PRODUCTION NOTE

University of Illinois at Urbana-Champaign Library
Technical Report No. 94

TRAINING SELF-CHECKING ROUTINES FOR ESTIMATING TEST READINESS: GENERALIZATION FROM LIST LEARNING TO PROSE RECALL

Ann L. Brown, Joseph C. Campione, and Craig R. Barclay

University of Illinois at Urbana-Champaign

July 1978

Center for the Study of Reading
Technical Report No. 94

TRAINING SELF-CHECKING ROUTINES FOR ESTIMATING
TEST READINESS: GENERALIZATION FROM
LIST LEARNING TO PROSE RECALL

Ann L. Brown, Joseph C. Campione,
and Craig R. Barclay

University of Illinois at Urbana-Champaign

July 1978

This research was supported in part by Grants HD 06864, HD 05951, and a
research career development award HD 00111 for the first author from the
National Institute of Child Health and Human Development, and in part by the
National Institute of Education under Contract No. US-NIE-C-400-76-0116. The
authors would like to extend their especial appreciation to Mrs. Wilma Noynaert,
Assistant Director of Special Education for the Peoria Public Schools. She
has continually provided support and encouragement and has made access to
special education classes readily available. Thanks are also due to the
following principals: Mr. William Jordan of Von Steuben School, Mr. Lee
Nugent of Hines School, Mr. Dennis Bitner of Trewyn School, Mr. William Ricca
of Roosevelt School, Mr. Robert Carruthers of Lincoln School, Mr. Robert
Sorensen of Kingman School, and to all the teachers for their generous and
willing cooperation. The authors would especially like to thank Mrs. Carolyn
Long for her patience and skill in testing the children and, of course, the
children themselves for their usually willing and always active participation
and their keen sense of humor.

Craig Barclay is now at State University College at Cortland, New York.
Abstract

Brown and Barclay (Child Development, 1976, 47, 71-80) trained educable retarded children to use either of two memory strategies, Anticipation or Rehearsal, involving a self-checking component. Following the training, both their free recall performance and their ability to estimate their readiness for a recall test improved significantly. In the present research, the students were tested for maintenance and generalization one year following the original training. The younger children (MA = 6 years) showed no effects of the training, whereas an older group (MA = 8 years) both maintained the trained strategies on the original rote recall task and generalized it effectively to a novel situation involving gist recall of prose passages. In comparison to a pair of control groups, the students trained in the use of self-checking routines took more time studying, recalled more idea units from the passages, and further, their recall was more clearly related to the thematic importance of the constituent idea units, a pattern characteristic of developmentally more advanced subjects.
Children with marginal academic skills, which render them at risk for special education, are found to experience particular problems in two main areas: strategic planning in school problem-solving tasks (including deliberate remembering) and reading effectively. Our interest in developing training routines to overcome some of these deficiencies stems from the belief that remediation aimed at marginal children can be the most fruitful in terms of obtaining worthwhile educational improvements. It also reflects our belief that average children acquire many academic skills without explicit training; repeated contact with a variety of tasks in school, all requiring the same basic strategies, is probably sufficient to inculcate at least the very simple routines of learning, remembering and beginning reading. We have considerable evidence that educable retarded (EMR) children do not acquire these rudimentary skills without intensive intervention (Belmont & Butterfield, 1971, 1977; Brown, 1974, 1975, 1978a, 1978b; Brown & Campione, 1977, 1978; Campione & Brown, 1977; Butterfield & Belmont, 1977).

