JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR, 8, 501–506 (1969)

Organization as a Determinant of Part-to-Whole Transfer in Free Recall

GORDON H. BOWER[1] AND ALAN M. LESGOLD

Stanford University, Stanford, California 94305

Tulving (1966) reported part-to-whole negative transfer in free recall on unrelated words, suggesting that the subjective organization induced by the part list persevered and was nonoptimal for the whole list. The present study first replicated Tulving's result with single-word presentations, finding too that *S*s clustered their whole-list recall according to old vs. new items (suggesting perseverance of prior organizations). Additional conditions were run in which the list-words were explicitly organized (grouped) for *S*, with the (forced) whole-list organization being compatible with the (forced) part-list organization. In these conditions with forced compatible organizations, part-to-whole transfer was highly positive compared to relevant control *S*s.

This experiment further investigates a most exceptional finding in free recall reported by Tulving (1966). Tulving showed that, contrary to common sense, prior training at recalling a part of a list had a *negative* or detrimental influence upon *S*'s ability to learn and recall the whole list subsequently. In Tulving's experiments, *S*s were pretrained either on an irrelevant control list or on a part list before being shifted to the whole list. Free recall of the whole list by control *S*s began at a lower level but increased more rapidly over trials, overtaking and surpassing recall of those *S*s pretrained with a part of the whole list.

This counterintuitive result shows how inappropriate is the view that free recall of a word depends upon some measure of its *independent* strength as calculated from prior frequency and recency factors. The result points to the role of interword dependencies in free recall; these dependencies are recognized in Tulving's concept of *S*-units or subjective groupings of the list words. The bases for aggregation of "unrelated" words may be multiple and idiosyncratic—similar categories of any kind, similar associations, homophonic or formal similarity or whatever—but numerous researches (cf. Tulving, 1968) leave little doubt that such *S*-units appear in multitrial free recall.

The explanation of part-whole negative transfer in terms of *S*-units goes somewhat as follows: In learning a part of the list, *S* develops particular groupings or *S*-units which he maintains when he is transferred to the whole list; but these *S*-units formed from the subset of the items are not the optimal ones for organizing the whole list. Rather than abandoning his former organization when confronted with the whole list, *S* instead persists in trying to force the new list into his old organization and this retards his discovering the optimal organization for the whole list. To construct a transparent example, in part learning, *S* may categorize *diamond* and *ruby* together in a gem category, but later find that in the whole list, gem is not a very large category and it would have been better to have separated *diamond* into a sports category and *ruby* into a color category, since sports and colors may be more inclusive categories for the whole list. This kind of account has testable implications about transfer in relation to the compatibility vs. incompatibility of organization of the part and whole lists.

[1] The research reported in this article was supported by a grant, MH-13950, to the senior author from the National Institute of Mental Health.

Example of compatible organizations might be (*a*) the same taxonomic categories are obviously present in the part and whole list (e.g., the whole list consists of a few more instances of each category) or (*b*) *S* is required to do serial recall, and the whole list is to be recalled by *S* chaining the new words onto the front or rear end of the part-list series he has formerly learned (see Mandler, 1967). Both these procedures should produce positive transfer from part-to-whole learning. However, the former method uses categorized word lists and the latter serial recall, thus differing in a critical detail from Tulving's demonstration which involved free recall of unrelated word lists.

The following experiment uses a method which we hoped would produce positive transfer in part-to-whole learning with free recall of unrelated words. The basic method is for *E* to impose arbitrary groupings upon the list words: Two or four concrete nouns are presented at a time and *S* is told to visualize an imaginary scene in which these objects are interacting in some vivid way. Prior research in free recall (Bower, Lesgold, & Tieman, 1969) has shown that this input method gives *E* virtually complete control over *S*'s low-level recall units. Subjects' recall is clustered completely by *E*'s input groupings. With this method for controlling *S*'s recall units, we should be able to produce positive part-to-whole transfer by arranging for the whole-list groupings to be compatible with the part-list groupings learned previously. We attempted to do this in two different ways in the following experiment. As an additional part of the study, Tulving's (1966) design was replicated with ungrouped (single word) input to assure ourselves that his result could be reproduced in our laboratory.

