EFFECTIVENESS OF RETRIEVAL CUES IN MEMORY FOR WORDS

ENDEL TULVING AND SHIRLEY OSLER

University of Toronto

Ss had to memorize lists of 24 to-be-remembered (TBR) words. The TBR words were exposed for study on a single input trial, in presence or absence of cue words—weak associates of the TBR words. Recall of TBR words was tested in presence or absence of these cue words. The findings showed that (a) cue words (retrieval cues) facilitated recall of TBR words when they were present both at input and output, (b) retrieval cues did not enhance recall of TBR words when they were present only at output, and (c) 2 retrieval cues presented simultaneously with each TBR word were no more effective in facilitating recall than single cues. The main conclusion was that specific retrieval cues facilitate recall if and only if the information about them and about their relation to the TBR words is stored at the same time as the information about the membership of the TBR words in a given list.

When a person studies a list of to-be-remembered (TBR) words with the intention of recalling them at a later time, appropriate mnemonic information is stored in his memory. This stored information is used at the time of attempted recall to reproduce the original input. The success of recall, broadly speaking, depends on two factors: the amount and organization of the relevant information about the TBR words in the store at the time of attempted recall (availability of information), and the nature and number of retrieval cues which provide access to the stored information (accessibility of information; Mandler, 1967; Tulving & Pearlstone, 1966). The distinction between these two factors can be demonstrated under conditions where different groups of Ss are treated identically (given identical instructions, presented with identical material, asked to engage in identical activity interpolated between input and output, etc.) up to the beginning of the recall period, and then provided with different kinds of retrieval cues. The availability of relevant information is equal for different groups under these conditions and consequently any variation in recall must be attributed to differences in accessibility of this information.

An earlier experiment in the present series (Tulving & Pearlstone, 1966) provided an experimental demonstration of the distinction between availability and accessibility. The Ss were presented with lists of TBR words which they had to memorize. At input the TBR words were accompanied by the names of conceptual categories of which the words were members. When these category names were given to Ss at output as retrieval cues, Ss recalled more words than when no experimentally manipulated retrieval cues were present at output. This finding demonstrates that retrieval de-
pends upon the completeness of reinstatement, at the time of output, of the stimulating conditions present at the time of input (Melton, 1963), but it does not provide much insight into the underlying mechanisms.

The experiment reported in this paper was designed to provide some empirical evidence relevant to speculations about the nature of these mechanisms. This evidence takes the form of answers to four specific questions. First, is it possible for cue words that are only weakly associated with the TBR words to facilitate recall of TBR words? In the Tulving and Pearlstone (1966) experiment, the average frequency of occurrence of TBR words, as responses, to category names as stimuli, according to the Connecticut Restricted Word Association norms (Cohen, Bousfield, & Whitmarsh, 1957) was 6.5%, and in every case the connection between the category name and the TBR word was quite obvious. What would happen if the strength of the associative connection is weaker and less obvious?

Second, given that a retrieval cue is effective if it is present both at input and at output, is it equally effective if it is provided to $S$ only at the time of attempted recall of the TBR word? It is conceivable that a preexperimental associative bond between the cue and the TBR word is sufficient to make the retrieval cue effective. On the other hand, it may be necessary that information about the relation of the retrieval cue to the TBR word be specifically stored at the time of input of the TBR word.

The third question is related to the second: If a TBR word is paired with a certain cue at input and its recall then tested in presence of a different but preexperimentally equivalent cue, would such a changed cue also be as effective as the original cue? To the extent that the effectiveness of a retrieval cue depends on the existence of a preexperimental associative bond between the cue and the TBR word, the changed cue would still be expected to facilitate recall. If, on the other hand, information about the relation of the cue to the TBR word has to be stored at the same time as information about the TBR word, the changed cue would not be expected to be effective.

