Strategy Training and Semantic Encoding in Mildly Retarded Children

RANDALL W. ENGLE
RICHARD J. NAGLE
University of South Carolina

Two experiments were performed with mildly retarded children. In the first experiment, three groups of mildly retarded pre-adolescents were matched on the basis of their performance on a free recall test. The groups were then given instruction in the use of a semantic encoding strategy, acoustic encoding strategy, or repetitive rehearsal strategy and practice using the memory strategy on two different lists of unblocked categorized items. Performance was uniformly higher for subjects using the semantic encoding strategy on both training lists and on unprompted tests up to seven days later. On a test seven months later, however, there were no differences in recall performance for the three groups until the appropriate strategy was prompted on the second trial; then, the semantic condition showed greater improvement than the other groups. Clustering data were also analyzed and were superior for the semantic condition on all tests except the first trial of the seven month delayed test. In the second experiment an incidental learning task was used with three different orienting questions. Enhanced performance with semantic processing was found in a group of 13-year olds but not in a group of 10-year olds. Implications for theory are discussed.

There is increasing evidence to support the hypothesis that the memory deficit frequently found with mildly retarded individuals is not a result of reduced memory capacity or other structural deficits. This viewpoint argues that their memory deficiencies are a result of either, (1) the failure to use any rehearsal strategy during acquisition, or (2) the utilization of an ineffective and unsophisticated strategy (Ellis, 1970; Brown, 1974; and Belmont & Butterfield, 1969).

A desirable feature of this theory has been the evolution of what Belmont and Butterfield (1977) call the "instructional approach." This approach is directed toward discovering and developing effective training methods for improving memory skills in the retarded. For example, Brown, in a series of articles (Brown, Campione, Bray & Wilcox, 1973; Brown, Campione &

Requests for reprints should be sent to Randall W. Engle, Psychology Department, University of South Carolina, Columbia, South Carolina 29208.

1) We wish to thank Dr. John May and the principals and teachers of Richland County Schools for their help and cooperation. Appreciation is also extended to Elizabeth Nelson, Eugene Lamb and Ben Thwaite for help at various stages of the project.
Murphy, 1974) trained retarded subjects to use a cumulative rehearsal strategy in a keeping–track task and found that training enhanced performance. Improvement of paired–associate learning has been obtained by Borkowski and Wanschura (1974) and Turnure and Thurlow (1973) using imagery and sentence elaboration techniques. Kellas, Ashcraft and Johnson (1973) used both free and serial recall tasks and found improvement in performance after their subjects were trained to use a cumulative rehearsal strategy.

The retention of the trained strategies has been an important issue in instructional research since strategy training is useless if it is not retained over prolonged periods of time. While Milgram (1967) was unable to find any retention of mediation instructions on the part of his retarded subjects, Turnure and Thurlow (1973), using two sessions of sentence elaboration instructions with a paired–associate task, found significant retention after seven days. Brown's studies on the keeping–track task found retention of the cumulative strategy after six months and Kellas, et al., found that subjects continued to use cumulative rehearsal in a free–recall task after a two week delay.

Our main concern in the present research, particularly the first study, is with the free–recall task since so much of the learning and remembering that children are required to do is analogous to this paradigm. It has been well established that retarded subjects can be trained in the effective use of a cumulative rehearsal strategy in a free–recall task and that his training will continue to enhance performance even over a six month delay. Craik and Watkins (1974) have argued, however, that cumulative rehearsal is one of the least efficient techniques for establishing a strong memory trace. Craik and Lockhart (1972) have further argued that the strength of a memory trace is determined by the dimensions along which a stimulus is encoded and the extent to which the internal representation of the stimulus is associated with existing memory traces. It remains to be demonstrated that retarded individuals can be trained in the effective use of more sophisticated strategies involving semantic encoding and association with existing memory traces. It also remains to be determined whether retarded individuals of different chronological ages show differential memory performance when the task requires that the stimulus be encoded along one of several different dimensions.

It was thus the purpose of the first study to determine whether retarded children would use a semantic encoding strategy of a highly general nature and whether they would retain the use of this strategy on a test given seven months later without the use of prompting (cf. Borkowski & Wanschura, 1974). Since undirected encoding in children is primarily acoustic in nature (Bach & Underwood, 1970; Bisanz, Kail, Pellegrino & Siegel, 1977) and those rehearsal strategies that are used tend to be simple rote repetition (Flavell, Beach & Chinsky, 1966), these two encoding strategies were adopted as control procedures.

