Continuity between Recall and Recognition

Endel Tulving and Michael J. Watkins
University of Toronto and Yale University

The discontinuity hypothesis is that recall and recognition are in some sense fundamentally different memory processes. The continuity hypothesis is that retrieval in both modes is essentially the same, a joint product of the information stored in the past and that in the immediate environment. Data from a simple experiment and a brief discussion of other studies in the literature support the proposal that the continuity view is both more parsimonious and more fruitful than the discontinuity view.

A critical problem of long standing in psychological study of memory is concerned with the relation between recall and recognition. In what sense are they the same, and in what sense are they different? The procedures involved in measuring unaided recall and recognition do exhibit obvious differences, as do instructions given to the subjects. But do recall and recognition procedures and measures reflect fundamentally discontinuous psychological processes, or are the processes underlying the two kinds of mnemonic performance essentially continuous?

Both points of view are amply represented in the literature. The discontinuity hypothesis has assumed many different forms, two of which seem to be currently more popular than others. The first is that in recall the context is provided by the experimenter and the subject must produce the missing focal element, while in recognition the focal element is given and the subject's task is to retrieve its former context (e.g., Anderson and Bower, 1972; Hollingworth, 1913; Norman, 1968). The second is that recall consists of two stages, search and decision, or response production and response identification, while recognition requires decision or identification only (e.g., Adams, 1967; Kintsch, 1970; Müller, 1913; Underwood, 1972).

The continuity hypothesis is implicit in explanations of superiority of recognition over recall as differentially sensitive measures of associative strength (e.g., Bahrick, 1965; Postman, 1963), and it has been accepted as a working assumption by those who think of memory as a joint product

American Journal of Psychology
1973, Vol. 86, No. 4, pp. 739–748

of the information stored in the past and information present in the person's immediate cognitive environment (e.g., Shiffrin and Atkinson, 1969; Tulving and Pearlstone, 1966; Tulving and Thomson, 1971).

In this paper, we report a simple experiment demonstrating one form of continuity between recall and recognition. In the experiment, the format and amount of information stored in memory about each word in a list are held constant and the retrieval of this information is observed as a function of number of retrieval cues present at the time of the test. After describing the experiment and its results, we will return to the question of the relation between recall and recognition and say why we prefer the continuity hypothesis.

METHOD

Each of 20 subjects, undergraduate students at Yale University, was given an immediate recall test on each of five lists of 28 five-letter words. Each list was presented for study on a single trial. One of the five lists was tested under the conditions of noncued recall, while component letters of the words were used as retrieval cues in tests of the other four lists. These retrieval cues consisted of the first two, three, four, or five letters of each word. Each of the four levels of this independent variable was represented in each of the four lists by seven words. The overall design was such that all lists yielded data equally frequently for all cue conditions, the noncued test of a list occurred equally frequently in all five possible list-order positions, and all words within a list were tested equally frequently with all cue sizes.

There were two sets of five lists each. They were composed by drawing words from two sets, A and B, of systematically related words. For each word in set A there was a word in set B that differed from it only with respect to the fifth letter. Thus, for instance, words like ALLOY, CRUST, GRAPE, SPARK, and TOWER in set A had their counterparts in set B in the words ALLOW, CRUSH, GRAPH, SPARE, and TOWEL. Five lists were composed of words drawn from set A by a random procedure, with the restriction that no two words in a list begin with the same two letters, and then corresponding words from set B were used to make up five corresponding yoked lists. The two sets of word lists were used equally frequently with two subgroups of subjects.

— Procedure — Subjects were tested in small groups of two to five persons. Each group was exposed to the same five lists in the same order, but each subject within the group received a different test of each list in keeping with the requirements of the design.

The first part of the instructions was: "I am going to show you a series of lists of words for you to remember. Each list will appear, one word at a time, on the television. After each list I shall give you a piece of paper with a list of clues written down the left-hand side; these clues may help you to recall the words. Your task will be to consider each clue in turn, and to write down any word it causes you to remember from the list alongside the clue. If a clue does not help you to recall a word then you must put a cross next to it. Do not guess,

and do not write down any words that the clues suggest unless you remember them as being in the list."

The rest of the instructions informed subjects of the procedure to be used in testing and told them they would be tested with a number of lists, each followed by a cued test. The subjects were also given four words as examples of the kinds of items to be remembered and a cue of different size for each word.

The words were presented on a closed-circuit TV screen, at the rate of 2 sec per word. At the end of the presentation of the list, the subjects proceeded to recall the words using a recall sheet. In the case of cued tests, the recall sheet contained seven cues at each level — the first two, three, four, or five letters of each word — ordered essentially haphazardly in a vertical column along the left-hand side of the sheet. The only restriction on the ordering of cues was that no word be cued for recall with fewer than seven input or output events interpolated between its presentation and test. The subject moved a card down the column of cues at an experimenter-signaled rate of 6 sec per cue and recorded his recall in writing.

