EFFECTS OF CONSTRAINED RECALL TRAINING ON CHILDREN'S PERFORMANCE IN A VERBAL MEMORY TASK

SYLVIA SCRIBNER and MICHAEL COLE

Rockefeller University

SCRIBNER, SYLVIA, and COLE, MICHAEL. Effects of Constrained Recall Training on Children's Performance in a Verbal Memory Task. CHILD DEVELOPMENT, 1972, 43, 845–857. *Second, fourth, and sixth graders were trained under conditions of Constrained and Cued recall on a list of randomly ordered nouns comprising 4 categories of things. Constrained groups were required to recall words by category on the first 3 trials, while Cued groups were reminded of the categories after list presentation but were permitted to recall in any order. Effects of these 2 practice methods were tested on a fourth trial with the original list and a transfer trial with a new list under standard free-recall conditions. Children with previous Constrained practice recalled more words than Cued children on the free trial with the old material, but did not maintain this advantage on the transfer trial when new material was introduced. On both fourth and transfer trials, children having Constrained practice showed greater categorical clustering than Cued children and continued to manifest errors characteristic of forced category recall.*

In studies of free recall, it has repeatedly been demonstrated that educated American adults commonly reorganize stimulus material in some

---

This work was supported by grant no. GM16735 to W. K. Estes and by grant no. OEG-0-71-1965 from the U.S. Office of Education to Michael Cole. We are indebted to Janet Throop for her help in developing procedures and analyzing the data, and to Vicky Carpenter, Elizabeth Dromey, Nancy Nylen, and Bonnie Thompson of Hampshire College for their careful work in conducting the experimental sessions and their thoughtful observation. The cooperation of the teachers and students of PS 40 and PS 6 of Manhattan were essential to the conduct of this research. Reprint requests should be sent to Sylvia Scribner, Rockefeller University, New York, New York 10021.

[*Child Development*, 1972, 43, 845–857.

systematic manner, and that such organization is generally associated with a superior level of recall (Bousfield 1953; Bower 1970; Cohen 1963; Tulving 1962). A standard finding is that when a randomly ordered list of words naming members of several categories of things (e.g., animals, food, furniture) is recalled, items from one category tend to appear together.

Such transformation of the input list according to semantic grouping occurs in laboratory learning studies utilizing diverse stimulus materials and under a wide variety of experimental conditions. Several recent lines of investigation, however, have shown that categorical reorganization of material is not characteristic of the mnemonic performance of all populations. Young children, for example, often display little semantic organization in their recall. A number of developmental studies involving American elementary school children have shown that categorical clustering increases with age (Bousfield, Esterson, & Whitmarsh 1958; Cole, Frankel, & Sharp 1971; Mandler & Stephens 1967; Moely, Olson, Halwes, & Flavell 1969; Neimark, Slotnick, & Ulrich 1972; Vaughan 1968). In studies in a nonindustrialized West African society (Cole, Gay, Glick, & Sharp 1971), the extent of category clustering was found to be more closely associated with participation in a Western-style education program than with age per se. Nonliterate adults failed to reorganize material to be remembered according to category membership under standard free-recall instructions, but high-school educated adults showed performance typical of similarly educated groups in the United States. Amount recalled by the educated adults was considerably above that of their nonschooled counterparts. These findings suggested that utilization of semantic structure as a mnemonic aid might usefully be considered a specific skill learned through specialized experiences provided in certain cultures. If this is indeed the case, investigation of the factors regulating the acquisition and employment of such a skill would be of both practical and theoretical interest—practical in its implications for educational programs, of theoretical importance in its potential contribution to the identification of processes underlying the learning of generalized "cognitive skills."

With these experimental interests, Cole, Gay, Glick, and Sharp (1971) explored features of the experimental situation that might induce the use of semantic structure as a mnemonic aid. They found they could secure the greatest effect on recall organization by requiring the subject to recall items by category as the experimenter named each in turn (hereafter referred to as constrained recall). West African adults who were given four trials of constrained recall on a 20-word, randomly arranged, categorized list recalled more words than comparison subjects who were simply reminded of the categories in the list at time of recall and were permitted to recall in any order they chose. On a fifth trial when recall was free for both groups, subjects with constrained recall practice continued to recall more words than the controls and showed greater categorical organization in their output.