Considerable effort has been expended in recent years in attempts to train mnemonic strategies in EMR students who would not think to introduce such mediation on their own volition. The large majority of these training attempts have concentrated on inculcating three specific mnemonics: rehearsal, categorization, and elaboration (Brown, 1974; Campione & Brown, 1977). But we have considerable reason to doubt that this is the best approach if one is interested in effecting a lasting improvement in memorization skills. The problem lies in the dubious success so far achieved by attempts to train these common memory skills. The general picture to emerge is that EMR children readily respond to appropriate training and evidence a variety of trained mnemonic skills accompanied by a reasonable increment in recall performance (Belmont & Butterfield, 1971; Borkowski & Wanschura, 1974;
In addition, following well-designed and extensive training, maintenance of the effects of this experience has been found even after periods of up to one year have elapsed since training was initiated (Brown, Campione & Murphy, 1974, 1977). Evidence for generalization to new situations, however, is not so easy to come by (Brown, 1974, 1978a; Campione & Brown, 1974, 1977). The problem of transfer of training is an old one, particularly in the context of the learning capabilities of retarded children. As the Soviet psychologist Shif (1969) has pointed out, one of the major problems experienced by those engaged in training mildly retarded children is that they tend to acquire information which is in some sense "welded" to the form in which it was learned. Flexible transfer to novel variants of the training task is not a cognitive characteristic of the slow learning child.

The absence of really convincing evidence of flexible generalization of trained mnemonic strategies has led some investigators (Brown, 1974, 1978a; Butterfield et al., 1973; Campione & Brown, 1977) to advance the view that training efforts should be directed at more general determinants of performance rather than the task-specific skills or strategies that have so far been the subject of extensive investigation. Instead of training a routine that is specifically tailored to the needs of a particular task, it might be useful to attempt to inculcate more general knowledge concerning strategies and their use. If we are interested in effecting improvement in the child's general performance on a variety of similar tasks, then we must consider both the specific gains from training (strategy use after training) and the general benefits (improved knowledge concerning memory tasks).

The argument is simple; if young children are totally unaware of the utility of, for example, mnemonic aids, why should they benefit from instruction? If trained to rehearse they will rehearse, especially if the situation
remains unchanged and they receive continual reminders. But why should they then be expected to use their new skills insightfully if the reason for the activity was never made clear? Several recent studies have illustrated this point. Young normal children will maintain the effects of training if they can be made aware of (a) the usefulness of the trained strategy (Borkowski, Levers & Gruenenfelder, 1976; Kennedy & Miller, 1976), or (b) the fact that the separate tasks in different phases of the experiment are related (Bullock & Gelman, 1977; Campione, 1973). But the utility of a prior solution or strategy to a new variant of the task does not seem to occur to them spontaneously.

For these reasons we believe that it would be judicious to rethink the type of skills we attempt to train. An alternative approach to training specific mnemonics would be to train skills which could have generality across a variety of problem-solving situations, skills such as checking, planning, asking questions, self-testing, and monitoring ongoing attempts to solve problems (Brown, 1978a; Brown & DeLoache, 1978). The types of general problem solving skills we are particularly interested in would be the ability to stop and think before attempting a problem, to ask questions of oneself and others to determine if one recognizes the problem, to check solutions against reality by asking not "is it right" but "is it reasonable," to monitor attempts to learn to see if they are working or are worth the effort. There are serious problems associated with this position, for it is certainly easier to suggest that training should be aimed at showing children "how to organize their knowledge" and "how to solve problems" (Norman, Gentner & Stevens, 1976) than it is to instantiate these suggestions in concrete training programs. But in view of the dismal failure to induce generalization in EMR children, an attempt to address the problem of training general skills seemed to be timely.
Our next step, therefore, was to attempt to train a very general metacognitive (Flavell & Wellman, 1977) strategy which could reasonably be supposed to have a wide range of application. Basically, we attempted to train a "stop-check-and-study" routine (Brown & Barclay, 1976). The specific task used, one of assessing readiness to recall, was adapted from Flavell's original work with nonretarded children (Flavell, Friedrichs & Hoyt, 1970). In our initial study (Brown & Barclay, 1976), the children were given a series of trials on which the task was to memorize in order a set of \( n \) items. The \( n \) for each subject was determined individually and was equal to one and one-half times the maximum number he or she had been able to recall consistently on a series of practice trials, i.e., \( 1\frac{1}{2} \) times his span on this task. He was instructed to continue studying the items until he was sure he could recall all of them perfectly, at which time he was to signal the experimenter. Judging recall readiness for such supr-span lists is an intriguing task, for it demands a complex form of self-evaluation, involving both the use of a specific mnemonic strategy (introduced to effect learning) and the ability to monitor its success; to both behave strategically and to "self-test" the success of the strategy in order to intelligently terminate study activity. In addition, it requires not only the ability to differentially study difficult items, another metamemory ability sensitive to cognitive maturity (Brown & Campione, 1977; Brown & Smiley, 1977a, 1977b), but it also requires that the subject engage in self-testing activities to determine which are the difficult items.