The fourth and final question has to do with the effectiveness of double cues. Given $A$ and $B$ as two different but associatively equivalent cues for a given TBR word, would the presentation of both $A$ and $B$ at input as well as at output produce greater facilitation of retrieval of the TBR word than each of them separately? Some data from an experiment on tachistoscopic identification of words—selection of words from the long-term store—would lead one to expect that double cues are more effective than single cues (Tulving, Mandler, & Bauml, 1964). It is also conceivable, however, that even though $E$ presents two words as both input cues and retrieval cues for a given TBR word, $S$ may treat the two words as a single unit, comparable to a single-word cue, and no greater facilitation of recall would occur.

**Method**

**Design.**—Lists of 24 TBR words were presented to $S$s for study and subsequent recall on a single trial. Four input conditions were combined factorially with five output conditions, except that one input condition was associated with only four output conditions. The input conditions were: 
(a) the TBR words were presented alone (Input Cond. 0),
(b) each TBR word was accompanied by Cue A (Input Cond. A),
(c) each TBR word was accompanied by Cue B (Input Cond. B),
and (d) each TBR word was accompanied by both Cues A and B (Input Cond. AB). The output condi-
tions were: (a) noncued recall of TBR words (Output Cond. 0), (b) recall of TBR words in presence of Cues A (Output Cond. A), (c) recall of TBR words in presence of Cues B (Output Cond. B), (d) recall of TBR words in presence of both Cues A and B (Output Cond. AB), and (e) free recall of TBR words and of all cues shown at the time of input (Output Cond. WC). The Output Cond. WC was not used in conjunction with Input Cond. 0, since in the latter condition no cues had been presented. Thus there was a total of 19 experimental treatment combinations. Each combination can be designated in terms of its input and output conditions. Thus, for instance, Cond. 0-0 was a standard free recall condition—TBR words shown alone and tested in absence of any cues; in Cond. B-AB each TBR word was presented in presence of Cue B and tested for recall in presence of both Cues A and B; in Cond. AB-WC each TBR word was presented in presence of both Cues A and B and had to recall TBR words and both Cues A and B in absence of any experimentally provided aids to retrieval, etc.

Independent groups of Ss were used in each of the 19 treatment combinations. The design is reflected in the organization of the data in Table 1.

Materials.—Two equivalent lists of TBR words (I and II) were used to provide for sampling of materials. Each list consisted of 24 words selected from among the stimulus words in the Minnesota Free Association Norms (Russell & Jenkins, 1954). For each TBR word two other words were selected to serve as cues. These were low-frequency responses from the associative hierarchy of each TBR word. Each cue had been given as a response to the stimulus words in the Minnesota norms by fewer than 1% of Ss. Of the two cues, thus selected for each TBR word, one was arbitrarily labelled as Cue A and the other as Cue B. The two sets of cues, A and B, for each list were thus approximately equally related to the TBR words. Some examples of A and B cues and TBR words are: fat, LEG—MUTTON; village, DIRTY—CITY; dark, GIRL—SHORT; body, VIGOR—HEALTH; empty, HURT—STOMACH; emblem, SOAR—EAGLE.

Independently of the present experiment, free association data were collected for cue words of this experiment as stimuli. The Ss were 278 high school students in the same school system in Metropolitan Toronto where the present experiment was carried out. One-third of these Ss were presented with 48 A cues, another third with 48 B cues, and the remaining one-third with each pair of corresponding A and B cues presented simultaneously (double cues). The Ss were instructed to write down for each single or double stimulus word some other word that the stimulus word or words made them think of. The observed proportions of responses corresponding to the TBR words used in this experiment (the total number of TBR words given to their respective cues, divided by the product of number of words and number of Ss) were as follows: A cues—.011, B cues—.015, and double cues—.025.

The Ss assignment to experimental conditions.—The Ss were 674 boys and girls from 23 eighth grade classes in 10 different elementary schools in the Metropolitan Toronto area. (Original design called for testing of Ss in 24 classes, but Ss were lost in 1 class because of apparatus failure.) They were tested in intact groups in their own classrooms, classes varying in size from 22 to 36. Each of the four input conditions was administered to 6 classes, each in a different school. Three classes in each input condition were tested with List I, and 3 with List II. (Because of the apparatus failure, Input Cond. 0 was given to only 2 classes with List I.) In each class, Ss were assigned to the five output conditions (four in the case of Input Cond. 0) on the basis of a random distribution of five (four) kinds of recall booklets to the seats in the classroom. Thus, each of the 19 experimental groups, corresponding to the 19 treatment combinations, consisted of Ss from 6 different classes, each in a different school. The sizes of the 19 experimental groups varied 32–39, with a median size of 36.