Method

Design. Half of the experiment involved a replication of Tulving's (1966) experiment. Let A, B, and C denote sets of 16 unrelated concrete nouns, and AB denote the 32-word list composed of sets A and B. Subjects first had four free recall trials, either with list A or with list C, then had six trials on the AB list. Words were presented singly for 2 sec., with order scrambled each trial. Comparison of recall by *S*s in the A-AB and C-AB conditions should replicate Tulving's result of negative part-to-whole transfer. The other half of the experiment was similar but it employed grouping methods: The 16 words in the A or C list were presented as eight pairs of words with imagery instructions, with the same pairings maintained over the four pretraining trials. The 32-word A-B list which followed for six trials was composed of eight quartets of words, the same quartets were composed in two different ways: (*a*) A former pair of A words was grouped with a new pair of B words, which condition will be abbreviated as (A)-(AB) where parentheses denote groups and each letter denotes two words, or (*b*) four quartets were composed by pairing the former eight A-pairs, and four by division of the 16 B words into quartets, this condition to be abbreviated as (A)-(AA)/(BB). The control *S*s who first learned eight C-pairs were transferred half to the (AA)/(BB) list and half to the (AB) list. For Control *S*s, these two lists are functionally equivalent, involving eight quartets of new words for them. Relative to these controls, we expected positive transfer for *S*s in groups (A)-(AB) and (A)-(AA)/(BB).

Procedure. Each *S* served in both parts of the experiment, doing one task with single-word presentations and one with grouped-word presentations. Different A, B, and C lists (unrelated concrete nouns) were used for each task, but word lists and order of occurrence of the tasks were counterbalanced over *S*s. The *S*s were run individually and all received mental imagery instructions at the start, even those who began with the single-word task. These instructions were repeated before the second task. The lists were presented on flash cards, at a rate of 2 sec. per word on the card. The *S* wrote his recall, having only 2 sec. per word of the list (total recall time was 32 sec. for A and C lists, 64 sec. for AB lists). After four input–output trials on list A or C, *S* had six trials on the appropriate AB list. The order of items in each list (singles, doublets, or quartets) was shuffled for each input trial. A two-minute rest period was interposed between the first and second tasks.

The *S*s were 24 undergraduates fulfilling a service requirement for an introductory psychology course. There were 12 males and 12 females. For the single-word task replicating Tulving's experiment, six females and six males were assigned to the A-AB condition and the remainder to the C-AB condition. For the grouped-work task, four males and four females were assigned to the (A)-(AB) condition, the same number to the (A)-(AA)/(BB) condition, and two males and two females to each of the control conditions, (C)-(AB) or (C)-(AA)/(BB).

Results

Recall scores did not vary significantly either with word list or with first vs. second order of tasks within the day. Therefore, data will be pooled over these incidental variables for the main comparisons of interest.

Single-Word Presentations. This part of the experiment was an attempt to replicate Tulving's result which was obtained with single-word presentations, which presumably permits *S* to impose his own subjective organization upon the materials. The main results of our experiment are shown in Figure 1, and the curves are almost identical

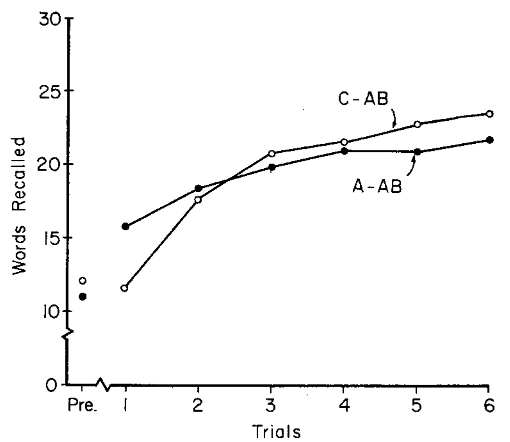

FIG. 1. Mean learning curves for the single-word presentation conditions.

to those reported by Tulving. The C-AB group begins at a lower recall level on the AB list, but by Trial 3 their recall overtakes and surpasses that of the A-AB group. The difference in recall on the last four trials is neither large nor statistically significant, $F(1, 22) = 1.66$, $p > .10$. What does differ between the groups is the slopes of the learning curves. A straight line was fit to the six recall scores of each *S*, as Tulving had done. The mean slope coefficient was .543 for A-AB *S*s and .864 for C-AB *S*s. The difference here is quite reliable, $F(1, 22) = 12.12$, $p < .01$.