An additional question this study attempts to address involves the influence of induced encoding strategies upon categorical clustering. Clustering is one reflection of the general tendency to organize information according to some set of dimensions which is important not just in storing and retrieving memories but in all forms of problem solving as well.

Spitz (1966) and Gerjuoy and Spitz (1966) have reported that retardates do not cluster items at recall according to taxonomic category as is typically done by normal subjects. They also found that retarded subjects could be induced to cluster their recall either by presenting the items blocked by category or by requesting that the subject recall items from particular categories at the time of recall. However, Bilsky, Evans and Gilbert (1972) found that training subjects on blocked lists had positive transfer only to a randomly presented (i.e., non–blocked) list of the same items and not to a list of new items from the old categories. Thus, a second goal for this first study was to determine whether instructions to encode in a general semantic and associative fashion without instructions to recall in any particular pattern would induce categorical clustering in randomly presented lists and whether this tendency to cluster would remain effective on delayed tests.

EXPERIMENT 1

Method

Subjects. Forty-two children were chosen from special classes designed for children with IQ scores in the 50–75 range. All the children were either 5th or 6th graders and were contacted only after parental permission was obtained and ethical criteria were satisfied.

Materials. A pool of high frequency concrete nouns and their corresponding pictures were chosen from the Peabody Picture Vocabulary Test and several picture books written for children. The items were chosen to represent the following taxonomic categories: furniture, vehicles, kitchen utensils, body parts, clothes, food, animals, weapons, room parts, musical instruments, toys, occupations, drinking utensils, and celestial objects.

Five lists of items were composed such that each list consisted of four words from each of five different categories. Some categories were repeated in non–adjacent lists but each presentation of a category used different exemplars. Pictures of the items, typically simple line drawings, were pasted on 4 × 6 inch cards for presentation to the subject.

A videotaped film was constructed for each of the three conditions which depicted a woman experimenter, and a subject receiving two trials on the task,
using the strategy appropriate for that particular condition. The subject in the film was a college-age woman chosen to look and talk in a manner similar to our real subjects. The film depicted the subject overtly using the designated strategy for each condition but her recall and clustering were perfect for all three films. Each film lasted 10–12 minutes.

**Conditions.** There were 14 subjects in each of the three encoding conditions. In the Semantic condition the subject was told that the best way to remember the words was to think of the meaning of each item as it was presented, to try to think of personal experience with the object, to think of the functions of the object, and to try to remember other objects in the list from the same category. The subjects in the Repetition condition were told the best way to remember the item was to repeat the verbal label over and over again either subvocally or overtly. In the Acoustic condition the subject was encouraged to think of the sound of the word and to repeat the initial sound over and over again either overtly or subvocally.

**Procedure.** All subjects were given a pre-test before any assignment to conditions. The pre-test consisted of three free-recall trials on the 20 item pre-test list. The items were presented in a different order on each of the trials at a 20 second per item rate. The list items were not blocked according to category. The experimenter said the name of the item out loud as she presented each picture to the children but said nothing else during presentation of the list. Prior to list presentation, the subject was instructed to recall orally in any order as many of the items as possible. The total number of words recalled on all three trials for each subject was used to assign subjects to the three groups so that each group had approximately the same mean number of words recalled on the pre-test ($F < 1.0$).

About 2 or 3 weeks after the pre-test, the experimenter returned to the school and conducted the training sessions and the immediate post-test. The training session consisted of showing the child the film appropriate for his or her particular condition. The film was followed by three training-memory trials on each of two different lists. Before the film, the experimenter explained to the child that she was going to try to teach him or her how to remember better and then explained the strategy to be used for that particular condition. The child was then shown the film and told to try to learn the words in the same manner as the child in the film. This was followed by the initial trial of the first training list which consisted of 20 items at a 10 seconds per item rate. As each picture was presented, the experimenter prompted the child as to the strategy to be used and asked the child questions relevant to the item condition. The unpaced and oral recall was not prompted by retrieval or clustering cues by the experimenter. The other two trials were different orderings of the same list and the experimenter prompted the subject during input as on the first trial.