The present study was designed to test the generality of the skills acquired during practice in constrained recall: Does such practice lead to superior recall on a new set of materials, or are its effects restricted to the original learning material? Because of the evident educational implications of this question, the investigation was conducted with elementary school children and as a more rigorous test of the effects of constrained recall practice, a child population of comparatively high scholastic performance was selected.

METHOD

Subjects.—Two New York City public elementary schools serving primarily children in families of middle or upper socioeconomic status provided the subject population. Both schools rated high in reading achievement and scholastic performance. Children of both sexes in the second-, fourth-, and sixth-grade classes were asked to secure parent consent for their participation in the study. All children on a given grade level who received parent approval became a population pool from which a subject group of 32 was randomly selected; remaining names were randomized and constituted a replacement pool. Subjects were randomly assigned within grades to the experimental groups. The study was conducted in the middle of the school year. Mean CA's were 7.11, 9.03, and 11.4 for children in the second, fourth, and sixth grades, respectively; 95 boys and 97 girls were included.

Design.—The experimental design included two schools, three grade levels, and two experimental conditions. The two experimental conditions differed in the method used by the experimenter to elicit recall and are designated as the Cue condition and the Constrained condition. Procedures employed in the Cue condition resembled, but were not identical with, the Cued recall used by Tulving and Pearlstone (1966) in that subjects were given category names as cues when the words were presented and again at the time of recall. In the Constrained condition, subjects were not only given category names at presentation and recall, but were required to recall words belonging to a given category as it was named by the experimenter.

Two experimental conditions combined with the three grade levels and two schools resulted in 12 experimental groups with 16 children in each: a total of 192 subjects.

Lists.—The stimulus material consisted of two 20-word lists of common nouns, each composed of four categories with five member items. List A, including food, clothing, kitchen utensils, and tools, had been extensively used in prior studies (Cole, Frankel, & Sharp 1971). List B was prepared from the Battig and Montague category norms (Battig & Montague 1969); an effort was made to select nouns having Thorndike-Lorge general and juvenile frequency counts similar to words in List A. Categories were people's names, sports, furniture, and animals (see table 1). Ten different orders

TABLE 1

STIMULUS WORDS AND THEIR ASSOCIATED THORNDIKE-LORGE (1944) FREQUENCIES

STIMULUS WORD	FREQUENCY	
	General	Juvenile
List A:		
Glass	AA	700
Cup	AA	700
Plate	A	329
Pan	A	290
Pot	47	205
Potatoes	A	330
Orange	A	291
Lemon	27	26
Banana	13	38
Onion	25	25
Saw	AA	M
File	43	200
Drill	21	140
Hammer	34	157
Ax	47	214
Hat	AA	700
Shoes	AA	534
Pants	6	35
Socks	12	14
Shirt	47	214
List B:		
Bill	AA	700
Mary	AA	430
Peter	A	300
Alice	34	72
Howard	19	28
Fishing	AA	700
Swimming	A	410
Football	26	92
Tennis	18	3
Hockey	3	5
Table	AA	M
Desk	A	316
Lamp	A	296
Couch	28	127
Television	1	...
Dog	AA	700
Cow	A	358
Lion	A	339
Elephant	35	287
Fox	25	116

were prepared for the words on each list by selecting the words randomly, subject to the restriction that no two words from any category appear adjacent to each other.

Procedure.—Each child was worked with in one individual session consisting of four successive presentations and recall of one list (original learning), followed by an interlude of 2–3 minutes of conversation, and then one presentation and recall of the second list (transfer).

Presentation was oral and at the rate of approximately 2 seconds per word. Recall was also oral and was recorded in writing. Children were given unlimited time for recall and, after indicating they were "through," received one prompt from the experimenter ("Can you remember any more things"?) and the standard comment, "Very good."

One-half of the subjects received List A as the original list and B as the transfer list, while the other half received the lists in reverse order. The 10 different serial orders for each list were randomized across trials and subjects.

Children in both experimental groups were told the specific categories contained in the lists before each presentation. On trials 1–3 of the first list, children in the Cue condition were reminded of the Category names after list presentation, but were instructed merely to "Tell me all the things you can remember." Children in the Constrained condition were required to recall by category: "Tell me all the food you can remember," "Tell me all the tools you can remember," etc.

On the fourth trial, Cue and Constrained subjects again received the category names as cues before list presentation, but recall was free, with no reminder of categories and no instructions on order of recall ("Tell me all the things you can remember"). The order in which the category names were given as cues was randomized across trials and subjects.