In the case of the one (unexpected) noncued test of a list, the subject was given, after the presentation of the list, an otherwise blank sheet with these instructions typed at the top: "*For this list only,* you will be given no clues at all. Your task is to try to remember as many words as you can, and write them down in the order that they occur to you. You therefore work at your own speed, and ignore the instructions given to the other subject[s] to attend to the next clue. [For the remaining lists you will be given clues, just as in the previous lists]."

Two common practice lists were given to all subjects before presentation of the five experimental lists. These practice lists were similar to the experimental lists. They contained 28 five-letter words and their recall was tested with the two-, three-, four-, and five-letter cues. Subjects were not alerted to the difference between the practice and experimental lists. The practice lists were used in an attempt to induce subjects to encode the list words appropriately with respect to the cues to be used. Recall protocols from the practice lists were not scored.

RESULTS

The main data are shown in Figure 1. The proportion of recalled words is plotted as a function of the number of letter cues from two to five. The noncued recall data are also included in the graph for whatever interest such comparison might have. The data depicted in Figure 1 describe a continuous function between two-letter cues and five-letter cues. Sign tests showed recall at each level of cue size to be significantly different from that at neighboring levels. Recall with two-letter cues, however, was not significantly different from noncued recall.

The intrusion data are summarized in Table 1. The total number of words produced that were not members of the list under test is shown for two classes of words: (*a*) those that would have been classified as correct if the list under test had been its yoked mate from the other set

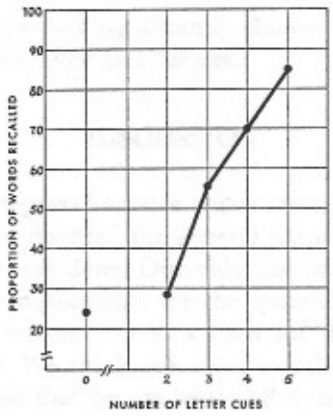

Fig. 1. Proportion of words recalled as a function of number of letter cues

and (*b*) all other words, mostly intrusions from earlier lists. The number of intrusions in the first (*a*) class could be used to correct the recall data, but whatever assumptions one would adopt in making such a correction, the correction would not change the finding of primary interest in this experiment, the continuous function relating the size of the cue to the probability of retrieval.

It should be noted that subjects could have obtained a perfect score on recall with five-letter cues by following the clever strategy of simply copying all five-letter cues. It is not known to what extent such a strategy was used. There were 6 subjects among the 20 who did obtain a perfect score of 28/28 (seven words in each of four lists) with the five-letter cues. Since their performance with other cue sizes, including noncued recall, was not discriminably different from that of the other 14 subjects, it is possible that they, or some of them, did use this memory-conserving stratagem. But again, it is difficult to see how any correction of the data

Table 1. Total number of intrusions in the recall protocols of 20 subjects

Type of intrusion	Number of letter cues				
	0	2	3	4	5
Complementary-set words	5	2	16	29	1[a]
Other words	18	42	36	19	1

[a] Cue THING rewritten by subject as THINK.

for the five-letter cues would significantly change the nature of the continuous relation between recall and cue size.

DISCUSSION

The data we have reported in this paper are trivial, in the sense that few people would have doubted the general nature of the outcome even before the experiment was done. Our only justification for reporting the experiment lies in its implications for the question of whether there is, in some theoretically relevant sense, a need for the distinction between recall and recognition. We think our data, together with other evidence in the literature, suggest that we are better off without such a distinction. We propose that the business of making sense out of the experimentally observed phenomena of memory would be simpler, and theory construction easier, if we adopted the continuity point of view.

Even at the level of pure description, the distinction causes unnecessary trouble. Consider the present data. We can refer to the test with five-letter cues as a recognition test, but what do we do with the other cue conditions? Where does recognition end and recall begin? Can we say that the test with four-letter cues was still essentially one of recognition?

One could, of course, assume that the subject used the four letters of the cue to generate all lexically meaningful completions of the cue and performed a 'recognition' test on generated alternatives. The extension of this argument, then, would be that recall drops with the reduction in cue size simply because the generation of candidate words for the implicit recognition test becomes increasingly difficult. A seductively simple theory of this kind is favored by many contemporary researchers (e.g., Bower, 1970; Kintsch, 1970; Underwood, 1972). It has at least two difficulties.

First, the difference in recall between three- and four-letter cues in our experiment was of the same magnitude as the difference between four- and five-letter cues, despite the fact that the probability of generating the target word from the first three letters is very much lower than the probability of generating it from the first four letters. Why are such differences in generation not reflected in the recall of the words?