Procedure.—Upon entering each classroom recall booklets were first distributed. These booklets contained pages for Ss' recording of their recall, different sections of booklets being of different colors and separated by blank pages. The E then gave some general information to Ss about the study of memory conducted by the psychology department of the University of Toronto and about the nature of the task, explaining how the material to be "learned" was to be presented, how Ss were to write their recall at the proper time in their recall booklets, and how they were to work independently. A prac-
tice list, consisting of 24 adjectives as used by Tulving and Pearlstone (1966) was then administered. Words were projected on the screen by means of a film-strip projector, at the rate of 2 sec/word, and Ss were allowed 3 min. to record their recall in the booklet. The Ss were then asked to turn to the next (blank) page in the recall booklet and were given study instructions with respect to the "second part of the experiment," memorization of the experimental list proper.

Study instructions varied according to input conditions. The Ss were told that they had to study and try to remember words projected in capital letters on the screen. In Input Cond. A, B, and AB, they were told that each of the capitalized words was to be accompanied by another word which was "somehow related to the capitalized word..." (two other words in Cond. AB) typed in smaller case letters above the capitalized word. They were told that although their memory was going to be tested only for words typed in capital letters (TBR words) they should also pay attention to the words typed in lower case letters (cue words), because these "may help you to remember the capitalized words later on." They were also told to try "to see how each word and its accompanying cue (cues) are related."

The list was projected on the screen in front of the classroom, at the rate of 5 sec/frame in all input conditions. Each frame contained one TBR word in capital letters, and, in Input Cond. A, B, and AB, one or two cue words in lower case letters above the TBR word. After all 24 frames had been shown, Ss were asked to turn to the next page in their recall booklet, read the recall instructions on the top of the page, and then to record their recall. Recall instructions in the booklets varied according to the output condition to which S had been assigned. The first sentence in all instructions was, "Now write down all the capitalized WORDS you remember," and the final sentence was, "The important thing is to get as many WORDS correct as possible." Some examples of the rest of the instructions follow:

Condition 0-0: "write them down in any order you like..."
Condition 0-A and 0-B: "the words you see written on this and on the following sheet may help you to remember the WORDS since each of them is related to one of the capitalized WORDS. If you can, put each of the WORDS you remember opposite the word to which it is related. If you find this too difficult, however, put down the WORDS you remember anyhow, anywhere on this sheet or on the following sheet."

Conditions A-A and B-B: "the cue words you saw are given on this sheet and on the following sheet. If you can, write down each WORD on the line opposite the cue word. If you remember a WORD but you do not remember which cue word it went with, put it down anyhow, anywhere on this sheet or the following sheet..."

Conditions A-B and B-A: "each of the words you see on this sheet and on the following sheet is related to one of the capitalized WORDS in more or less the same way as the cue words which appeared on the slide together with the WORDS you had to remember. If you can, put each of the WORDS you remember opposite the word to which it is related. If you find this too difficult, however, put down the WORDS you remember anyhow, anywhere on this sheet or the following sheet..."

Conditions A-WC and B-WC: "write them down in any order you like in the blank spaces on the right hand side of this sheet and the following sheet. If you can, you should also write down as many of the cue words as you remember. Write the cue words in the blank spaces on the left side of this sheet and the following sheet. If you can, write each capitalized WORD on the line opposite the cue word that went with it. But if you do not remember the connection between a WORD and its cue word, write both of them anyhow, anywhere on this sheet or on the following sheet..."