Not surprisingly, EMR children performed very poorly on the initial pretesting phase of the study. Only 4% of children with MAs of around six years and 12% of students with MAs of around eight years gave even one perfect recall, poor performance considering the children were allowed as much study time as they wanted. One reason why the developmentally young perform so poorly on this task could be that they do not tend to introduce strategies of deliberate
memorization involving self-testing elements which would alert them to their readiness for a test. If the children do not use such self-testing devices, they can hardly be expected to monitor their own stage of learning.

For this reason we trained groups of children in the use of three strategies of remembering: anticipation and rehearsal, both of which involve self-testing elements; and labeling, which does not. The labeling condition essentially served as a control treatment. The subjects were required to go through each list once, naming each picture. This labeling trial was followed by a series of three more trials on which the procedures differed between the groups. Those in the Anticipation group were trained to anticipate the next picture by saying its name before exposing it. The Rehearsal subjects were trained to rehearse the items in sets of three. Finally, the Label group was told to go through the list three more times, labeling each item. All groups were further encouraged to continue with the instructed activity until they were sure they could recall all items. Training was continued for two days.

Following training, three posttests were given, a prompted posttest (one day after training) on which individuals were instructed to continue the trained strategy, and two unprompted posttests given one day and approximately two weeks later. Both the younger and older children in the Anticipation and Rehearsal groups improved their recall scores significantly on the prompted posttest, as can be seen in Figure 1; further, 72% of the younger subjects and 92% of the older subjects gave at least one perfect recall. Thus, training

Insert Figure 1 about here

the useful self-testing strategies resulted in both enhanced performance (percent recall data) and improved monitoring (data on number of perfect
recalls) compared with the control Label group.

Also shown in Figure 1 are the data from the unprompted posttests (2 and 3). The MA 6 and MA 8 groups differed considerably on the unprompted posttests. For the younger group, performance on posttests 2 and 3 was not significantly different from the pretraining level, whereas for the older group, performance on all posttests differed significantly from the pretraining level. Thus, as in previous studies concerned with direct training of a strategy, training facilitates performance, with the effect being somewhat durable for the older children but transitory for the younger ones (Campione & Brown, 1977).

In the present research, we extended the results of the original Brown and Barclay study. There we had shown successful maintenance of training, at least in an older sample of EMR students (MA = 8 years). Here we report the findings of a series of long-term maintenance tests and a generalization phase introduced over a year after the end of training. We were interested in a number of questions. The first concerned the extent to which the MA 8 children would show durable effects of training. For the younger group, we did not expect any evidence of such maintenance for the obvious reason that they had reverted to pretraining levels during the original study (see Figure 1). For these subjects, therefore, we were concerned only with the effect of prompting them in the use of the trained strategy. Specifically, would such prompting result in enhanced performance even after a one-year delay since training, and more interestingly, would the effects of prompting be maintained on a subsequent series of unprompted trials? Finally, if there were reasonable evidence for maintenance, we planned to probe for generalization of a self-testing strategy to a different recall-readiness task.

**Maintenance Phase**

**Method**

**Subjects.** In the original study, there were 66 EMR subjects. We were
able to locate 58 of them for the one-year maintenance tests. Of the original 27 MA 6 students we retested 21 and of the original 39 MA 8 students we retested 37. The complete descriptive statement for these samples can be found in Brown and Barclay, (1976). The mean IQ of the two groups was approximately 70, mean CA 10 and 13, and the mean MA 6-7 and 8-9 at the time of retesting. Although the MA scores obviously did increase over the one-year period, we will continue to refer to the groups as MA 6 and MA 8, respectively, for convenience and consistency.