The second and perhaps more serious difficulty lies in the observation that generating word completions, even with four-letter cues, is a vexing and time-consuming task, especially if the subject has studied one alternative completion in the list and is then asked both to recall the seen alternative from the list and to generate other possible completions from

semantic memory. We suggest that a skeptical reader conduct a simple experiment to convince himself or herself of this matter.

Let the reader have his subject study a list of, say, five-letter words, and then test him with three- and four-letter cues. Then have him study another list, say, consisting of the words AWARD, CHIMP, HEAVY, SHELF, BREAD, DREAM, ADORE, PRONE, THROB, PLUME, ALLOY, CROWD, APPLY, BLANK, QUILT, SHORE, AUDIO, BASIN, SCALP, CHINK, BOOTY, CHESS, SCORE, CIVIL, FLOOR, GRAPH, CHARM, STEAM. Give him, say, 2 min to learn all these words, presented on a sheet. In the test for this list, provide the first four letters of each word as a cue and ask the subject to recall the word, and at the same time, to generate from his semantic memory another five-letter nonplural word that has the same first four letters. The results will show the great difficulty of this task. The subject will take a long time to complete it, much longer than he needed to recall words when he had four-letter cues. If the internal generation of response alternatives is part of the rather fast recall process, why is the overt production of the alternatives already generated so slow?

The data we have described do not, of course, prove that recall and recognition, in the historical sense of the two words, are not discontinuous. No single experiment can do that. But they do add to the growing list of findings and observations that make the continuity position a more attractive one.

What is the continuity position? It is the view that retrieval always depends both on the availability of information in the memory store and on the accessibility of that information through appropriate retrieval cues, the latter being fragmentary knowledge the system possesses before retrieval about the material to be retrieved. While the exact mechanism remains to be specified, we find it helpful to think that the information contained in the retrieval cues somehow actively combines or interacts with the stored information to create the memory of a previously experienced event. Retrieval cues may vary greatly in their effectiveness, depending on the relation between the format of stored information and the encoding of the cues. The retrieval cues in a typical recognition task are more appropriate to the stored information than are the cues in typical recall tasks, and hence recognition is usually easier than recall. But conditions can be readily created where retrieval cues in what would be normally classified as a recall task are more effective than those in a recognition task (e.g., Tulving, 1968; Tulving and Thomson, 1972). The consequent reversal of the usual relation between recall and recognition measures is difficult to reconcile with the discontinuity hypothesis, yet it is entirely compatible with the continuity point of view.

We can further illustrate the continuity point of view by considering how it would handle four types of arguments that have been used to support the discontinuity hypothesis.

The first sort of argument is that in the recognition test the list item is physically present and the subject needs only to check it as 'old,' while in recall the correct response must be produced by the subject, either in absence of any specific cues or in response to various cues other than the test word itself. From the continuity point of view, however, the dichotomization of all retrieval cues into two mutually exclusive categories — copies of input items versus all other things — and the conclusion that there is some fundamental difference between processes that corresponds to the logical dichotomy between the two kinds of 'fragmentary knowledge' make little intuitive sense. For instance, consider a case where the to-be-remembered word is PLANT. The following retrieval cues can be expected to be effective in various degrees: _____, P____, PL___, PLAN_, _LANT, SLANT, PLANT, PLANTS, PLANET, PLANTER, PLANTERS, PLANTATION, FACTORY, FLOWER. The continuity theorist finds it somewhat difficult to imagine that Mother Nature arranged matters of memory retrieval in such a way that the process is basically different in the case of one of these cues — the cue _____, for example, or the cue PLANT — than it is in the case of all the others. If a person sees, say, a human face in one orientation, does he recognize it when it appears again in the same orientation and do something different when the orientation is changed? The continuity theorist says no.

The second type of argument is that the question to the subject in the recall task is, What was that word? whereas in the recognition task it is, Is this that word? (Underwood, 1972, p. 6). The continuity point of view holds that the intended question, despite the traditional differences in its wording, is always the same, namely, Does this cue remind you of the item or items you are trying to remember? The question, like its two discontinuity versions, must always be presented in a context of general retrieval instructions that specify, implicitly or explicitly, the larger set (list, category, pair, or whatnot) of target items of which the person is trying to retrieve one or more elements.

The third argument is based on introspective evidence available to experimenters as organisms rather similar to their subjects, as well as on verbal comments made by the subjects about their introspective observations. This evidence frequently suggests that recognition is usually effortless, sure, and fast, while recall sometimes succeeds only after a laborious, halting, and slow process. The less said about such introspections the better; introspectively the earth is flat and the sun goes around it exactly

like the moon. Objectively measurable retrieval times, on the other hand, depend on many variables and therefore do not provide a reliable guide to dichotomization of retrieval processes.