A straight-line fit may give a spurious slope advantage to the C-AB *S*s because they have a lower value on Trial 1. To adjust in another manner for the possible effect of starting value upon learning rate, we fitted the exponential curve of statistical learning theory to the six recalls of each *S*. The equation is

$$p_n = 1 - (1 - p_1)(1 - \theta)^{n-1}, \qquad (1)$$

where p_n is the proportion of the AB list recalled on Trial n and θ is the learning rate. The average θ value was .126 for A-AB *S*s and .228 for C-AB *S*s. These differ significantly, $F(1, 22) = 11.00$, $p < .01$. Thus, the difference in learning rates remains when the influence of different initial recall values (p_1's) is removed by this equation.

The slower learning rate for the A-AB *S*s was in large measure attributable to the very slow growth in recall of their old A words, whereas recall of the new B words increased at a rate only slightly slower than that of the C-AB *S*s. Mean θ values computed for the A words during AB learning was .08; for the B words it was .16, to be compared with a mean θ of .23 for the C-AB *S*s.

Clustering by A's and B's. The next question we asked was whether the A-AB *S*s tended to distinguish and keep separate in recall their old A words and the new B words, even though these were thoroughly intermixed in the AB input list. So we computed clustering scores with respect to the A set and B set for all *S*s. This score is essentially a count of the number of AA and BB consecutive pairs in the recall protocols. The C-AB *S*s should cluster the A and B words only at the chance level, whereas the A-AB *S*s should cluster beyond chance.

Some complications arise for a direct comparison, however, since the A-AB *S*s began by recalling more A's than B's and this would inflate their chance clustering score. With two categories, however, the null hypothesis of no clustering of the n_1 A's and n_2 B's which *S* recalls is just the complement of the null hypothesis of the Wald and Wolfowitz (1940) runs test, since the number of repetitions plus runs equal the number of words recalled. Therefore, the values of n_1, n_2 and the number of runs of A's and B's for each protocol were entered into tables from Swed and Eisenhart (1943) to obtain the probability of that protocol on the null hypothesis of no clustering by A's and B's.

These clustering probabilities clearly showed the anticipated effect. Since $-2 \log p_i$ is chi-square distributed, these transformed probabilities were pooled over trials and *S*s within each condition. For the A-AB *S*s, the total chi-square was 228 ($df = 144$), giving $p < .0001$ on the null hypothesis; for the C-AB *S*s, the chi-square was 101 ($df = 144$), giving $p > .75$. This means that A-AB *S*s cluster significantly in recall according to A's and B's, whereas C-AB *S*s do not (as they should not). These A/B cluster scores showed no reliable trends over the six trials, but it was observed that male *S*s clustered considerably more than did female *S*s. Average recall of the AB list was also reliably less for the male than the female *S*s in condition A-AB.

Grouped Input Lists. In these conditions, *S*s first learned doublet groupings on list A or C, then quartet groupings on list AB, either in the form (AB) or (AA)/(BB). The recall results in this portion of the experiment are shown in Figure 2. Here we find substantial positive transfer going from (A) to either (AB) or to (AA)/(BB) relative to the control

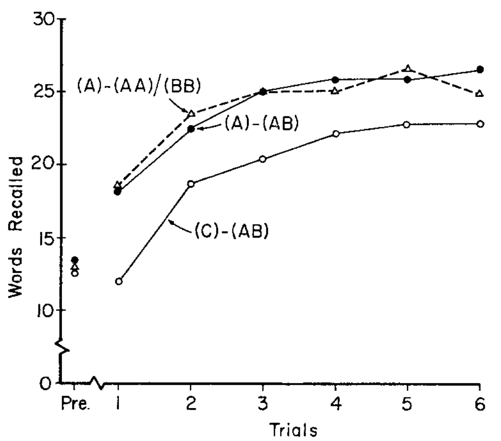

FIG. 2. Mean learning curves for the multiple-word presentation conditions.

