Journal of Experimental Psychology
1970, Vol. 86, No. 2, 255–262

ASSOCIATIVE ENCODING AND RETRIEVAL: WEAK AND STRONG CUES [1]

DONALD M. THOMSON AND ENDEL TULVING [2]

University of Toronto

Data from three experiments are reported in support of the encoding specificity hypothesis of retrieval: the effectiveness of retrieval cues depends upon the specific format of encoding of the to-be-remembered (TBR) words at the time of their storage, regardless of how strongly the cues are associated with the TBR words in other situations. In the critical experimental conditions, TBR words were presented for study in presence of weakly associated cue words. Recall of the TBR words in the presence of these cues was greatly facilitated in comparison with noncued recall; recall of the TBR words in presence of their strongest normative associates, which had not been seen at input, did not differ from noncued recall.

This paper is concerned with the problem of the relation between storage and retrieval of information in a simple event-memory experiment. It follows two earlier papers in the series. The first (Tulving & Pearlstone, 1966) demonstrated that many items available in the memory store that cannot be recalled under noncued recall conditions do become accessible in presence of appropriate retrieval cues. The second (Tulving & Osler, 1968) provided experimental evidence in support of the inference that a retrieval cue is effective if, and only if, the information about its relation to the to-be-remembered (TBR) item is stored at the same time as the TBR item itself. Thus, a specific encoding format of the TBR item seems to constitute a prerequisite for the effectiveness of any particular retrieval cue. The point of view reflected in this inference from the data can be referred to as the encoding specificity hypothesis.

The encoding specificity hypothesis, among other things, clearly implies that no cue, however strongly associated with the TBR item or otherwise related to it, can be effective unless the TBR item is specifically encoded with respect to that cue at the time of its storage. This inference contrasts starkly with comparable derivations from what we refer to as the associative continuity hypothesis. This hypothesis forms the core of the explanation of effectiveness of retrieval cues offered by Bilodeau and his associates (Bilodeau & Blick, 1965; Fox, Blick, & Bilodeau, 1964), as well as by Bahrick (1969, 1970). According to the associative continuity hypothesis, if a strong preexperimental association exists between Verbal Units A and B, then A can serve as an effective retrieval cue for B, and vice versa, simply by virtue of the existence of such an association and regardless of the specific nature of encoding events occurring at the time of the storage of the TBR unit.

An anonymous reviewer of the Tulving and Osler (1968) paper had, as the Editor of the *Journal of Experimental Psychology* put it in correspondence, "vehement objections to the conclusions drawn from the experiment," probably because of his own faith in the transsituational associations between nominally identical verbal units. In addition to finding several of the conclusions trivial, the reviewer was especially critical of Tulving and Osler's suggestion that strong preexperimental associates of TBR words are effective retrieval cues only to the extent that these cues overlap with the subjective encoding pattern into which the TBR item is embedded at input. He concluded his long and detailed criticism of the manuscript

[1] This research was supported by the National Research Council of Canada, Grant No. APA-39, and the National Science Foundation, Grant No. GB-3710 and GB-24171X.

[2] Requests for reprints should be sent to E. Tulving, Department of Psychology, Yale University, 333 Cedar Street, New Haven, Connecticut 06510.

with the following query: "Do the authors really wish to say that the *S*'s pre-experimental history cannot be effectively utilized by *E* in controlling recall and that prior experiments have really in effect capitalized upon chance correspondences between the word employed as cue and the differential and idiosyncratic responses made by the *S* while he is in training?" [3]

The three experiments reported in this paper provide an answer to our unknown colleague and critic. The answer, in the form of interpretation of data from three new experiments, consists of three parts: (*a*) *S*'s preexperimental history is indeed important, but only insofar as it determines the encoding of a given TBR item at input; (*b*) effectiveness of cueing at recall is strongly determined by specific encoding of the TBR item at input, and, therefore, (*c*) *S*'s preexperimental history has little effect on recall of an event, such as the occurrence of an otherwise very familiar word in an unfamiliar list, unless this history has influenced the encoding of that event.