**Apparatus.** An automated apparatus (Scott, 1970) was housed in a tworoom experimental suite located in a motor home. The apparatus featured a multiple window display panel which contained two horizontal arrays of 5 x 5 cm windows, with a separation of 1.9 cm between windows. Each window had a separate rear projection screen with a shutter mounted on the back. When the shutter was operated by the subject's pressing the window, the stimulus appeared in the window and remained visible until the subject pressed another window, at which time the stimulus in the second window appeared while the original disappeared.

**Stimulus Material.** The stimulus pool consisted of 12 slides each containing 12 colored pictures of common objects, one for each window of the apparatus. The 12 slides were different from the 24 comparable slides that were used in the original testing phases (Brown & Barclay, 1976). The pictures were deliberately selected to minimize inter-item clusterability. All of the items were readily labeled by comparable subjects and usually elicited a one-word response. In all phases of the experiment one of the 12 slides was randomly selected from the pool, without replacement, in such a way that all subjects saw new pictures on each trial, and the order was randomized across subjects and conditions.

**Procedure.** Each subject was seen individually on a series of four
consecutive days. On Days 1 and 2 the experimenter was not informed of the subject's training history. On each day the subject received three trials, a trial consisting of one n-item list to learn. As in the earlier phases of this research, the n for each subject was equal to 1½ times the subject's original span. (The 12 - n remaining windows of the apparatus were covered with black posterboard.) On Days 1, 2, and 4, the subject was given unprompted posttests identical to those given in the Brown and Barclay study. The subject and experimenter sat in front of the apparatus, and the subject was instructed to try to learn the pictures' names in order. He was told to push the windows to expose the pictures in the list and that, "you can look at any one you need to and in any way you want, but I'm going to want you to say them back to me in this order (E pointing to each window in a forward serial order). You can look at the pictures as long as you want. When you are really sure that you know them all very, very well and can say them all back to me, ring this bell (desk bell) and try to say them to me." It was further emphasized that they should be very careful that they could say them all without error before terminating study. No mention was made to the subject of his prior training on this task. If, however, the subject mentioned that it was an old game, the experimenter assented but did not give any description of the desirable strategies.

On the third posttest day, the experimenter was informed of the subject's prior training and she proceeded to follow the procedure of the prompted posttests of the Brown and Barclay study. The procedure was the same as the unprompted tests, but the experimenter specifically reminded the subjects of their trained strategy. That is, subjects trained to anticipate were told that they had played this game before and it helped them to go through the list one by one trying to guess the name of the picture before they exposed it. Similarly, the subjects in the cumulative
rehearsal group were told to remember the rehearsal method. The experimenter gave one demonstration trial on an old list and then gave the three new slides for that day, continually prompting the subject to try and guess the picture names before exposing them (Anticipation group) -- say the names over and over in bunches (Rehearsal group) -- or say the names when the pictures were seen in the window (Label group).

On the remaining day (Day 4) of the maintenance phase, the subjects were given their final three tests without prompting. Again the experimenter answered all queries in a noncommittal fashion saying that the subjects should do whatever they thought necessary to help them remember all the pictures. During all stages of the experiment, the experimenter recorded accuracy of response, study time (the total time from the exposure of the first item to recall-readiness indication) and any overt signs of remembering activity. All subjects were rewarded with candy or small gifts at the end of each session.

Results and Discussion

The first question concerns the extent to which the effects of training were durable given that one year had elapsed since the study was initiated. The relevant data are included in Figure 1 separately for the two MA groups. The first four tests (Pre-test through Post-test 3) are from the original Brown and Barclay study. Recall that Post-test 1 was a prompted recall test. The data shown for Post-test 4 are the means of the first two days of this study, i.e., performance on the initial unprompted trials.

Considering first the Young group, the data were as expected. There was no evidence for maintenance, as performance remained at the pre-training level. For the Old groups, the data were more encouraging, as clear evidence of maintenance was obtained. The Groups effect is still large, and
performance on Post-test 4 is comparable to that on Post-test 3, still well above pre-training levels. The MA 8 students in the two conditions where self-testing strategies (anticipation and rehearsal) were trained continue to outperform those students trained to use an inappropriate strategy (labeling) that lacked a self-testing feature.

The complete data from the maintenance phase are shown in Table 1, which includes the mean proportion of correct recall as a function of MA, Group, and Test Days.