(C)-(AB). The contrast between mean recall of AB of the control and the other two groups is statistically reliable, $F(1, 21) = 5.35, p < .05$, but the (A)-(AB) and (A)-(AA)/(BB) groups do not differ.

The advantage for the part-learning groups in this case results mainly from their starting at a higher recall level on Trial 1 with the AB list, i.e., in the value of the p_1 parameter of Equation 1. By contrast, the overall learning rates for the AB list were similar: $\theta = .21$ for (A)-(AA)/(BB), $\theta = .23$ for (A)-(AB), and $\theta = .24$ for (C)-(AB). The advantage in p_1 appeared for both old A words and new B words in the part-learning groups. Pooling the two part-learning groups, mean recalls of A and B words over the six trials were 12.4 and 11.6, respectively, out of 16; for the (C)-(AB) control condition, mean recall for A and B words was 8.6 and 8.9, respectively. The separate recall of A words and of B words was significantly higher for the part-learners than for the control *S*s, ($t(22) = 3.5$, $p < .001$ for A's and $t = 2.3$, $p < .02$ for B's). With these grouping procedures, therefore, part-learning of the A words leads to better recall of both A and B words when gauged relative to the control *S*s.

The input groupings can be shown to determine *S*'s recall groupings. During first learning of (A) or (C), if the two words of an input group were recalled, the probability was .93 that they were recalled adjacent to one another. By any measure of recall consistency, either Tulving's (1962) S. O. measure or Bousfield, Puff and Cowan's (1964) measure of intertrial repetition, *S*s getting the paired (A) or (C) lists were much more organized in their recall than those getting the singly presented A or C lists. The same holds comparing AB recall of the C-AB and (C)-(AB) conditions. Despite stronger clustering, however, the overall level of recall was similar for these paired- vs. singlet-list comparisons. Thus, strong organization at this level does not necessarily imply better overall recall, contrary to what Tulving seems to imply.

Recall of the grouped AB list was predominately clustered according to the presented quartets. A modified repetition ratio score was used to measure clustering, viz.,

$$\text{MRR} = \frac{\sum_i n_{ii}}{\sum_i (n_i - 1)} \quad \text{for all } n_i > 1,$$

where n_i is the number of words recalled from quartet or group i, and n_{ii} is the number of consecutive recalls of words in group i (max $n_{ii} = n_i - 1$). The MRR score is 1 for perfect clustering. For the first two trials on the AB list, the average MRR score was .86 for the (C)-(AB) *S*s, .93 for the (A)-(AB) *S*s, and .91 for the (A)-(AA)/(BB) *S*s. By the last two trials, the mean MRR scores had changed only slightly to .81, .95, and .98, respectively, for the three conditions. Insofar as output clustering followed input groupings, there was no residual clustering of recall for the (A)-(AB) *S*s in terms of their old A words vs. their new B words. For the (A)-(AA)/(BB) *S*s receiving four A quartets (composed of old A pairs) and four new B quartets, MRR scores calculated for only the A quartets were identical to those calculated for only the B quartets. The mean recall of the B quartets was lower than the A quartets initially, but the degree of clustering when the B groups were recalled was very high.

Discussion

We have replicated Tulving's result of negative part-to-whole transfer when *S* is presented words singly and permitted to develop his own *S*-units. Apparently, these persist into the whole-list task and retard *S*'s finding more optimal groupings of the whole list of words. In partial support of this persistence, the results showed marked clustering of recall of A-AB *S*s in terms of their old A words vs. their new B words, whereas C-AB *S*s, of course, make no such distinctions in their recall. The pernicious persistence of initial categorizations or hypotheses in the face of conflicting evidence is a well-documented phenomenon in the areas of problem solving (Wyatt & Campbell, 1951), impression formation (Anderson, 1965), clinical diagnosis (Rubin & Shontz, 1960), and perceptual recognition of noisy signals or blurred pictures (Blake & Vanderplas, 1950; Davison, 1964). The part-to-whole negative transfer in free recall may simply be another manifestation of this tendency to persist with an initial description or set of categories, which in the present case are doing a fair but nonoptimal job of classifying or organizing the whole list of words.