In Exp. I, the critical experimental condition—in which the TBR words were accompanied by weakly related associative cues at input and were tested in presence of strongly associated cues at output—constituted one part of an overall design that rather closely followed that used by Tulving and Osler (1968).

Experiment I

Method

Design.—Lists of 24 TBR words were presented to *S*s for study and subsequent recall on a single trial. Three input conditions were combined factorially with four output conditions to yield 12 different experimental treatment conditions. The input conditions were: (*a*) the TBR words were presented alone (Input Cond. O), (*b*) each TBR word was accompanied by a weakly associated cue word (Input Cond. W), and (*c*) each TBR word was accompanied by a strongly associated cue word (Input Cond. S). The output conditions were: (*a*) noncued recall of TBR words (Output Cond. O), (*b*) recall of TBR words in presence of weakly associated cue words (Output Cond. W), (*c*) recall of TBR words in presence of strongly associated cue words (Output Cond. S), and (*d*) free recall of both TBR and cue words (Output Cond. FR). Each of the resulting 12 treatment combinations can be designated in terms of its input and output conditions. Thus, for instance, Cond. O-O was a standard free recall condition, as was Cond. O-FR; in Cond. W-W, TBR words were accompanied by weakly associated cue words at input, and their recall tested in presence of the same cues; in Cond. S-O, each TBR word was accompanied by a strong cue at input, but recall of TBR words was tested in absence of any cues; in Cond. S-FR, *S*s studied TBR words in presence of strongly associated cues and recalled as many cues and TBR words as possible under standard free recall conditions, etc.

Independent groups of 15 *S*s served in each of the 12 treatment combinations. Thus, there were 180 *S*s, first-year female students at the University of Toronto, meeting the service requirement of their introductory psychology course. The assignment of *S*s to treatment conditions was free from any known systematic bias, occurring on a haphazard basis subject to certain restrictions mentioned subsequently.

Materials.—To construct lists of TBR words, 48 response words were selected from the Bilodeau and Howell (1965) free association norms. Each selected TBR word occurred in the norms twice: once as a high-frequency response to a strong-cue stimulus word, and once as a low frequency response to a different weak-cue stimulus word. The two stimulus words eliciting each TBR word in the norms constituted the input and output cues in the experiment. The selection of cues and TBR words was further constrained by the important requirement that the weak and the strong cue of each TBR word be associatively and semantically unrelated to each other on the basis of the two authors' judgment.

Two lists of 24 TBR words (A and B), together with the corresponding sets of weak and strong cue words, were used in order to provide a basis for the generalization of the results over a wider selection of materials. Each list was used with approximately one half of the *S*s in each treatment condition. The mean normative strength of associations between cues and TBR words, in both lists, was 42% for strong cues and 1% for weak cues. Some examples of weak and strong cues and their corresponding TBR words, listed in this order for each triplet, were as follows: train, white, BLACK; knife, meat, STEAK; lamb, dumb, STUPID; hand, woman, MAN; blow, ice, COLD; head, dark, LIGHT.

A practice list, given to all *S*s prior to the experimental list, consisted of 24 proper nouns—names of oceans, rivers, countries, cities, politicans, and monarchs.

Procedure.—Usually four *S*s were tested at a time, although some variations from this procedure occurred owing to differential availability of *S*s. In a given session all *S*s saw the same input list, but they were tested under different output conditions, with the restriction that all output conditions contain an equal number of *S*s in the whole experiment.

[3] D. A. Grant, personal communication, August 28, 1967.

All *S*s were first given the practice list under free recall conditions. The 24 names were shown on a TV screen in front of the room, at the rate of 2 sec/word, and *S*s recorded their recall in recall booklets that had been distributed at the beginning of the session.

Instructions given to *S*s for the experimental list varied according to input and output conditions. The instructions prior to the presentation of the list were read by *E*. Briefly, all *S*s were told they would next see another list on the TV screen, consisting of words typed in capital letters, and that later on they would have to recall as many of these words as they could. The *S*s in Input Cond. W and S were also told that each capitalized word would be accompanied by a related word which should be studied as a possible aid in recalling the capitalized word.