Insert Table 1 about here

Days 1, 2, and 4 of the maintenance phase were unprompted and Day 3 involved prompting by the experimenter. Given the different patterns obtained in the original study, the data from the two MA groups were analyzed separately, each via a 3 (Groups) x 4 (Days) mixed analysis of variance. For the Young group, Days produced the only significant contrast, $F(3,54) = 22.90$, $p < .001$. Scheffé follow-ups indicated that performance on Day 3, the prompted test day, was superior ($p < .05$) to that on each of the remaining days, which themselves did not differ.

Thus for the Young group, the results were straightforward. There was no evidence of maintenance on the initial unprompted tests. Prompting resulted in a large and reliable improvement in performance; however, on the fourth day of testing, performance was back to its original level. In this regard, the data replicate those of the earlier study: (Brown & Barclay, 1976): intervention leads to enhanced recall, but the effects of the intervention are transitory. Without continual prompting, the younger children show little evidence of the effects of intensive training. The failure of the younger students to maintain their enhanced performance on the final
nonprompted test is in keeping with previous studies where immature trainees abandon a trained strategy when no longer specifically instructed to continue in its use (Keeney, Cannizzo & Flavell, 1967). It should be noted that at least with normal children of comparable MA (CA 6-7) this tendency to relinquish a trained strategy can be overcome by the simple expedient of providing explicit feedback that the strategy does indeed enhance performance (Kennedy & Miller, 1976). The EMR students in the present study received a good deal of feedback. The aim of the game was to achieve perfect recall and the students knew this for they constantly complained of their failure to recall the entire list. In addition, the experimenter revealed the entire list again at the end of a trial to relieve their frustration by telling them what were the missed items. Explicit statements to the effect that their recall was much improved on strategy-use trials, similar to that given in the Kennedy and Miller experiment, were not part of the feedback procedure. It is possible that such a minor addition could overcome the recalcitrance of the MA 6 EMR trainee, but we have reason to doubt this as prior attempts to provide explicit feedback in a span-estimation task did not provide lasting benefits for a comparable MA 6 sample (Brown, Campione & Murphy, 1977).

Turning to the maintenance data of the Old group the effects due to Conditions, $F(2,32) = 4.09, p < .025$, and Days, $F(3,96) = 10.47, p < .001$, were both reliable. As can be seen in Table 1, the Conditions effect indicated that Anticipation (mean recall = 82%) and Rehearsal (79%) groups outperformed the Label (63%) group. The mean recall scores for the four test days were 69%, 69%, 83%, and 78%, respectively. Scheffé tests indicated that there were significant ($p < .05$) differences between performance on Days 1 and 2 and Days 3 and 4. The Day 1-Day 2 and Day 3-Day 4 differences did not approach significance.
Thus, the data obtained from the Old group include impressive evidence for maintenance. Even one year after training, the effects of the original instruction are readily discernable. Considering the Day 3 and Day 4 data, it is again the case that they provide a limited replication of the earlier study. Intervention results in enhanced performance, and the effects of that training are detectable on the day following that training, i.e., the improvement induced by prompting on Day 3 was maintained on Day 4.

**Generalization Phase**

Given the poor performance of the younger students we made no attempt to test these children for evidence of generalization. The older students were clearly more promising, however, so we decided to see whether they would show the benefits of the recall-readiness training on quite a different task. Systematically studying material until it is judged to be well enough known to risk a test, is, of course, a very general strategy, as any student could attest. Therefore, we were hoping that even with very different materials, the children who had received extensive training would show some generalized benefits.