When the groupings of the list words are given directly to *S* and these are compatible for the A and AB lists, then part-learning leads only to positive transfer in learning the whole list. The advantage appears not only in better recall of the pretrained words, but also in better recall of the new words relative to controls pretrained on irrelevant words. For the (A)-(AB) condition, the superior recall of the new B words is explicable in terms of the higher availability of the A words and the associations established at input between the A and B words.

It must be remembered that the time permitted for written recall in this task was very short (2 sec. per word), as in Tulving's earlier experiment with oral recall. Many *S*s volunteered the information that they knew many more words but they simply were not given enough time to write down all they knew. Despite *S*s' subjective feeling that they were writing as fast as they could, the (A)-(AB) and (A)-(AA)/(BB) *S*s were still recalling reliably more words than *S*s in the other three conditions. However, it is possible that the main difference among these several conditions is in the *rate* of recall rather than in the total number of words that could be recalled over a prolonged test period. This would not be an unprecedented outcome, since Postman, Egan, and Davis (1948) have shown that rate of recall can be a more sensitive measure of interference than is total recall. Earlier results (e.g., Pollio, 1968) on sequential latencies of recalled words show that words in an *S*-unit are recalled in a rapid burst with short inter-response times, with longer pauses between adjacent *S*-units. If the (A)-(AB) and (A)-(AA)/(BB) conditions create larger and better integrated *S*-units than in the A-AB conditions, then the advantage of the former conditions in rate of recall would be explained by these observations on accessing times within *S*-units.

References

ANDERSON, N. H. Primacy effects in personality impression formation using a generalized order effect paradigm. *Journal of Personality and Social Psychology*, 1965, **2**, 1–9.

BLAKE, R. R., & VANDERPLAS, J. M. The effect of pre-recognition hypotheses on veridical recognition thresholds in auditory perception. *Journal of Personality*, 1950, **19**, 95–115.

BOUSFIELD, W. A., PUFF, C. R., & COWAN, T. M. The development of constancies in sequential organization during repeated free recall. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 449–459.

BOWER, G. H., LESGOLD, A. M., & TIEMAN, D. G. Grouping operations in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 481–493.

DAVISON, G. C. The negative effects of early exposure to suboptimal visual stimuli. *Journal of Personality*, 1964, **32**, 278–295.

MANDLER, G. Traditional serial learning: A waste of time? Paper read at meetings of Psychonomic Society. St. Louis, 1967.

POLLIO, H. Associative structure and verbal behavior. In T. R. Dixon and D. L. Horton (Eds.), *Verbal Behavior and General Behavior Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968. Pp. 37–66.

POSTMAN, L., EGAN, J. O., & DAVIS, J. Rate of recall as a measure of learning: I. The effects of retroactive inhibition. *Journal of Experimental Psychology*, 1948, **38**, 535–546.

RUBIN, M. & SHONTZ, F. C. Diagnostic prototypes and diagnostic processes of clinical psychologists. *Journal of Consulting Psychology*, 1960, **24**, 234–239.

SWED, F. S., & EISENHART, C. Tables for testing randomness of grouping in a sequence of alternatives. *Annals of Mathematical Statistics*, 1943, **14**, 66–87.

TULVING, E. Subjective organization in free recall of "unrelated" words. *Psychological Review*, 1962, **69**, 344–354.

TULVING, E. Subjective organization and effects of repetition in multitrial free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 193–197.

TULVING, E. The effects of presentation and recall of material in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 178–184.

TULVING, E. Theoretical issues in free recall. In T. R. Dixon and D. L. Horton (Eds.), *Verbal Behavior and General Behavior Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968. Pp. 2–36.

WALD, A., & WOLFOWITZ, J. On a test whether two samples are from the same population. *Annals of Mathematical Statistics*, 1940, **11**, 147–162.

WYATT, D. F., & CAMPBELL, D. T. On the liability of stereotype or hypothesis. *Journal of Abnormal and Social Psychology*, 1951, **46**, 496–500.

(Received December 27, 1968)