Each TBR word, and each cue-TBR word pair, was presented for 3 sec., for a total time of 72 sec/list. Each TBR word was printed in capital letters, with the cue word, if present at input, appearing to the left of it in lowercase letters. Recall instructions, printed at the top of the appropriate page in the recall booklet, varied according to output conditions. The instructions for all *S*s were: "Now write down all the capitalized WORDS you remember." For *S*s being tested under Output Cond. W and S, the instructions continued: "The words you see typed on the 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 opposite the word to which it is related. If you find this too difficult, however, put down the words you remember anyhow, anywhere on the sheet. The most important thing is to get as many WORDS correct as possible."

The *S*s turned to the recall page only after the list had been presented. Thus, *S*s did not know how their recall would be tested. It can be assumed, therefore, that the availability of TBR words—the amount and organization of the appropriate information in the store—was constant, within limits of random variation, in each group of *S*s studying the material under a given input condition.

The recall sheets in recall booklets contained 24 consecutively numbered lines on two successive pages. In Output Cond. W and S, 12 cue words were typed on a page, with space beside each cue for the appropriate TBR word. In Output Cond. FR, two consecutive pages contained two columns of lines, one for recording cue words and the other for TBR words.

The *S*s were given 5 min. for recording the recall.

Results and Discussion

Lenient scoring was used throughout: *S* was given credit for recall of a TBR word regardless of whether or not the word was paired with its appropriate cue when output cues were provided.

TABLE 1

MEAN NUMBER OF WORDS RECALLED IN TWELVE CONDITIONS OF EXPERIMENT I

Input condition	Output condition			
	O	W	S	FR
O				
M	14.1	11.1	19.0	14.5
SD	2.9	3.2	3.1	3.2
W				
M	10.7	15.7	13.9	9.8 + 6.9[a]
SD	2.4	4.0	3.6	2.1
S				
M	12.2	9.2	20.2	11.1 + 9.4[a]
SD	3.6	3.0	3.4	2.3

[a] Refers to number of cue words recalled.

The mean number of TBR words recalled, with the data pooled for treatment conditions, was 14.03 for List A and 13.03 for List B. Analysis of variance showed this to be a significant difference at the .05 level, but since all interactions involving lists yielded *F* ratios smaller than unity, the data to be described were pooled for both lists.

The mean numbers of words recalled in the 12 conditions of Exp. I, and their corresponding standard deviations, are shown in Table 1. The important features of these results, all statistically reliable at least at the .05 level, can be summarized as follows:

1. The presence of strong cues at output facilitated retrieval of TBR words, both under the condition where TBR words alone were shown in the input list (mean of 19.0 words for Cond. O-S, compared with the mean of 14.1 words for Cond. O-O) and the condition where the cues accompanied TBR words at input (20.2 for Cond. S-S, compared with 12.2 for Cond. S-O). These data confirm earlier similar results (Bilodeau & Blick, 1965; Fox, Blick, & Bilodeau, 1964; Wood, 1967).

2. Weak cues presented at output facilitated retrieval of TBR words, provided that the same cues had accompanied TBR words at input (15.7 for Cond. W-W vs. 10.7 for Cond. W-O). This finding corroborates that of Tulving and Osler (1968).

3. Weak cues presented at output did not facilitate recall of TBR words when they had not been present at input (11.1 for

Cond. O-W vs. 14.1 for Cond. O-O, and 9.2 for Cond. S-W vs. 12.2 for Cond. S-O). These data, too, confirm and extend the generality of Tulving and Osler's (1968) results.

4. Finally, the results of Exp. I showed that strong associative cues, present at output, facilitated recall of TBR words even when the TBR words had been accompanied by different, weak cues at input (mean of 13.9 for Cond. W-S vs. 10.7 for Cond. W-O), but that this facilitative effect was smaller than the facilitative effect of strong cues at output following no cues or strong cues at input (mean of 13.9 vs. 19.0 and 20.2, respectively).

It is this fourth finding that is critical for the evaluation of the respective merits of the encoding specificity and associative continuity hypotheses. The strict interpretation of the former says that recall under Cond. W-S should not have been higher than under Cond. W-O; the strict interpretation of the latter would have to be that recall under Cond. W-S should have been as high as in Cond. O-S. The actual pattern of results clearly vindicates neither of these positions. As frequently happens in crucial experiments, the critical datum fell somewhere between the two extremes predicted from the two points of view.