For the second training list, the experimenter reiterated the strategy at the beginning of each trial but said nothing about it during list presentation. Again the list was presented for three trials at a 10 seconds/item rate.

Immediately after the second training list the subject received the first post-test (PT1) which consisted of three trials on a new list at the 10 seconds/item rate. The experimenter gave no reference to the strategy at any time during the test. Seven days later the experimenter returned to the school and conducted a second post-test (PT2) which consisted of three trials on a new word list. Approximately seven months later a different experimenter (male) returned to the school and conducted a third post-test (PT3) which consisted of two trials on the same list of items used on the pre-test. For the first trial on PT3 the experimenter made no reference to the strategy the subject had been encouraged to use earlier but did ask if they remembered how to do the task. For the second trial of PT3 the experimenter reminded the child of the strategy and gave prompts for the strategy during list presentation.

**RESULTS**

The two dependent variables of primary interest in this study are mean words recalled and clustering of recall according to taxonomic category.

**Recall Data.** The recall data for the two training lists and three post-tests are shown in Figure 1. A separate analysis of variance was performed on each

![FIG. 1. Mean number of words recalled for each trial of the two training sessions and the three test sessions.](image)
of the five lists with Groups (3) and Trials (3) as factors. The analyses of PT3 are discussed separately since an analysis with unequal number of subjects was required. For both training lists and PT1 and PT2 the Semantic condition was superior in performance to the Repetition and Acoustic conditions while performance in the latter two conditions did not differ. These findings were supported by the significant Conditions main effect for Training 1 (F(2,39) = 17.2, p < .01), Training 2 (F(2,39) = 21.9, p < .01), PT1 (F(2,39) = 5.9, p < .01), PT2 (F(2,39) = 4.32, p < .01) and subsequent Tukey tests at the .05 level. The Trials main effect simply reflected improvement in recall over trials for Training 1 (F(2,78) = 32.8, F(2,78) = 68.6, p < .01).

While the Semantic condition was superior to the other two conditions from the first trial of each list, the rate of learning was not increased in this condition. The Condition × Trials interaction was nonsignificant for the two training trials and the first two post-test trials (F < 1.0, F < 1.0, F < 1.0, and F (2,78) = 1.5, respectively) indicating that the Semantic superiority was comparable over all three trials for these four lists.

Another analysis of variance was performed on the total number of words recalled, summed over all three trials on PT1 and PT2, that is, a Conditions (3) × Test (2) analysis. The purpose of this analysis was to determine whether the superiority of the Semantic strategy had diminished from PT1 to PT2. If the effect of Semantic strategy had diminished, a significant Conditions × Test interaction would have resulted. But this interaction was not significant, F (2,39) = 1.8, p > .10, indicating that Semantic strategy training had about as much effect after seven days as immediately after the training session.

On the seven month delayed post-test (PT3) many of the children had changed schools and several were unavailable for testing which resulted in unequal N’s for the three groups: Semantic, N = 12; Acoustic, N = 10; and Repetition, N = 11. An unequal N’s analyzed unweighted analysis of variance was performed on these data. The first trial that was unprompted in any fashion as to particular strategy showed nearly equivalent performance for the three conditions, but for trial 2, the prompted trial, the Semantic group performed much higher than the other two groups. Thus, on the unprompted trial seven months after training, the strategy manipulation did not have differential effects on performance. Only one trial, however, was necessary to restate the superiority of Semantic encoding. The above findings were reflected in a non-significant Conditions effect (F < 1.0), significant Trials effect (F (2,30) = 77.7, p < .01) and a significant Conditions × Trials interaction (F (2,30) = 7.3, p < .01).

Category Clustering. The recall data were scored for clustering, using formulae proposed by Bousfield and Bousfield (1966). One measure was the observed clustering score and the other was the clustering score expected by chance. Figure 2 shows the observed data for all lists and trials while the expected scores are only shown for the three post-tests (PT1, PT2, and PT3).

