

Relation between recognition and recognition failure of recallable words

ENDEL TULVING and SANDOR WISEMAN
University of Toronto, Toronto, Canada M5S 1A1

Data are presented from 12 different previously published experiments to demonstrate a systematic relation between recognition of all studied list words and recognition failure of recallable words. This relation appears to be independent of many specific procedural details of the experiments analyzed. The data support the statement that recognition failure of recallable words appears to be a general phenomenon that, as far as is known, always occurs whenever recognition is imperfect.

In several recent experiments, recall of studied list words has been found to be higher than their recognition (Tulving, 1974; Tulving & Thomson, 1973; Watkins & Tulving, 1975; Wiseman & Tulving, in press). The experimental paradigm under which such recall superiority has been observed has the following components: (a) The to-be-remembered (TBR) word is presented on a single trial in the context of another word, the list cue, which is meaningfully related to the TBR word. (b) The subject studies the TBR word in the expectation that recall of the TBR word is going to be tested in the presence of the list cue. (c) Following the study trial, and usually after some interpolated activity, the subject is given a recognition test in which copies of TBR words are mixed with lure words and the subject is required to identify the TBR words. (d) After the recognition test, the list cues are presented to the subject, and he is asked to recall the corresponding TBR words. Specific details of this procedure, such as the nature of the activity interpolated between the study list and the recognition test, the nature of the recognition test, and the extent of subjects' knowledge of the test sequence may be varied without affecting the superiority of recall over recognition (Watkins & Tulving, 1975; Wiseman & Tulving, in press). However, Reder, Anderson, and Bjork (1974) have shown that recall superiority is reversed when TBR items are low-frequency words, and Postman (in press) has also reported *recognition* superiority, although in his experiment the reasons for the discrepant results are not identifiable.

The phenomenon of recall superiority is of considerable theoretical interest. The fact that recall is higher than recognition, under the conditions of the paradigm as described, necessarily means that subjects fail to recognize some words which they subsequently can recall. This finding of recognition failure of recallable words, entailed in the recall superiority effect, has implications for the form that theories of

Table 1
Recognition/Recall Contingency Table

Recognition	Recall		Total
	Success	Failure	
Success	R_n, R_c	$\overline{R}_n, \overline{R}_c$	R_n
Failure	\overline{R}_n, R_c	$\overline{R}_n, \overline{R}_c$	\overline{R}_n
Total	R_c	\overline{R}_c	

retrieval must assume; indeed, it already has led to some modifications of the currently popular generation-recognition models of recall (e.g., Anderson & Bower, 1974; Kintsch, 1974; Reder et al., 1974).

The purpose of this short article is twofold: (a) we summarize evidence from a number of experiments in support of the statement that recognition failure of recallable words appears to be a general phenomenon that, as far as we know, always occurs to some degree in the paradigm under discussion here, and (b) we demonstrate that the magnitude of the phenomenon in any particular experimental condition is quite accurately predicted by the overall level of recognition in that condition.

The magnitude of recognition failure of recallable words can be indexed by the conditional probability that a TBR item is not recognized, given that it is recalled (Watkins & Tulving, 1975). This conditional probability can be calculated from the data tabulated in a fourfold contingency table such as the one shown in Table 1. The cell entries in Table 1 represent joint probabilities of two possible outcomes for a TBR word in the recognition test, successful recognition (R_n) or failure to recognize (\overline{R}_n), and two parallel outcomes in recall, success (R_c) or failure (\overline{R}_c). The conditional probability of recognition failure, given recall success, equals the ratio of joint probability of recognition failure and recall success (\overline{R}_n, R_c) to the simple probability of recall success (R_c): $P(\overline{R}_n/R_c) = P(\overline{R}_n, R_c)/P(R_c)$. Instead of using this measure, however, it is, for certain purposes, more convenient to work with its complement, recognition success of recallable words. The recognition success measure is provided by the

Table 2
Recognition/Recall Contingency Data*

Recognition	Recall		Total
	Rc	$\bar{R}c$	
Rn	.26	.06	.32
$\bar{R}n$.37	.31	.68
Total	.63	.37	

*Watkins and Tulving, 1975, Experiment 3, subject-made test.

conditional probability that an item is recognized, given that it is recalled: $P(Rn/Rc) = P(Rn,Rc)/P(Rc)$. As mentioned above, it is the complement of the recognition failure measure: $P(\bar{R}n/Rc) = 1 - [P(Rn,Rc)/P(Rc)]$. We will describe the results of interest in this article in terms of this measure of recognition success of recallable words. Theoretical implications of the data reported in this fashion, of course, are not changed in any way. Thus, for instance, the traditional conceptualization of the relation between recall and recognition (e.g., Bahrick, 1970; Kintsch, 1970) which implies that the conditional probability of recognition failure, given recall, should be zero also implies that the conditional probability of recognition success, given recall, should be 1.0.