The encoding specificity hypothesis claims that no cue, regardless of how strongly it might be associated with the TBR item in other situations, can facilitate retrieval of the TBR item in absence of appropriate prior encoding of that item. The W-S condition in Exp. I was meant to bring about specific encoding of TBR words in relation to weak cues and to preclude subjective encoding of TBR words with respect to strong cues. The ambiguous outcome in this critical W-S condition may mean that the encoding specificity hypothesis, or at least its strict interpretation, is not tenable. But it may also mean that the attempt to manipulate the encoding pattern of TBR words under weak-cue input conditions was not quite successful. Suppose that some *S*s ignored all weak input cues in Cond. W-S, or that all *S*s ignored these cues for some of the TBR words. The encoding of such TBR words would then be functionally equivalent to encoding of these words under Input Cond. O, with the consequence that strong cues at output would facilitate their recall. Before rejecting or revising the encoding specificity hypothesis, therefore, it seemed desirable to make a more incisive attempt at experimental manipulation of encoding of TBR items.

Experiment II describes such an attempt. Here the critical condition—the switch from weak cues at input to strong cues at output—was introduced only after *S*s had been tested in several lists with output cues that completely matched the input cues. It was hoped that repeated testing of *S*s under one and the same set of cue conditions would encourage them to pay close attention to the cues present at input, and discourage them from encoding TBR items in the normal manner characteristic of Input Cond. O and S.

Experiment II

Method

Design.—The design of Exp. II is schematically shown in Table 2. Each of six independent groups of *S*s was tested with four successive lists, each containing 24 TBR words. Each list was presented once, each TBR word either occurring alone (Input Cond. O) or accompanied by a weakly associated cue word (Input Cond. W). Recall of TBR words was tested in absence of any cues, in presence of weakly associated cues, or in presence of strongly associated cue words. These three output conditions are labelled O, W, and S, respectively. Each experimental treatment condition can again be described by the combination of input and output cueing conditions, as in Exp. I.

Table 2 shows that Groups 1, 2, and 3 studied and recalled the first two lists under the O-O treatment condition, while Groups 4, 5, and 6 did the same under the W-W condition. With respect to the treatment given to the groups on the third list, the experiment can be thought of as a 2×3 factorial, in which 2 input-cue conditions, O and W, perfectly confounded with input and output conditions in the first two lists, were orthogonally combined with 3 output-cue conditions (O, W, and S).

The critical conditions were created by List 3 in Group 6 and List 4 in Group 5. In these conditions, *S*s studied TBR words in presence of weak cues and recalled them in presence of strong cues.

Materials and procedure.—The two lists of 24 TBR words and their corresponding cue words

constructed for Exp. I were used as Lists 3 and 4 in this experiment. Two additional lists of 24 words and weak cues were constructed from two sets of free association norms (Bilodeau & Howell, 1965; Riegel, 1965) to serve as Lists 1 and 2 in this experiment.

The procedure was identical with that of Exp. I with respect to all important features, with the exception that (*a*) no practice list was given in Exp. II, and (*b*) each *S* was now tested with four successive lists. Words were presented visually on a closed-circuit TV screen, each TBR word and each cue-TBR word pair appearing for 3 sec. The TBR words were again printed in capital letters, cue words on their left in lowercase letters. Three minutes were provided for a written recall test on each list, *S*s recording their responses in booklets containing, depending upon output conditions, 24 numbered lines, 24 weak cues, or 24 strong cues. The *S*s were not explicitly told how their recall was going to be tested on any particular list. Instructions before and after the presentation of a list were essentially the same as those used in Exp. I.

Results and Discussion

Again, lenient scoring was employed. The mean numbers of TBR words recalled by all six groups on each of the four lists, together with standard deviations, are recorded in Table 2.