The single trial of the new list (transfer trial) proceeded in the same way as the fourth trial on the old list: both Cue and Constrained groups were given the category names before list presentation, but recall was free. This trial constituted a test of whether children in the Constrained condition would spontaneously categorize new material on which the categories as well as the individual words differed from the original list.

RESULTS

Since no significant differences were found on any performance measure between schools, data from both are combined in the following presentation and discussion of results. Lists were also included as a factor in the statistical analysis of number of words recalled. No list effects or interactions appeared except on the transfer trial, and findings will be discussed without reference to list or list order except in relation to recall performance on the transfer trial.

Amount recalled.—Amount recalled by grade and experimental group is presented in figure 1.

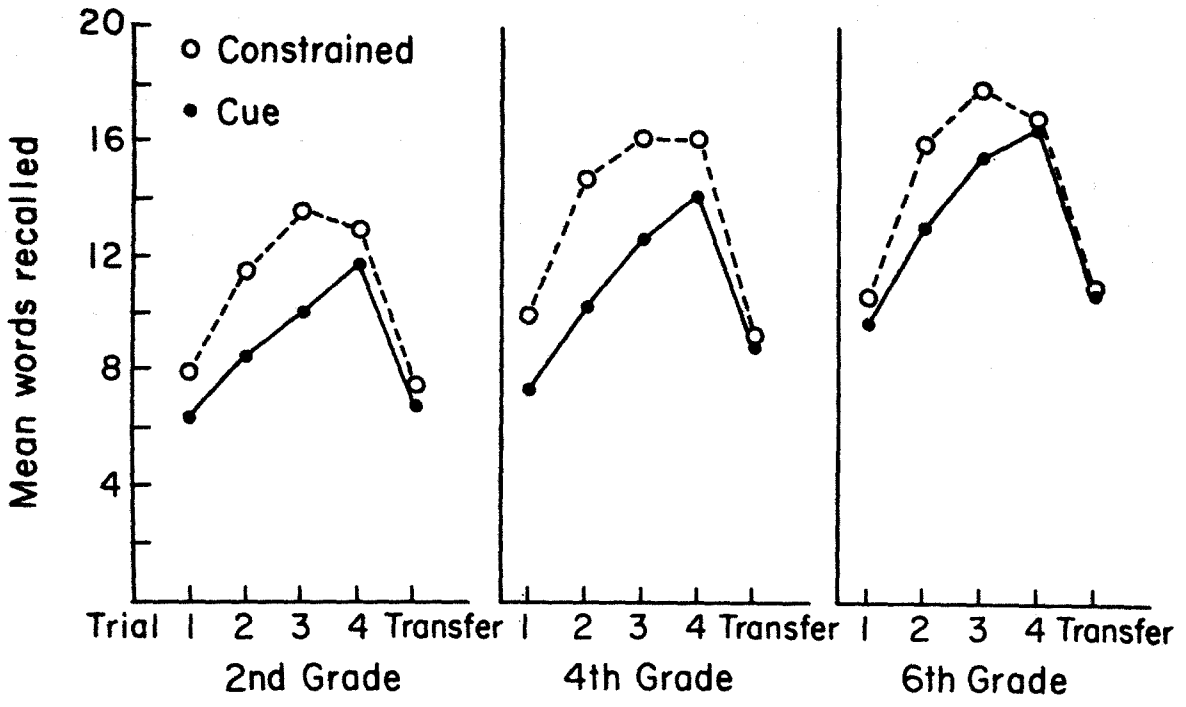

FIG. 1.—Mean number of words correctly recalled per trial, by experimental groups and grades.

The most meaningful interpretation of the data can be achieved by looking separately at the children's performance under conditions of Cued and Constrained recall (trials 1–3) and under free recall (trial 4 and transfer trial). On the first three trials, learning was very rapid, $F(2,336) = 493.06$, $p < .01$; recall improved with age, $F(2,168) = 45.58$, $p < .01$; and was superior under Constrained recall, $F(1,168) = 68.75$, $p < .01$. A condition × trial interaction, $F(2,336) = 12.52$, $p < .01$, reflects the fact that children in the Constrained groups made more substantial gains in amount recalled between the first and second trials than did the Cued groups. The more rapid learning of older children on the early trials is shown in a grade × trial interaction, $F(4,336) = 4.25$, $p < .01$.