The fourth sort of argument derives from some experimental evidence that certain independent variables, such as word frequency, intentionality of learning, and organization of material, interact with the method of testing, that is, with recall and recognition, in nontrivial ways. Such evidence is thought by some to be more compatible with the discontinuity hypothesis than the continuity point of view (e.g., Anderson and Bower, 1972; Kintsch, 1970; McCormack, 1972; Underwood, 1972). This is not the place to consider the relevance of this evidence as thoroughly as its alleged importance obviously requires. The general stance of the continuity point of view vis-à-vis this kind of evidence, however, can be indicated: these data illustrate, even if they do not particularly illuminate, the principle that the effectiveness of a retrieval cue depends on the *encoding format* of the target item.

For instance, first, if the initial letter of a word is to serve as an effective retrieval cue, the word must be stored as having that initial letter (e.g., Earhard, 1967; 1969). Second, a word semantically-associatively related to a to-be-remembered word facilitates the latter's recall if the subject is permitted to encode the to-be-remembered word in his own way, but not if he is induced to encode it in a specific relation to a third word (Thomson and Tulving, 1970, Experiment II). Third, if a meaningful word is encoded as a part of a larger memory unit, it can be recalled as a part of that unit, but not always recognized independently of the unit (Tulving, 1968). Fourth, one and the same retrieval cue, or category name, facilitates or inhibits recall, in relation to the free-recall baseline, depending on the input context of target words (experiment by Bobrow and Light reported by Bower, 1970). And fifth, subjects shown pictures and names of objects and prepared for a recall test may do better on the recall test and worse on the recognition test than do subjects who have been prepared for the recognition test (Tversky, 1973). The dependence of mnemonic performance on the relation between the encoding format of the to-be-remembered event and specific retrieval cues, whether experimentally specifiable or not, thus manifests itself not only in situations comparing recall and recognition but also in situations where only recall measures are taken. Studies of the interaction between mode of retrieval and other variables may indeed be relevant to the problem of the relation between recall and recognition, but that relevance pales in the light of the broader significance of data like those we have just outlined.

We prefer the continuity point of view to the discontinuity hypothesis for two reasons. It permits a meaningful and consistent interpretation of all the data that the discontinuity view can handle, and then some. And it creates many interesting problems shut out by the acceptance of the discontinuity hypothesis.

Notes

This research was supported by Grant 24171X from the National Science Foundation. The authors are grateful to Olga Watkins for experimental assistance. E. Tulving's dual affiliation is with the University of Toronto and Yale, M. J. Watkins is at Yale. Received for publication December 27, 1972.

References

Adams, J. A. 1967. *Human memory.* New York: McGraw-Hill.

Anderson, J. R., and Bower, G. H. 1972. Recognition and retrieval processes in free recall. *Psychological Review* 79:97–123.

Bahrick, H. P. 1965. The ebb of retention. *Psychological Review* 72:60–73.

Bower, G. H. 1970. Organizational factors in memory. *Cognitive Psychology* 1:18–46.

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

Earhard, M. 1969. Storage and retrieval of words encoded in memory. *Journal of Experimental Psychology* 80:412–418.

Hollingworth, H. C. 1913. Characteristic differences between recall and recognition. *American Journal of Psychology* 24:532–544.

Kintsch, W. 1970. Models for free recall and recognition. In *Models of human memory,* ed. D. A. Norman, pp. 331–373. New York: Academic Press.

McCormack, P. D. 1972. Recognition memory: How complex a retrieval system? *Canadian Journal of Psychology* 26:19–41.

Müller, G. E. 1913. Zur Analyse der Gedächtnistätigkeit und des Vorstellungsverlaufes. *Zeitschrift für Psychologie* suppl. vol. 8.

Norman, D. A. 1968. Toward a theory of memory and attention. *Psychological Review* 75:522–536.

Postman, L. 1963. One-trial learning. In *Verbal behavior and verbal learning: Problems and processes,* ed. C. N. Cofer and Barbara S. Musgrave, pp. 295–321. New York: McGraw-Hill.

Shiffrin, R. M., and Atkinson, R. C. 1969. Storage and retrieval processes in long-term memory. *Psychological Review* 76:179–193.

Thomson, D. M., and Tulving, E. 1970. Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology* 86:255–262.

Tulving, E. 1968. When is recall higher than recognition? *Psychonomic Science* 10:53–54.

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

Tulving, E., and Thomson, D. M. 1971. Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology* 87:116–124.

Tulving, E., and Thomson, D. M. 1972. Word-blindness in episodic memory. *Psychonomic Science* 29:262.

Tversky, B. 1973. Encoding processes in recognition and recall. *Cognitive Psychology* 5:275–287.

Underwood, B. J. 1972. Are we overloading memory? In *Coding processes in human memory*, ed. A. W. Melton and E. Martin, pp. 1–23. New York: Academic Press.