The important facts summarized in Table 2 are the following:

1. The *S*s in Groups 1, 2, and 3 recalled, on the average, 11.0 TBR words from Lists 1 and 2. These lists were studied and tested under noncued conditions. The unexpected switch to testing of TBR words in presence of strong cues (List 4 in Group 1, and List 3 in Group 3) produced a sizable facilitation in recall, the means being 16.8 and 16.3. Similar switch to testing of TBR words in presence of weak cues, however, produced no facilitation: the mean number of words recalled from List 3 by Group 2 was 10.3. The pattern of these data is identical with Exp. I.

2. The *S*s in Groups 4, 5, and 6 recalled, on the average, 18.3 TBR words from Lists 1 and 2. The TBR words in these lists were studied and tested in presence of weak cues. The unexpected switch to noncued recall condition on List 3 for Group 4 produced a striking reduction in performance, the mean number of words recalled being 7.1. The unexpected switch to testing of TBR words in presence of strong cues (List 4 in Group 5, and List 3 in Group 6) produced an equally striking loss in recall, the two means being reduced to 5.4 and 5.5, respectively. These two means were not significantly different from the mean number of words recalled by Group 4 on List 3 under noncued conditions.

The critical finding here again has to do with the effects of strong cues on recall. Strong cues clearly facilitated recall under conditions where *S*s were left free to subjectively encode the TBR words with a view to expected noncued recall test. Identical strong cues, however, completely failed to augment recall when the TBR words

TABLE 2

DESIGN OF EXPERIMENT II AND SUMMARY OF THE RECALL DATA

Group	Successive lists			
	1	2	3	4
1				
Input condition	0	0	0	0
Output condition	0	0	0	S
$\bar{X}$ TBR recalled	10.6	9.6	11.7	16.8
SD	1.6	3.5	1.8	3.8
2				
Input condition	0	0	0	0
Output condition	0	0	W	0
$\bar{X}$ TBR recalled	12.6	10.9	10.3	13.4
SD	4.5	3.0	2.8	5.1
3				
Input condition	0	0	0	0
Output condition	0	0	S	0
$\bar{X}$ TBR recalled	11.5	10.5	16.3	10.7
SD	2.0	3.1	3.7	4.0
4				
Input condition	W	W	W	W
Output condition	W	W	0	W
$\bar{X}$ TBR recalled	17.2	19.4	7.1	16.5
SD	3.7	3.1	3.4	5.3
5				
Input condition	W	W	W	W
Output condition	W	W	W	S
$\bar{X}$ TBR recalled	17.1	19.5	19.8	5.4
SD	3.4	2.9	3.1	4.7
6				
Input condition	W	W	W	W
Output condition	W	W	S	W
$\bar{X}$ TBR recalled	18.0	18.4	5.5	15.6
SD	3.0	2.8	4.1	2.5

Note.—Input and output conditions are abbreviated as follows: O, no cues; W, weak cues; and S, strong cues.

were presumably encoded specifically with respect to their accompanying weak cues.

The outcome of Exp. II thus seems consistent with the implications of the encoding specificity hypothesis, and inconsistent with the associative continuity hypothesis. It seems reasonable to assume, therefore, that the failure to completely demonstrate the highly critical role of encoding processes in Exp. I may have reflected the failure of intended manipulation of these processes in Cond. W-S. However, it might be argued that strong retrieval cues failed to facilitate recall in the W-S conditions in Exp. II for reasons other than encoding specificity. It could be assumed, for instance, that *S*s in the W-W conditions, in the course of studying and recalling the first two or three lists, developed a set to respond to retrieval cues with weak associates. If this set persisted when strong cues were provided in the critical W-S conditions, *S*s could not have responded with strong associates of these cues as correct TBR words. It could also be assumed that the sudden unannounced switch to the strong cues in the W-S conditions in some other way confused *S*s and that this confusion prevented them from taking maximum advantage of stored information about TBR words at the time of the recall test.

Experiment III represented an attempt to test the reality of this set or confusion notion, using a mixed-list paradigm.