There is clear evidence here that Constrained recall, which was first observed to provide a dramatic increase in amount recalled among nonliterate West Africans, yields a similar, although lesser, increase in recall among middle-class urban schoolchildren. Figure 1 shows that at every grade level, more words were recalled by the Constrained than the Cue groups. In fact, this method of eliciting recall eliminated the customary differential between adjacent age levels: second graders in the Constrained condition performed as well as fourth graders in the Cue condition; similarly, fourth graders in the Constrained condition performed as well as sixth graders in the Cue condition.

On the fourth trial (free-recall), children trained under Constrained recall showed some decline in performance but continued to remember significantly more words than children in the Cue condition, $F(1,168) = 7.84$, $p < .01$, demonstrating a persisting effect of the constraining technique with the original learning material.

On the transfer trial, when children were presented with new material, there were no differences between the Cued and Constrained groups in number of words correctly recalled. The level of recall was affected, however, by the list used in the transfer test, $F(1,168) = 7,36$, $p < .01$; on the average, more words were remembered from List A than from List B. List effects did not interact with recall condition in any consistent way, but the effect of list varied with school, $F(1,168)$ for school × list × condition = 4,08, $p < .05$. In one school, both experimental groups recalled less on List B than List A, whereas in the other school, the Cued group did less well on B while the Constrained Group's recall on B was equivalent to that on A.

Organization of recall.—Recall was analyzed in terms of the extent to which output lists were organized according to category membership. The measure of organization used in this analysis was the Z score discussed by Frankel and Cole (1971). A positive Z score indicates that items of one category tend to cluster together on the recall list more than might be expected by chance.

Comparison of Cued and Constrained recall conditions while the constraints were in effect are meaningless because perfect clustering was assured the Constrained subjects. Table 2 shows that on the fourth trial (free recall with the original material) for all grade levels, groups which had been constrained previously showed greater categorization of output than the groups without such practice, $F(1,144) = 51,88$, $p < .01$. Cluster scores of Constrained groups remained higher than those of Cue groups, $F(1,144) =$

TABLE 2

MEAN CLUSTERING SCORE (*Z*) FOR EXPERIMENTAL GROUPS

	EXPERIMENTAL CONDITION	
	Cue	Constrained
Second grade:		
Trial 4	1.060	2.560
Transfer	.473	1.035
Fourth grade:		
Trial 4	1.912	4.224
Transfer	.754	1.455
Sixth grade:		
Trial 4	2.866	4.385
Transfer	1.509	2.602

15.97, $p < .01$, for all grade levels on the transfer trial (free recall on new material) as well.

A further indication of group differences in organization on the fourth and transfer trials is shown in table 3 which presents the frequency of

TABLE 3

FREQUENCY DISTRIBUTION OF RUN-SIZES IN FREE RECALL

	RUN-SIZES			
	2	3	4	5
Fourth trial:				
Second grade, cue	62	20	7	0
Second grade, constrained	55	34	19	2
Fourth grade, cue	57	32	18	0
Fourth grade, constrained	39	26	43	25
Sixth grade, cue	64	29	26	13
Sixth grade, constrained	32	15	30	47
Transfer trial:				
Second grade, cue	29	7	0	0
Second grade, constrained	39	14	4	0
Fourth grade, cue	38	17	3	1
Fourth grade, constrained	47	17	8	1
Sixth grade, cue	43	24	11	0
Sixth grade, constrained	44	30	19	3

runs of intracategory responses of different lengths. Runs of four and five were heavily concentrated in the recall output of the two higher grades. At each grade level, however, a greater number of four- and five-item runs was exhibited by the groups with prior practice in Constrained recall.

The distribution of five-item runs is especially interesting since these represent full category retrieval and provide stronger evidence of the deliberate use of categories as a mnemonic device than incomplete sequences of category members. On the fourth trial, second graders produced only two five-item runs and these were both in the Constrained condition; fourth graders produced 25 complete runs, again entirely in the Constrained condition; sixth graders produced a sizable number of such runs in both conditions, although more than three times as many in the Constrained as in the Cue condition. On the transfer trial, there was a marked reduction of four- and five-word runs, a finding which is to be expected on the first recall of new material. Nevertheless, this trial resulted in the same relationships as the fourth trial: longer runs occurred more frequently in the upper grades and in the Constrained conditions of all grades.