Because of considerable heterogeneity of variances for the observed data from the two training sessions, observed clustering scores summed over the three trials of each training session were submitted to a Kruskal-Wallis One-Way Analysis of Variance, with the three encoding conditions being the units of analysis. The first training session data resulted in H = 152.1, df = 2, p < .001; while the data from the second session resulted in H = 154.7, df = 2, p < .001. Inspection of Figure 2 shows that the Semantic strategy condition exhibited much more clustering than the other two conditions which were noted to be approximately equal. Also of interest is the fact that not only was the sizable clustering for the Semantic group present from the initial trial of the first training session, but that it increased steadily over trials for both training sessions.

The data from the three post-tests were analyzed by separate ANOVA’s with conditions and trials as factors. In correspondence with the recall data, the clustering data from the Semantic condition were superior to those for the Acoustic and Repetition conditions over all three post-tests while these latter two conditions were nearly equivalent. The Conditions main effects for PT1, PT2, and PT3 respectively, were F(2,39) = 8.6, p < .01, F(2,39) = 5.4, p < .01 and F (2,30) = 5.3, p < .05.

For PT2 and PT3, the clustering scores of the Semantic Condition increased over trials while no such increase was shown by the other two conditions. This was reflected in significant Conditions × Trials interactions for PT2 (F (2,78) = 4.62, p < .05) and PT3 (F(2,30) = 13.7, p < .01).
particularly interesting finding for PT2, the test session given seven days after training and in which no prompts were given for the subject to use any particular storage or retrieval strategy.

Lest the observed clustering scores for the Semantic group be considered an artifact of higher recall performance for this group, Figure 2 shows that, while the observed scores for the Acoustic and Repetition conditions approximate the expected scores for these groups, the Semantic observed scores are considerably higher than those values expected by chance. This is true for all trials of all three post-test sessions except for the unprompted trial 1 of PT3.

Discussion

The purpose of this first study was to determine whether mildly retarded children would use a rather sophisticated semantic encoding strategy in a memory task, whether this strategy would enhance recall and category clustering and whether the child would continue to use the strategy in unprompted test sessions delayed for up to seven months. The present study establishes that retarded pre-adolescents can and will use the semantic strategy. The recall data showed enhanced performance by the Semantic condition from the very first trial of the first training session. Of interest, however, was the fact that use of the semantic strategy did not increase the number of items mastered on each trial after the first.

For the first training session there was considerable prompting of strategy for each item of the list on each trial. For the second training session, however, the subject was simply reminded of the strategy at the beginning of each list presentation and the Semantic group performed as well on all three trials in this session as for training session 1.

When no prompting of any kind was used, on PT1, recall performance dropped for the Semantic condition by 2-3 items per trial but it was still considerably higher than the other two conditions. The fact that the Acoustic and Repetition conditions did not drop on PT1 when no prompting was used probably reflects that these strategies are similar to those that would be used by these subjects regularly in a non-prompted situation.

The data from the seven day post-test are the most interesting in showing that even without the benefit of prompting, subjects in the Semantic condition continued to perform at a high level of recall. We infer from this that the subjects were continuing to use the instructed strategy. Only this indirect inference is allowed because no direct measures of strategy usage were made. The clustering data for PT2, however, lend support to this inference since the subjects in the Semantic condition clustered above chance seven days after training in the absence of any prompts to do so. These findings are particularly encouraging in light of the minimal amount of training given our youngsters.

The PT3 data are fairly clear in showing no lingering effect of the encoding treatment on an unprompted test seven months after training. The mean recall on trial 1 of PT3 is virtually identical for the three groups. Prompting on trial 2, however, was again effective in quickly reinstating the encoding set and trial 2 performance rose considerably for the Semantic condition. Also interesting from the PT3 data is the overall high level of performance for all subjects on the unprompted trial 1. This could result because of some retention of the items from the pre-test nearly eight months prior, since the same items were used on the pre-test and PT3.

The clustering data also seem to demonstrate clearly that the deep level semantic strategy was effective at instating a set to cluster items at recall according to taxonomic category in spite of the fact that presentation was random and no cues were given at the time of recall. Therefore, these findings suggest that in contrast to previous research (i.e., Bilsky, et al., 1972; Gerjouy & Spitz, 1966; Spitz, 1966), retardates are capable of categorical clustering in the absence of both blocked item presentation and cueing of categories at recall if given training in sophisticated semantic encoding strategies. The interaction between Conditions and Trials on the observed scores of PT2 is interesting because, not only did the Semantic groups show significant clustering on the first unprompted trial seven days after training, but this group showed significantly accelerated rates of clustering over trials with respect to the other two groups. This might mean that over trials, the subjects in this group remembered more and more the proper strategy to use in retrieving items in the list.