The 24 words in each of two lists were shown in the same constant order to all Ss. In Output Cond. 0, the recall sheets in the booklet consisted of 24 consecutively numbered lines on two successive pages. In Output Cond. WC, there were two columns of 24 numbered lines, the left one for cue (cues) and the right one for TBR words. In Output Cond. A, B, and AB, the cues were shown in the same order in which their corresponding words had appeared in the input list. In each group, 6 min. were allowed for recall. Recall booklets were then collected and Ss permitted to ask questions they had
about the experiment, problems of memory, or psychology. Finally, Ss were asked not to discuss the experiment with other students at the school until the next day and reasons for this request were given.

Results

The mean number of words correctly recalled from the practice list of 24 adjectives was 6.19 for all 674 Ss. The same list administered under comparable conditions to 948 high school Ss in the Tulving and Pearlstone (1966) study had yielded a mean recall score of 9.48. This difference indicates that the authors are here dealing with a different population of Ss and that therefore no direct comparisons of the main data between the two experiments would be very useful. The mean scores on the practice list for different experimental groups varied from 5.74 to 6.69.

The primary data are provided by the mean number of TBR words correctly recalled. "Lenient" scoring was used throughout. The S was given credit for recall of a TBR word regardless of whether or not the word was paired with its appropriate cue.

The data on correct recall of TBR words were pooled over Lists I and II and are summarized in Table 1. Table 1 shows the number of Ss in each experimental group (n), the mean number of TBR words recalled (M), and the standard deviation of the distribution of the correct recall scores (SD). The last column in Table 1, Output Cond. WC, also shows the mean number of cue words correctly recalled. Thus, for instance, the entry 8.45 + 5.88 in the cell corresponding to Cond. A-WC means that the mean number of TBR words recalled was 8.45 and the mean number of cues recalled was 5.88. The figure for the standard deviation, however, applies only to the distribution of TBR word scores.

Table 1 contains all the data needed to answer the four questions posed in the introduction. First, is it possible for cue words that are only weakly associated with the TBR words to facilitate recall of TBR words? The answer to the question is affirmative. When single cues were present at input, cued recall (Cond. A-A and B-B) was approximately 70% higher than noncued recall (Cond. A-0 and B-0). With the data pooled over Input Cond. A and B, the two means are 14.93 and 8.73, respectively.

Second, does an otherwise effective retrieval cue still facilitate recall if it is provided to S only at the time of output? This question receives a negative answer. When single or double cues were present at output, but not at input (Cond. 0-A, 0-B, and 0-AB), recall was lower than in the absence of cues at both input and output (Cond. 0-0). With the data pooled over relevant conditions, the two means are 8.49 and 10.62, respectively.

Third, would a changed cue, pre-

<table>
<thead>
<tr>
<th>Input Condition</th>
<th>Output Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>n</td>
<td>37</td>
</tr>
<tr>
<td>M</td>
<td>10.62</td>
</tr>
<tr>
<td>SD</td>
<td>(2.79)</td>
</tr>
<tr>
<td>n</td>
<td>37</td>
</tr>
<tr>
<td>M</td>
<td>9.00</td>
</tr>
<tr>
<td>SD</td>
<td>(2.96)</td>
</tr>
<tr>
<td>n</td>
<td>36</td>
</tr>
<tr>
<td>M</td>
<td>8.44</td>
</tr>
<tr>
<td>SD</td>
<td>(2.94)</td>
</tr>
<tr>
<td>n</td>
<td>32</td>
</tr>
<tr>
<td>M</td>
<td>9.06</td>
</tr>
<tr>
<td>SD</td>
<td>(4.08)</td>
</tr>
</tbody>
</table>

Note.—For explanation of entries in column WC, see the text.
experimentally equivalent to the cue paired with the TBR word at input, be as effective as the original cue? This question is also answered negatively. Recall of TBR words in Cond. A-B and B-A was not only considerably lower than in Cond. A-A and B-B (means of 7.45 and 14.93, respectively), but also lower than in Cond. A-0, A-WC, B-0, and B-WC (pooled mean of 8.70) in which the cue present at input was simply omitted at output.