Errors.—Responses which were not a part of the input list (errors) were analyzed to uncover possible modifications of mnemonic activities produced by manipulation of the recall method. Errors were classified into

three categories: (1) repetitions of list words, (2) intrusions of words belonging to the same categories as the list words (categorical intrusions), (3) and intrusions of other words (noncategorical intrusions). Errors falling into the first two classes comprised 92% of the total error. Error was related to output by expressing the number of wrong words given on each trial by each experimental group as a ratio of the number of correct words given on that trial.

Table 4 shows that in every grade Constrained groups began the learning task (trial 1) by making proportionately more errors than the Cue groups, but they concluded it (trial 4) with a lower error rate. This differential relationship between accuracy and practice follows from the fact that each experimental condition exhibited a characteristic form of error. Constrained recall predominantly manifested categorical intrusions and Cued recall manifested repetitions. Repetition and category errors took a different course over learning. Under both Cued and Constrained conditions, category errors initially exceeded repetition errors, but declined over trials (see table 4). This was an absolute decline. Fewer category errors were made as correct output increased. Repetition errors, in contrast, increased over trials both absolutely and proportionately. These changes held for all grades tested and account for the fact that Constrained groups turned in more accurate performances at the end of learning than the Cue groups.

Error performance on the transfer trial is especially suggestive of the influence of recall condition on subject activities during the recall task. The group previously trained on categorized recall (Constrained group) showed the characteristic predominance of category errors on this trial, even though new categories were introduced with the new list. Groups with previous cued recall showed more repetition errors (fourth and sixth grades), or a nearly equal number of repetition and category errors (second grade).

DISCUSSION

Requiring subjects to recall words by category is a facilitatory memory technique for urban middle-class American schoolchildren as well as for the African villagers with whom it was first tested (Cole, Gay, Glick, & Sharp 1971). Improvement in recall produced by constraining techniques is all the more striking in the present study because it was obtained in comparison with cued, rather than free, recall. Tulving and Pearlstone (1966) and other investigators have found that presentation of category names as cues at presentation and recall typically enhances the amount recalled. Constraining recall appears to have "extra-facilitating" effects, over and above those of cueing.

While the superiority of induced recall is clear, the mechanisms by which it yields a higher output are not. The Constrained groups' superior recall on trial 1 must be attributed to enhance *retrieval*, since the Cue and

TABLE 4

TYPE OF ERROR AS A PERCENTAGE OF CORRECT RECALL

	SECOND GRADE				FOURTH GRADE				SIXTH GRADE			
	Cue		Constrained		Cue		Constrained		Cue		Constrained	
TRIAL	R	C	R	C	R	C	R	C	R	C	R	C
1	1.5	6.0	1.6	18.3	1.3	9.4	0.3	15.5	3.0	4.3	0.0	12.9
2	7.0	4.8	1.4	10.1	5.2	2.8	1.9	9.6	5.8	4.1	0.4	4.3
3	13.2	3.1	0.7	8.3	6.2	1.7	0.4	6.4	8.6	2.6	0.9	1.9
4	12.7	2.6	5.8	6.1	10.8	1.3	6.1	2.9	15.0	1.1	5.1	1.7
Transfer	4.6	5.5	3.2	13.7	5.7	2.5	2.4	10.5	6.3	5.1	0.9	6.5

NOTE.—R = repetitions; C = category intrusions.

Constrained groups are not differentially treated until after items are presented on trial 1. On trials 2 and 3, constrained recall may have influenced the manner in which subjects organized the list at the time of input as well as the way they retrieved the items at output. We have no basis in this study for making inferences about the possible effects of constrained recall on input processes, but our findings warrant some observations about effects on retrieval processes.

Tulving and Pearlstone (1966) suggest that recall of a categorized list involves two independent retrieval processes, one concerned with the accessibility of higher-order memory units and the other with the accessibility of items within higher-order units. "Accessibility of higher-order units," according to these investigators, depends on appropriate retrieval cues. Our findings suggest the need for distinguishing between the *accessibility* of higher-order units and their actual *utilization* in retrieval.

This distinction is similar to that made by Mandler (1966), who observed that lack of clustering in free recall might be due either to the subject's failure to discover the specific rule relating list items to one another or to his inability to use the rule adequately, once it is discovered. He considered the second alternative possible, but less likely than the failure of discovery.