Experiment III

Method

Design.—Each of 24 *S*s was treated exactly alike, being tested with three successive lists. Each list contained 24 TBR words, one half of them accompanied by weak cues and the other half by strong cues, both at input and test. In Lists 1 and 2, the output cue of a given TBR word always matched its input cue. Thus, Lists 1 and 2 contained two experimental conditions, W-W and S-S, each represented by 12 TBR words. In the third list, half of the weak input cues were changed to appropriate strong cues at output, and half of the strong input cues were changed to appropriate weak cues. The remaining TBR words were cued at the recall test with their input cues. Thus, List 3 generated data for four different intralist experimental conditions, W-W, W-S, S-S, and S-W, each represented by 6 TBR words. Two different sets of three lists were used, each set with one half of the *S*s. In addition, there were four alternative versions of List 3 in both sets, such that each TBR word occurred in each of the four treatment conditions equally often. The order of presenting TBR words in the study list and the cues in the recall booklets was determined randomly.

Materials and procedure.—The two lists of 24 TBR words and their associated cue words constructed for Exp. I constituted the two sets of List 3 words in this experiment. Four other lists of 24 TBR words and cues were constructed from the association norms (Bilodeau & Howell, 1965; Riegel, 1965) to serve as the first two lists.

In most aspects the procedure was identical to that of Exp. I and II. Usually two or three *S*s were tested at a time.

The *S*s were first shown a short practice list of four TBR words, the names of famous people and places, together with related cue words, before the three experimental lists were presented. Prior to the presentation of the first experimental list, instructions given in Exp. I and II for Input Cond. W were read to *S*s. Each cue-TBR word was presented on the closed-circuit TV screen for 3 sec. Recall instructions, typed at the top of a page in the recall booklets, were essentially the same as Output Cond. W instructions in Exp. I and II. In addition, for List 3, the recall instructions stated that half of the cues had appeared with the TBR words in the input list and half were new but related words. The two types of cues were identified for *S*s by being presented in columns headed "old" and "new." The *S*s were given as much time as they wanted for the written recall test. They hardly ever took more than 3 min. for recall.

Results and Discussion

The mean proportion of TBR words recalled (lenient scoring) for the different input and output conditions in the three lists, together with standard deviations of these proportions, are presented in Table 3.

The main features of the data were as follows:

1. In Lists 1 and 2, recall of TBR words was higher in Cond. S-S than in Cond. W-W. The same tendency in List 3 was not quite statistically significant.

2. Recall in presence of strong output cues was much lower in Cond. W-S than in Cond. W-W—proportions of .33 vs. .73, respectively. The proportion of .33 in Cond. W-S, equivalent to 8 words out of 24, is clearly of the same order of magnitude as noncued recall in Exp. I and II.

3. Recall in Cond. S-W was practically zero. Only 5 *S*s out of 24 recalled one word

TABLE 3

PROPORTION OF WORDS CORRECTLY RECALLED IN EXPERIMENT III

Input cue	Output cue					
	Strong			Weak		
	List					
	1	2	3	1	2	3
Strong						
Proportion	.840	.840	.833	—	—	.035
SD	.143	.195	.170			.085
Weak						
Proportion	—	—	.326	.750	.753	.729
SD			.297	.207	.235	.234

each in this condition, for a mean proportion of .03.

Since the important fact—that strong associative cues do not facilitate recall if specific encoding of TBR items is not appropriate to these cues—was again clearly suggested by the overall pattern of the data, there is no support for the set or the confusion notion as a plausible explanation for the critical results of Exp. II. Strong cues presented at recall in Cond. W-S were not effective in augmenting access to stored TBR words, despite the fact that in Exp. III no set to respond only with weak associates of retrieval cues could have developed, and despite the fact that any confusion *S*s may have experienced as a consequence of switching cues was presumably minimized by recall instructions in List 3.

The striking finding that recall failed almost completely in the S-W condition in Exp. 3 is at variance with a much higher level of recall under the same condition in Exp. 1. The reasons for *S*s' failure to recall any words in presence of weak cues after they had been shown in the company of strong cues at input are not clear.

DISCUSSION

The purpose of these experiments was to evaluate two theoretical views from which explanations of the operation of retrieval cues in event memory can be derived: the associative continuity and encoding specificity hypotheses. While in many experimental situations the two hypotheses make identical predictions about cued recall, in certain cases the predictions are different. In all three experiments reported in this paper, the critical conditions were those in which *S*s studied TBR words in presence of their weak normative associates and then recalled these words (*a*) in a noncued recall test, (*b*) in presence of the previously seen weak input cues, or (*c*) in presence of strong normative associates not seen in the input list. The associative continuity hypothesis would predict that recall in the *c* condition should be considerably higher than in the *a* condition, while the encoding specificity hypothesis predicts no difference.