Inspection of various data of the relation between recall and recognition, reported in a number of experiments under the present paradigm, suggested that recognition failure, and, hence, recognition success of recallable words, is correlated with recognition of all words. We summarize the results of a systematic correlational analysis here. The source of the data for the analysis is the set of the following experiments: (a) Tulving and Thomson (1973, Experiments 1 and 2), (b) Tulving (1974, Experiments 1 and 2), (c) Watkins and Tulving (1975, Experiments 1, 3, 4, 5, and 6), (d) Reder et al. (1974, Experiment 2)¹, (e) Wiseman and Tulving (in press), and (f) Postman (in press, Experiment 2). The data we describe entail free-choice recognition tests only.

Each separate experimental condition in a given experiment provided one bivariate data point, defined by two coordinates. One of the coordinates was the overall level of recognition (the observed "hit" rate), $P(Rn)$, and the other was the recognition success of recallable words, $P(Rn/Rc)$, with the observations pooled over all subjects and all items in a given experimental condition. In experiments in which the data were analyzed and reported separately with respect to different independent variables (e.g., Watkins & Tulving, 1975, Experiment 5), data points for the present purposes were calculated for all conditions, even though this means that not all data points are independent in such experiments. Also, for the Reder et al. (1974) experiment, we analyzed the data separately for the first and second lists learned by subjects.

To illustrate our procedure, consider Experiment 3 in Watkins and Tulving (1975). In that experiment, there were two experimental conditions, differing only with respect to the nature of the recognition test. In one, recognition tests were generated by subjects in a free-association task (the procedure is described in detail in Tulving & Thomson, 1973); in the other, the recognition tests given to the subjects were prepared in advance of the experimental session by the experimenter. The data from the former condition are given in Row 2 in Table 5 in Watkins and Tulving (1975) and are reproduced here in Table 2 in the form of a fourfold contingency table. (Minor discrepancies are due to rounding of figures).

The single bivariate data point derived from Table 2 has the coordinates of .32 as the measure of recognition of all words, or the overall "hit" rate, and .41 (i.e., .26/.63) as the measure of recognition of recallable words, or the conditional probability of recognition given recall. This data point has been plotted, along with the same data points from all the other experiments listed above, in Figure 1. The abscissa in the graph in Figure 1 represents the unconditional probability of recognition, and the ordinate represents the conditional probability of recognition, given recall.

There are a total of 40 data points plotted in Figure 1 from 40 different experimental conditions in 12 different experiments. The data points seem to describe a reasonably systematic function between probability of recognition of all words and probability of recognition

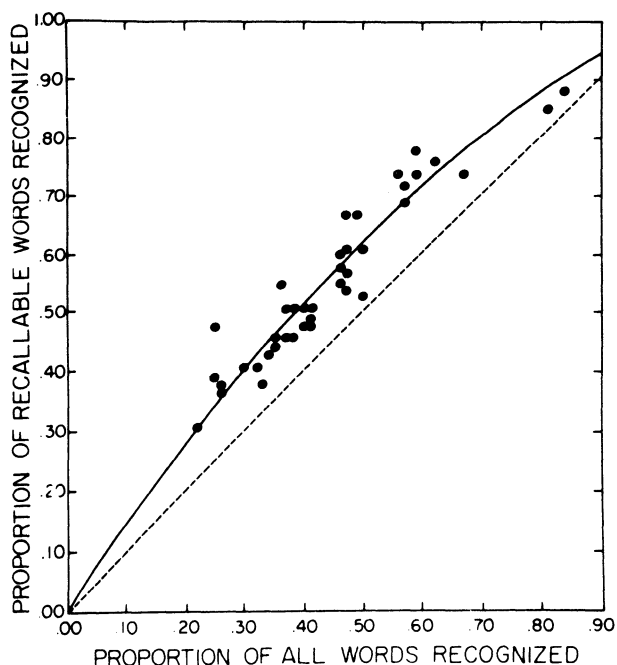


Figure 1. Relation between overall recognition and recognition of recallable words. (Data combined from 12 different experiments).

of recallable words. A mathematical equation that fits the data quite well is given by:

$$P(R_n/R_c) = P(R_n) + c[P(R_n) - P(R_n)^2] \quad (1)$$

The solid line drawn through the data points in Figure 1 corresponds to Equation (1), with the value of the single constant c set equal to .50, according to an estimate provided by a least-squares solution of Equation (1). The product-moment correlation between the actual data points and those predicted by the function is .95; the standard error of estimate is .028.

The data in Figure 1 indicate that, at least in the experiments under consideration here, recognition success of recallable words is quite well predicted by the overall recognition hit rate. The experiments varied considerably from each other in specific procedural details. The close fit between the data points and the function described by Equation (1) thus suggests that the relation between recognition and recognition failure is independent of these details.

