previous experiments (e.g., Shuell, 1968; Thompson & Poling, 1969; Winograd, 1968) have demonstrated reduction of recall of words from categories common to two or more lists, when such recall is compared with recall of words from categories that are not repeated from list to list. Indeed, common category membership of words has been thought to represent a source of retroactive inhibition. In our experiments, the two list halves contained identical and different categories, and had any retroactive inhibition occurred in the list, recall of words from the first half of the list should have been lower when the category of the words was repeated with different instances in the second half. There was, as we have seen, no trace of any intralist retroactive inhibition. On the contrary, recall of first-half list words from a category repeated by related items was higher than recall of second-half words from the same category in all three experiments, and very much higher than recall of first-half list words from nonrepeated categories. It is not clear why the presentation of words belonging to the same category produces a large amount of mutual facilitation when the category is repeated within a list, and some retroactive interference when the categories are repeated between lists. One obvious difference between the two conditions has to do with an interpolated explicit recall test, absent in a within-list repetition of categories and present in the typical experiment in which categories are repeated from list to list. It may well be that this difference in procedure is partly responsible for different outcomes. The hypothesis receives some support from observations (Tulving & Watkins, 1974) about importance of explicit recall tests in determining the amount of negative transfer in the AB, AC paradigm. The matter is not entirely clear, however, since Winograd (1968) did demonstrate retroactive inhibition under conditions where the first list was not tested before the second one was presented, a situation somewhat similar to that studied in the present experiments. For the time being, therefore, the problem of facilitation versus interference in recall of words from a repeated category remains without solution.

References

Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D.C.: Winston, 1973.

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**, No. 3, Part 2.

Borges, M. A., & Mandler, G. Effect of within-category spacing on free recall. *Journal of Experimental Psychology*, 1972, **92**, 207–214.

Bower, G. A multicomponent theory of the memory trace. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation*, Vol. 1. New York: Academic Press, 1967.

Bower, G. H. Stimulus-sampling theory of encoding variability. In A. W. Melton and E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972.

Buschke, H., & Lazar, G. Cue encoding and recognition in facilitation of recall. *Journal of Experimental Psychology*, 1973, **97**, 75–78.

Deese, J. Frequency of usage and number of words in free recall: The role of association. *Psychological Reports*, 1960, **7**, 337–344.

Estes, W. K. An associative basis for coding and organization. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972.

Glanzer, M. Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 105–111.

Johnson, N. F. Organization and the concept of a memory code. In A. W. Melton and E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972.

Madigan, S. A. Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 829–835.

Martin, E. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review*, 1968, **75**, 421–441.

Martin, E. Stimulus component independence. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 715–721.

Melton, A. W. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 596–606.

Postman, L. Organization and interference. *Psychological Review*, 1971, **78**, 290–302.

Shiffrin, R. M. Memory search. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

Shuell, T. J. Retroactive inhibition in free-recall learning of categorized lists. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 797–805.

Schwartz, A. K. An analysis of the recall of sequences of related words. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 32–42.

Slamecka, N. J. The question of associative growth in the learning of categorized material. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 324–332.

Thompson, C. P., & Poling, D. R. Free recall of successive lists: Effect of category size and category repetition. *Psychonomic Science*, 1969, **16**, 201–202.

Tulving, E., & Patterson, R. D. Functional units and retrieval processes in free recall. *Journal of Experimental Psychology*, 1968, **77**, 239–248.

Tulving, E., & Pearlstone, Z. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 381–391.

Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, **80**, 352–373.

Tulving, E., & Watkins, M. J. On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, in press.

Underwood, B. J. Attributes of memory. *Psychological Review*, 1969, **76**, 559–573. (a)

Underwood, B. J. Some correlates of item repetition in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 83–94. (b)

Weist, R. M., & Powell, A. Blocked presentation in multitrial free recall. *Journal of Experimental Psychology*, 1972, **93**, 398–403.

Wickens, D. D. Encoding categories of words: An empirical approach to meaning. *Psychological Review*, 1970, **77**, 1–15.

Winograd, E. List differentiation, recall, and category similarity. *Journal of Experimental Psychology*, 1968, **78**, 510–515.

(Received July 9, 1973)