The overall pattern of results favored the encoding specificity hypothesis: when *S*s were induced to encode TBR words with respect to weak cues at input, preexperimentally defined strong cues introduced at output failed to facilitate recall. According to the encoding specificity hypothesis, retrieval of event information can only be effected by retrieval cues corresponding to a part of the total encoding pattern representing the perceptual cognitive registration of the occurrence of the event. Thus, the cue "white" cannot provide access to stored information about the occurrence of BLACK as a TBR word, if BLACK has been encoded as part of the "train-BLACK" complex, or as part of a unique event in a series of unique events. The two lexical units, BLACK and BLACK, are identical, but the encoded engram of the unique event BLACK, in the context of "train," and in the context of a specific set of TBR events, may be as different from the pattern of neural excitation corresponding to the generalized concept of BLACK as a beautiful and talented actress receiving an Oscar is different from any one of millions of stars twinkling in the endless night.

To the extent that our data support the encoding specificity hypothesis, they rule out the associative continuity hypothesis. We will have to leave it to the proponents of that hypothesis to explain how it might be brought into line with the overall pattern of these data, because we cannot think of any reasonably simple way of doing it. This does not mean that the encoding specificity point of view can readily account for all details of the data. It cannot. For instance, the asymmetry in recall scores between the W-S and S-W conditions in Exp. III appears to be somewhat more compatible with associative continuity than with encoding specificity, suggesting that further thought on the distinction between the two views is needed.

It is only on balance that the encoding hypothesis appears to emerge less ruffled from this contact with the experimental data than does the associative continuity hypothesis.

It is not immediately clear why our data are apparently at variance with, and how the encoding specificity hypothesis could explain, Bahrick's (1969, 1970) findings revealing considerable facilitation of recall by prompts (cues) of various strengths presented only at output. It is possible that Bahrick's data reflect the influence of overlap between the encoding patterns of some weak and strong cues, such as "child" and "boy," "weather" and "hot," "parade" and "banner," "small" and "long," and "hospital" and "physician" (Bahrick, 1969). In our present experiments, as we noted earlier, weak and strong cues were selected so as to be semantically unrelated. It is also possible that the specific instructions given to Bahrick's *S*s—that, if necessary, they should guess in response to all prompts—are partly responsible for the increased correspondence, under conditions of prompting, between *S*s' responses and *E*'s tally sheet. But the evaluation of these and other possible reasons for the apparent discrepancies between Bahrick's findings and ours is difficult at the present time. Further experimental and theoretical analysis suggested by these and other gaps in our knowledge and understanding of effectiveness of retrieval cues should also include a critical appraisal of appropriateness of different methods of measuring this effectiveness, another source of as yet unresolved disagreement between associative continuity and encoding specificity positions.

REFERENCES

BAHRICK, H. P. Measurement of memory by prompted recall. *Journal of Experimental Psychology,* 1969, **79**, 213–219.

BAHRICK, H. P. Two-phase model for prompted recall. *Psychological Review,* 1970, **77**, 215–222.

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. No. 6-V16), 1173–1192.

BILODEAU, E. A., & HOWELL, D. C. *Free association norms.* (Catalog No. D210.2:F87) Washington, D.C.: United States Government Printing Office, 1965.

FOX, P. W., BLICK, K. A., & BILODEAU, E. A. Stimulation and prediction of verbal recall and misrecall. *Journal of Experimental Psychology,* 1964, **68**, 321–322.

RIEGEL, K. F. *Free associative responses to the 200 stimuli of the Michigan restricted association norms.* (USPHS Rep. No. 8) Ann Arbor, Mich.: University of Michigan, 1965. (United States Public Health Service Grant MH 07619.)

TULVING, E., & OSLER, S. Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology,* 1968, **77**, 593–601.

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

WOOD, G. Category names as cues for the recall of category instances. *Psychonomic Science,* 1967, **9**, 323–324.

(Received May 16, 1970)