EXPERIMENT 2

In the first experiment we have shown that a group of 5th and 6th grade educable mentally retarded children can encode words along the semantic dimension when prompted and can, essentially, be taught to prompt themselves to retrieve information about a to-be-remembered item such as its taxonomic category and former experiences with the item. An important question raised by this study is whether younger EMR children would or would not make use of a prompted semantic encoding to facilitate recall. If younger children can not use a semantic prompt when provided by the experimenter it is unlikely they could be taught to make effective use of self-prompted and thus would not benefit from the kind of instructional set used in Experiment 1.

To answer this question we decided to use a procedure made popular by Craik and Tulving (1975) and recently used in a study with children by Weiss, Robinson and Hastie (1977). With this procedure the subject is presented with a word and then asked some question about the word. The nature of the question is varied to require the encoding of various dimensions of the
stimulus item. The standard finding on a surprise recall test with adults and older children is that questions that elicit semantically related information, such as “Is this object an article of clothing?”, lead to better recall of the items than questions that focus on the more primitive or structural properties of the item, such as “Does this word rhyme with shoot?”. Our assumption in using this procedure was that any child that did not show differential recall for items given at least nominally different levels of encoding would probably not benefit from instruction of the type used in Experiment I.

*Method*

**Subjects.** Thirty-six children were selected from special classes serving educable mentally retarded children (IQ 55 to 75). Subjects were divided into two groups on the basis of age with the mean IQ of both groups being 65. The subjects in one group ranged in age from 9–11 and had a median age of 10 years 7 months. The subjects in the second group ranged in age from 12–14 and had a median age of 13 years 6 months. The subjects were contacted only after parental permission was received and all ethical criteria were satisfied.

**Materials.** Thirty-nine high frequency concrete nouns and their corresponding pictures were chosen from the Peabody Picture Vocabulary Test. The pictures of the items were pasted on 4 × 6 inch index cards for presentation to the subject. The set of 39 words was divided randomly into three different encoding or question-asking conditions.

**Conditions.** All thirty-six subjects participated in three encoding conditions. In the *Category* condition the subject was asked whether the object was a member of a particular taxonomic category. For the *Liking* condition the child was asked whether or not he or she liked the object. In the *Letter* condition the question inquired as to whether a particular letter was present in the word.

In the Category and Letter conditions the total number of yes and no responses was approximately equal (6 or 7 of each). No such control was possible with the Liking condition. The pairing of the three sublists with the three encoding conditions was counterbalanced across subjects.

**Procedure.** Each subject was seen individually by a male experimenter for one session. After adequate rapport was established, the following instructions were given to the subject: “We are going to play a game about words. I will say a word and show you a picture of that word. After each word I will ask you some questions about it. Take your time and think before you answer. This is just a game between you and me, and you will not be graded for it.”

The experimenter said the name of the item aloud as he presented each picture to the child. This was followed by a question appropriate to the encoding condition. The thirty-nine words were presented consecutively to the child at a 5-second rate and the order of the three conditions was counterbalanced.

Upon completion of the entire list, the subject was engaged in conversation for one minute and then given an unexpected request to recall orally as many of the thirty-nine words as he or she could remember. Each subject was given five minutes for recall, after two minutes of which he or she was given an encouragement to try to remember some more words.

**Results**

The mean number of items recalled for the two age groups at each of the three encoding conditions is shown in Figure 3. These recall data were submitted to an Age (2) × Condition (3) analysis of variance with age as a between-subjects factor and Condition as a within-subject factor. The only significant difference between the means plotted in Figure 3 was the greater recall performance for the 13-year old group in the Category condition than the other five Age by Condition means. This one difference was sufficient to yield significant effects of Age (F[1, 34] = 5.2, p < .05), Condition (F[2, 68] = 14.3, p < .001) and Age × Condition (F[12, 68] = 4.8, p < .025). While the younger subjects showed a slight trend toward higher performance for the Category condition, the difference was not close to significance.