In this study, subjects in both experimental conditions were told the rule at the outset and repeatedly on each trial. The categories were made accessible to both groups at time of recall.[1] In the Constrained condition, however, subjects were required to use the accessible categories while in the Cue condition this was optional, and in addition, subjects had to determine the order of utilization.

The failure of many Cue subjects to make full use of categories in a retrieval plan, even with continued practice on the list, is demonstrated by the fact that their cluster scores remained substantially below corresponding scores of subjects in the Constrained condition on all training trials. Observational data substantiate the view that mere provision of information by the experimenter about higher-order units in the list does not assure their functional significance. Some children in the Cue condition asked the experimenter not to repeat the category names at time of recall, and some ignored the cueing altogether, proceeding to give the words they remembered while the experimenter was still reminding them of the categories in

[1] It might be argued that subjects in the Cue condition ignored the information about categories, so that at the time of recall, this information was not available to them. Data on this point are not available from the present study, but in an unpublished study using essentially the same procedures except that no transfer trial was given, and in which the same results were obtained, we questioned subjects following the free trial about the names of the categories in the list. Recall of category names was virtually perfect for both groups.

the list. This accumulated evidence makes it clear that "list structure" and "retrieval structure" are not inseparably linked, although they may appear so in the performance of educated American adults.

The effect of constrained recall training on subsequent free recall is ambiguous. While Constrained subjects were able to maintain a higher level of recall than Cued subjects on the original material, they lost this advantage as soon as new material was introduced. Transfer measured in terms of amount recalled—the crucial criterion in judging memory performance or memory skills—failed to occur. Nonetheless, there are several lines of evidence that the mnemonic activities of Constrained subjects were modified during the training procedure and that these modifications persisted with the new as well as the old material. The greater degree of clustering for constrained groups on both trial 4 and the transfer trial is one piece of evidence. The more frequent appearance of intracategory runs of four and five words among Constrained subjects on both these trials is another. A final indication of the influence of constrained recall training on mnemonic activities is the persistence of characteristic category intrusion errors among trained subjects on the transfer trial.

These performance measures would seem to indicate that the list categories played a more salient role for previously constrained children on both old and new material. Nevertheless, there was a marked drop-off in all measures of organization when the training procedures were terminated, indicating that under these procedures the children did not learn to direct their own retrieval as it was previously directed by the experimenter. Much more detailed and extensive training in category retrieval is clearly required to help children develop and fully use categorization as an effective mnemonic technique.

REFERENCES

Battig, W. F., & Montague, W. E. Category norms for verbal items in 56 categories. *Journal of Experimental Psychology,* 1969, **80** (3, Pt. 2).

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology,* 1953, **49**, 229–240.

Bousfield, W. A.; Esterson, J.; & Whitmarsh, G. A. A study of developmental changes in conceptual and perceptual associative clustering. *Journal of Genetic Psychology,* 1958, **92**, 95–102.

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

Cohen, B. H. An investigation of recoding in free recall. *Journal of Experimental Psychology,* 1963, **65**, 368–376.

Cole, M.; Frankel, F.; & Sharp, D. Free recall learning in children. *Developmental Psychology,* 1971, **4**, 109–123.

Cole, M.; Gay, J.; Glick, J.; & Sharp, D. M. *The cultural context of learning and thinking.* New York: Basic, 1971.

Frankel, F., & Cole, M. Measures of organization in free recall. *Psychological Bulletin,* 1971, **76,** 39–44.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation.* Vol. 1. New York: Academic, 1966.

Mandler, G., & Stephens, D. The development of free and constrained conceptualization and subsequent verbal memory. *Journal of Experimental Child Psychology,* 1967, **5,** 87–93.

Moely, B. E.; Olson, F. A.; Halwes, T. G.; & Flavell, J. H. Production deficiency in young children's clustered recall. *Developmental Psychology,* 1969, **1,** 26–34.

Neimark, E.; Slotnick, N. S.; & Ulrich, T. The development of memorization strategies. *Developmental Psychology,* 1972, **5,** 428–432.

Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words.* New York: Bureau of Publications, Teachers College, Columbia University, 1944.

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

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

Vaughan, M. E. Clustering, age, and incidental learning. *Journal of Experimental Child Psychology,* 1968, **6,** 323–334.