The transfer task selected was one which we believed to be more representative of the type of study activity required in the classroom. Most studying requires the student to extract the main ideas of prose passages and regurgitate the gist of the ideas in his own words. Our question was, would training recall-readiness on the simple rote-list learning task help children on the more typical school study activity of preparing for gist recall of prose passages? Although strategies of anticipation and rehearsal are useful on a prose learning task they would have to be modified considerably from the straightforward procedures suitable for learning lists of words. Rehearsal or anticipation of individual words would be inefficient.
and the subject would have to attempt anticipation or rehearsal of longer chunks of material. In addition, the criteria for judging readiness are much more subtle. In the rote recall task readiness is reached when the learner can recall all items verbatim and it is relatively easy for the learner to check this prior to recall attempts. But in the gist recall tasks, the learner must gauge when he has grasped the main ideas of the material, for verbatim recall is not required. Thus, the training and transfer tasks were quite different in their strategy-use and strategy-monitoring requirements even though they demanded the same general "stop-check-study-recheck" routine. We reasoned that if we could find transfer under these conditions our training would really have been effective and have had practical utility; if we did not find transfer we could always revert to less ambitious transfer tasks, those more like the training vehicle (Brown, 1978a; Campione & Brown, 1977).

Method

Subjects. Of the 37 older students tested for maintenance we were able to locate 33 at the time of the generalization phase. Of the original 13 subjects in each of the Anticipation, Rehearsal, and Label groups, we tested 12, 11, and 10 respectively on the generalization problems. In addition to the 33 trained students we tested 17 EMR students who were sharing the same junior high school classes as the trained subjects. These additional students had no previous experience in our testing program. They were matched for CA, IQ, and MA with the trained students. In addition, all subjects in this study, naïve and experienced, were reading at least at the second grade level.

Stimulus Material. A set of 12 simple stories were written, based on the "I can read series," but rewritten so that they conformed to a Dale-Chall readability score of second grade difficulty. The stories were divided into
idea units by independent groups of college students. They ranged in length from 10-19 idea units with a mean of 15. The stories were then retyped with one idea unit per line and given to other independent groups of college students who were asked to rate the idea units in terms of their importance to the theme of the story. The procedure adopted was that introduced by Johnson (1970) and used by Brown & Smiley (1977a). The raters were told to delete one third of the units that were least important to the theme (using a red pencil). Then they were asked to delete a further third of the units (with a green pen) leaving one third exposed. It was explained that the units exposed at the end were the most important to the theme, those deleted first were the least important, and the remaining third were of intermediate value.¹

Procedure. The subjects were seen individually on a total of six days, approximately three months after the end of the maintenance phase. The experimenter was not informed of the prior training condition of the subject or indeed of the fact that this "experiment" was related to the prior phases of the study. She was simply told that we wanted to test all subjects who met certain selection criteria concerning IQ, MA, CA, reading scores, and class placement. The experimenter was led to believe that the study represented the beginning of a new series of experiments concerned with beginning reading skills.

No mention of the prior testing was made to the subjects either. On each day they were given two of the twelve stories randomly selected from the pool without replacement. First the experimenter read each story (printed in primary type) to the subjects and then the subjects read it back twice with the experimenter sounding out and explaining any words the subjects could not read or understand. Then the subjects were told to read the story again, as many
times as they liked, in order to try to remember everything that happened. They were told that they could do anything they liked to help them remember and they were provided with pencils, pens, felt-pens, paper, string, a stopwatch, the three key words of the story on cards, and a picture of the main action of the story. They were told that they could take as long as they liked learning the story. When they were sure they could tell all that happened in their own words they were to ring the desk bell and then try to tell the story. All subjects had pretraining at gist-recalling a simple story and recording their recalls on a tape recorder. The features of gist recall were fully explained. The subjects' recalls were played back to them and they were asked if they had anything to add.

This procedure was repeated for two stories a day for six days. The procedure was identical on Days 1, 2, 5, and 6, the unprompted days. On Days 3 and 4, however, the experimenter added the extra prompt that the pens, paper, story pictures, and key words were there as study aids. The subjects were informed that it helps some people to underline, mark the paper, look at the pictures to check if they are ready for a test, take notes, etc. It was pointed out that the subjects could do all or any of these activities if they wanted, in order to make remembering easier.

The experimenter recorded the study time (from the last reading aloud of the story until the bell indicated recall-readiness) and attempted to record any overt signs of deliberate attempts at remembering, such as looking away and self testing, repeating words or phrases, etc. Any attempts to underline or take notes were retained and attached to the transcript of the subject's recall of that story. The recalls were transcribed onto index cards and scored for gist recall by two independent raters, blind to the purpose of the study and experimental condition of the subject.
Results and Discussion

For each subject, the proportion of idea units recalled at each level of importance was computed for each story. To reduce the variability in the data, the scores from successive pairs of days were combined, resulting in a variable (Tests) with three levels. Test 1 consisted of the four recalls given on Days 1 and 2 and constituted an unprompted test. Test 2 included performance on the prompted days (3 and 4), and Test 3 involved performance on the final two (unprompted) days.