Fourth, does the presentation of two cues at input as well as at output facilitate recall of TBR words to a greater extent than do the single cues? Again the answer is negative. The mean number of TBR words recalled in Cond. AB-AB was 14.33, while in Cond. AB-0 and AB-WC it was 8.68. Thus, with input conditions held constant, recall in presence of two retrieval cues per TBR word was approximately 65% higher than in absence of these cues. The facilitation of recall by double cues thus is of the same order of magnitude as that observed in the case of single-cue input conditions (70%).

Statistical analyses, in the form of one-way analyses of variance within each input condition and subsequent contrasts of individual means using Scheffé’s method, were found to support all of the above statements at least at the .05 level of significance and will not be reported in detail. Some additional observations, however, may be of interest and will be mentioned briefly.

In Cond. A-A and B-B, e.g., over 96% of all TBR words correctly recalled were paired with their cues on recall sheets, while in Cond. 0-A and 0-B only 39% of recalled words, and in Cond. A-B and B-A 34% were so paired. The two latter figures are probably inflated because of the correspondence between input positions of words and the ordering of cues on the recall sheets. These data again demonstrate that otherwise potent retrieval cues are quite ineffective in facilitating recall if they are not presented together with the TBR words at input.

When two cues were presented with each TBR word at input, but recall was tested in presence of only one of those cues (Cond. AB-A and AB-B), recall was lower than in presence of both cues (Cond. AB-AB). This finding is reminiscent of stimulus selection in paired associate learning (Underwood, 1963), although alternative interpretations are possible.

When TBR words were accompanied by cues at input but had to be recalled in absence of any cues, recall was approximately equal in Output Cond. 0 and WC. The mean recall of TBR words in Cond. A-0, B-0, and AB-0 was 8.83, and in Cond. A-WC, B-WC, and AB-WC it was 8.55. Thus, the requirement that S retrieve cue words in addition to TBR words produced the same level of recall as the requirement that S select for recall only the TBR words from the total input. The lower recall of cues than of TBR words in Output Cond. WC, however, does suggest that Ss paid less attention to cues than to TBR words at input, i.e., that they believed E’s instructions. Of all the cues recalled, 93% were paired with their corresponding TBR words.

Finally, the number of repetitions of TBR words was negligible (674 Ss gave a grand total of 106 repetitions), but the frequency of extralist intrusions was relatively high. In the seven uncued recall conditions the mean number of extralist intrusions per S was .58, while in the 12 cued recall conditions this average was 1.77.
DISCUSSION

An earlier experiment in the present series (Tulving & Pearlstone, 1966) demonstrated that category names of TBR words can serve as effective retrieval cues. In that experiment category names accompanied the TBR words at input, and recall of TBR words was tested either in presence of category names (cued recall) or in absence of these cues (noncued recall). Cued recall was higher—in some cases considerably higher—than noncued recall. A part of the design of the present experiment replicated the paradigm used by Tulving and Pearlstone, namely Cond. A-A and B-B (cued recall) and A-0 and B-0 (noncued recall). The data from this part of the experiment fully corroborated the earlier findings.

Some other evidence available in the literature (Earhard, 1967a, 1967b; Tulving, 1962) indicates that initial letters of TBR words can also function as potent retrieval cues. In these experiments, the multistratified free-recall paradigm was used, and Ss were instructed to think about the initial letters of TBR words at input and to generate the letters of the alphabet as retrieval cues on their own at output. Such “alphabetic recall” was found to be higher than free recall. In addition, several smaller experiments the authors have done at Toronto have shown that synonomic cues of TBR words (bent—twisted, bridge—bond, soft—pliable, etc.), as well as descriptions of graphemic features of TBR words (a long word—understanding, a word ending in ly—intimately, a word with a double consonant in the middle—summer, etc.), also facilitate recall if the cues are present both at input and at output. It thus appears that a wide variety of experimentally manipulable specific retrieval cues that are meaningfully related to the TBR words can provide access to stored information about the TBR words that is available but not accessible under the noncued recall conditions.