We have examined data from a number of other, as yet unpublished, experiments that conform to the general paradigm. So far, we have not found any data describing the relation between recognition and recognition of recallable words that appear to deviate greatly from the function shown in Figure 1. When similar analyses are carried out on data from forced-choice recognition tests, no striking deviation from the function is observed; if anything, the data points move closer to the broken straight line that goes through the origin and has a slope of +1.0. This latter function represents a state of complete stochastic independence between recognition and recall and, hence, complete stochastic independence between recognition of recallable words and recognition of all words. We have also analyzed data, in the manner we have just described, for postselected groups of subjects, as well as for postselected groups of words, in several individual experiments. The assignment of subjects or words to different groups in these analyses was based on their overall level of recognition. The function relating recognition of recallable words to recognition of all words for these groups of subjects and groups of words, graded by overall hit rates, seems to be virtually identical to that shown in Figure 1 for different experimental conditions and different experiments.

It may be worth noting that the data points in Figure 1 seem to constitute a continuous function, regardless of whether recall was higher or lower than recognition in any particular experiment or experimental condition. The data points plotted in Figure 1, as we have indicated, include those reported by Reder et al. (1974), as well as those described by Postman (in press, Experiment 2). In Postman's experiment, there were four experimental conditions for which the fourfold contingency data were reported. Since recall in each of

the four conditions was scored by both lenient and stringent criteria, Postman's experiment provided eight data points to the collection shown in Figure 1. In each of these eight conditions, recall was *lower* than recognition; nevertheless, the eight data points from the experiment seem to be indistinguishable from the rest in their adherence to the general function described.

As we mentioned earlier, the traditional theory of the relation between recognition and recall claims that the probability of recognition failure of recallable words is zero and that the probability of recognition success of recallable words is unity. If this theory was correct, then all the data points in Figure 1 would lie, perhaps with some unsystematic deviation, on a line parallel to the abscissa at the ordinate of 1.00. This is clearly not the case. No extant theory has yet been proposed that claims that recall and recognition are completely independent, in the sense that the probability of recognition of a recallable item is identical with the probability of recognition of a nonrecallable item. If such a hypothetical theory were true, then all the data in Figure 1 would lie along the broken straight line. This is also not true of the actual data. Thus, while the data agree with neither the traditional nor the independence theory, they seem to be closer to the latter, encouraging the inference that recall and recognition in the paradigm under scrutiny are at least to some extent independent.

Recognition here refers to a situation in which the specific retrieval cue is a nominal copy of the target item, or a copy cue, while recall refers to a situation in which retrieval of the TBR item takes place in the presence of the list cue. Applying the encoding specificity principle (Tulving & Thomson, 1973) to the quasi-independence of recognition and recall, we can conclude that the retrieval information contained in the copy cue is to a considerable extent independent of the retrieval information contained in the list cue. This inference, however, is based on the assumption that the stored information about TBR items is the same at the time of the recall test as it is at the time of the earlier recognition test. Since the recognition test may change the stored information, this assumption may not be valid (cf., Tulving & Watkins, in press).

Full implications of what appears to be an orderly relation between recognition of all words and recognition failure of recallable words in the paradigm with which we are concerned remain to be worked out and explored. But the existing data we have summarized suggest that recognition failure of recallable words is a general phenomenon that invariably occurs whenever overall level of recognition is less than perfect. Many variables may affect recognition of studied items and may thus determine whether recall or recognition is higher in any particular situation. But these variables apparently do not affect the relation between recognition and recognition failure of recallable words. The explanation of the magnitude of the phenomenon

of recognition failure of recallable words can thus be sought, at the empirical level, in variables that determine overall recognition. At the theoretical level, the understanding of the phenomenon seems to require the assumption of rather specific encoding of TBR items presented as members of study pairs.

REFERENCES

- Anderson, J. R., & Bower, G. H. A propositional theory of recognition memory. *Memory & Cognition*, 1974, 2, 406-412.
- Bahrick, H. P. A two-phase model for prompted recall. *Psychological Review*, 1970, 77, 215-222.
- Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- Kintsch, W. *The representation of meaning in memory*. Hillsdale, N.J: Erlbaum, 1974.
- Postman, L. Tests of the generality of the principle of encoding specificity. *Memory & Cognition*, in press.
- Reder, L. M., Anderson, J. R., & Bjork, R. A. A semantic interpretation of encoding specificity. *Journal of Experimental Psychology*, 1974, 102, 648-656.
- Tulving, E. Recall and recognition of semantically encoded words. *Journal of Experimental Psychology*, 1974, 102, 778-787.
- Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.
- Tulving, E., & Watkins, M. J. Structure of memory traces. *Psychological Review*, 1975, 82, in press.
- Watkins, M. J., & Tulving, E. Episodic memory: When recognition fails. *Journal of Experimental Psychology: General*, 1975, 104, 5-29.
- Wiseman, S., & Tulving, E. A test of confusion theory of encoding specificity. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, in press.

NOTE

1. We are grateful to Lynne Reder for making her data available for analysis.

(Received for publication April 21, 1975.)