To determine whether recall performance was affected by whether the word was given a yes or no as a correct response, the recall data were analyzed through an Age (2) × Condition (2) × Question Type (2) analysis of variance. The Liking condition was not included in this analysis since we had no control over the number of yes and no responses. The analysis showed that there was

![FIG. 3. Mean words recalled for the 10- and 13-year olds in the three encoding conditions.](image)
no difference in the recall of Yes and No words ($F(1,34) = 1.3, p > .10$) nor did Question type interact with Age ($F(1,34) = 1.3, p > .10$) or Condition $F < 1.0$). The three-way interaction was also nonsignificant ($F < 1.0$).

Discussion

As mentioned above, the instructional set used in the Semantic condition of Exp. 1 depends on the subjects 1) being able to use prompts of a semantic nature given by the experimenter and 2) learning to provide those prompts to themselves at a later time. The present experiment would suggest that the children in the younger group would not benefit from such an instructional set since the prompts would do no good in facilitating recall even if the child did learn to self-initiate them. The present data also are of some interest to theoretical arguments regarding the levels of processing notion in general. This approach, at least as presented by Craik and Lockhart (1972), assumes that a category decision would require more elaborate processing than would a letter decision. It is not clear, given that assumption, why the younger subjects failed to have their recall performance facilitated by this more elaborate processing.

It is also unclear why both of our groups showed no difference in recall of “Yes” words and “No” words. A highly replicable finding (Craik & Tulving, 1975), at least with normal adults, is that “Yes” items are better recalled than “No” items. In accord with the study reported by Weiss, Robinson and Hastie (1977) with normal children, our study showed no difference in the recall of affirmative and negative response items.

The current studies both seem to be relevant to the argument between those classes of theories proposing a structural deficit in the memory system of mildly retarded subjects and those theories arguing for strategic or control deficits as the basis for memory deficiencies (Robinson & Robinson, 1976). By structural deficits, most theorists seem to mean actual neuropathology in some form (cf. Spitz, 1963) whereas control deficits refer to the subject not having available those often complex strategies for coding, storing, and retrieving information in order to remember it. We would argue that there is a deficit of another kind that is not quite but almost structural and not quite but almost strategic. That deficit is in the amount and complexity of knowledge existing in the retarded child's long-term memory (what Rumelhart, Lindsay & Norman, 1972, refer to as “declarative knowledge”) and the number of associations that are automatically elicited when a verbal item is presented. Experiment 2 in the present research suggests that although the semantic network may not be sufficiently rich to allow a category decision to facilitate recall at age 10 in the EMR child or at age 6 in the normal child (Weiss, et al, 1977) it is sufficiently rich by the time the EMR child is 13 or the normal child is 8 years of age. This deficit in the richness of the semantic network is not really structural in the sense that it is physiologically based but it would be structural in the sense that it is based on an automatic process (e.g., automatic spreading activation) not under the conscious control of the subject; that is, it does not require processing by the limited capacity attentional mechanism.

This deficit is not really strategic in the sense that it is some subroutine called into action under the conscious control of the subject but it is strategic in the sense that this declarative knowledge is obviously composed of learned associations and at one time they were not automatically elicited as they would be in the 13-year old EMR child or 8-year old normal child. To the extent that eliciting these associations requires limited capacity attention they would be classified as being strategic. With practice the elicitation of this semantic information would become automatic and not require limited capacity attention. Evidence for the development of this automaticity in normal adults is provided by Schneider and Shiffrin (1977).

It has been well documented (cf. Belmont & Butterfield, 1971) that there is an increase in the number and sophistication of subroutine type strategies with increases in intelligence. We would argue that, in addition, there is an increase in the number and richness of automatically elicited associations to a given stimulus. With increases in intelligence comes a more rich and automatically elicited semantic network. Thus, even though our 10-year old subjects could correctly respond that boots are an article of clothing, the richness of semantic network automatically activated in the course of that decision would be much less than in the more intelligent 13-year olds. If these ideas are correct then the approach to remediation of memory deficiencies that Belmont and Butterfield (1977) have labeled the “instructional approach” must be revised somewhat. To get the EMR child to remember as well as the normal child it is not sufficient to equate the size and quality of their strategic repertoires. We must also equate the extent, richness and automaticity of their declarative knowledge.

REFERENCES