While the meaningfulness of the connection between the cue and the TBR word—the meaningfulness obviously being determined by Ss’ preexperimental knowledge of the language—may be a necessary condition for the effectiveness of retrieval cues, it does not seem to be a sufficient condition. It may be necessary in that a random pairing of cues and TBR words will probably not enhance cued recall when compared with noncued recall, and it is not sufficient in that the presence of cues only at output (Cond. 0-A, 0-B, and 0-AB of the present experiment) does not facilitate recall of TBR words. The overall pattern of the data reported in this paper was completely consistent in showing that whenever the cues accompanied the TBR words at input, their presence at output facilitated recall, and whenever they were absent at input, their presence at output did not serve any useful purpose. In fact, the presence of cues only at output, or changing of cues from input to output, appeared to interfere with recall of the TBR words. This phenomenon may merit further study, but for the present purposes the important finding is the lack of recall facilitation by cues presented to Ss for the first time at the time of recall. This finding, in conjunction with the finding that the same cues were quite effective when presented at both input and output, suggests that specific retrieval cues facilitate recall if and only if the information about them and about their relation to the TBR words is stored at the same time as the information about the membership of the TBR words in a given list. The authors would like to offer this suggestion as the main conclusion of the present experiment.

At first blush, this conclusion may appear to be inconsistent with the results of experiments (e.g., Bilodeau & Blick, 1965; Fox, Blick, & Bilodeau, 1964; Lloyd, 1964) in which retrieval cues have been provided to Ss only at the time of recall and which have showed such retrieval cues to facilitate recall. The inconsistency disappears, however, if it is remembered that appropriate coding of input words may take place even if E does not explicitly suggest to S how he
is to code the TBR items, that is, what additional information he has to store with the TBR item at the time of input. If the TBR word is bulb—to use an example given by Bilodeau and Blick (1965)—at least some Ss are quite likely to think of it as something to do with light. If “light” is then presented by E as a retrieval cue, it is effective for those Ss in the same way as it would have been if it had been presented together with bulb at input.

Thus, if E leaves S free to code the input subjectively, or lets S make his own differential responses to stimuli (Postman, Adams, & Phillips, 1955), the effectiveness of specific retrieval cues provided by E at output presumably depends on the extent of the overlap between the cues and such subjective coding responses that have occurred at input. Experimental manipulation of cues at the time of the presentation of TBR items simply restricts the ways in which various Ss code the input and thus provides E with greater control over what is stored, but the underlying mechanisms are probably the same in both cases. Regardless of whether S codes the TBR items subjectively or follows the suggestions for coding given by E in the form of input cues, a retrieval cue is effective only if the information about it and its relation to the TBR item is stored at the same time with the TBR items. This conclusion is quite consistent with the principle that retrieval depends upon the completeness of reinstatement of original stimuli at the time of recall (Melton, 1963).

Finally, a few words about the effectiveness of double cues. To the extent that specific retrieval cues provide access to the information about TBR items not accessible in absence of such cues, one might expect that the use of multiple retrieval cues would lead to more effective recall performance than the use of single cues. The attempt to demonstrate this relation in the present experiment, however, ended in failure. Double-word cues were found to be no more effective than single-word cues, even though each member of the double-word cue was shown to be quite effective in other parts of the experiment, and even though the double-word cue produced a somewhat higher frequency of “correct guesses” in the free-association task. It looks as if Ss in the present experiment treated the double-word cue as a single unit of information which was as potent in effecting retrieval of the TBR word as was a single-word unit. It is still conceivable, however, that double cues are more effective than single cues if the two members of the double cue are presented separately, rather than simultaneously, at input. Bevan, Dukes, and Avant (1966) have shown, in the terminology of the present paper, that non-cued recall of TBR words repeated within a list is higher for words accompanied by multiple cues than for words accompanied by single cues. The same phenomenon may well hold for cued recall as well.

REFERENCES


Bilodeau, E. A., & Blick, K. A. Courses of misrecall over long-term retention intervals as related to strength of preexperimental habits of word association. Psychological Reports, 1965, 16 (Monogr. Suppl. 6).


EARHARD, M. Cued recall and free recall as a function of the number of items per cue. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 257–263. (a)


Lloyd, K. E. Short term retention as a function of recall point coding. *Psychological Reports*, 1964, 14, 752–754.


(Received July 24, 1967)