Separate analyses were conducted on the recall data and on study time. The recall analysis was a 4 (Groups) x 3 (Test) x 3 (Importance Level) mixed analysis of variance. Reliable effects were obtained due to Groups, $F(3,46) = 15.31, p < .001$; Importance Level, $F(2,92) = 309.47, p < .001$; and the Groups x Importance Level interaction, $F(6,92) = 5.76, p < .001$. These data are shown in Figure 2, where it can be seen that the Groups effect was due to the Anticipation (mean recall = 50%) and Rehearsal (49%) subjects' outperforming the Label (35%) and New (37%) subjects. Thus, students trained in task appropriate self-testing strategies did recall more than the two control groups.

The Importance Level main effect indicated that recall improved as thematic importance decreased. While the latter trend was reliable for all groups ($p < .001$), as indicated by simple effects analyses, the pattern differed across groups. Scheffé follow-ups to the simple effects analyses indicated that for the Anticipation and Rehearsal groups, there were significant ($p < .05$) differences between Level 1 and Level 2 units, as well as between Level 2 and
Level 3 units. For the Label and New groups, however, the Level 1 and Level 2 units produced comparable recall levels. For these groups, Level 3 units were recalled more frequently than Level 2 units. This pattern fits in nicely with results from prior studies of prose recall in children. Students of third grade or above differentiate all levels of importance in their recall of stories (Brown & Smiley, 1977a, 1977b); however, less mature children, while they do favor the most important elements somewhat, fail to differentiate the lower levels of importance in recall (Smiley, Oakley, Worthen, Campione & Brown, 1977). The less mature pattern has been found in poor readers of normal intelligence, EMR children and normal children younger than third grade. Thus, the appropriately trained MA 8 subjects in this study both recalled more and demonstrated a more mature pattern of recall than did the inappropriately trained or naive controls.

Turning to study time, a 4 (Groups) x 3 (Tests) mixed analysis of variance was performed on the median time per test, and the only reliable effect was due to Groups, F(3,46) = 5.28, p < .01. The Anticipation subjects studied each passage for an average of 103.7 seconds before attempting recall, and the corresponding figure for the Rehearsal subjects was 105.6 sec. These values were considerably larger than those for the Label (57.6) and New (62.4) subjects. Thus, the appropriately trained subjects did study longer than the two control groups, another index of more effective monitoring of recall-readiness.

We had some difficulty with the observational data. First, only one observer recorded evidence of activity and as this observer was also the sole experimenter she had insufficient time to do this as thoroughly as we would like. Also, as we had no means of checking the reliability of her ratings we do not place much reliance on these data and will only mention the outcome
briefly here.

There was very little evidence of any sustained overt study activity. In addition, there was no evidence of a general trend to increase overt activity in the prompted phase of the generalization tests. A few subjects did begin to underline sporadically when prompted (a few responded to instructions by underlining all the text!) but again a few (not necessarily the same) subjects underlined sporadically before the prompt. One thing is clear, it will take more than a mild prompt before EMR children will adopt a common study strategy such as underlining or note taking. This is not surprising because we have found that even normal children below seventh grade do not respond effectively to mild prompting to underline or take notes (Brown & Smiley, 1977b).

One observational measure from this study was suggestive. The number of subjects who showed any evidence of a variant of the trained strategies of anticipation or rehearsal did differentiate the groups. Scoring very leniently, and including any subject who was seen, even once, to be undertaking these activities, we found that 67% of the Anticipation and 64% of the Rehearsal groups did show some observable evidence of strategy maintenance. Such evidence included looking away and self-testing, repeating chunks of material over and over, and self-admonitions to stop and check before ceasing study activity. By contrast, only 40% of the Label and 29% of the New groups were observed to engage in these strategies. As mentioned previously, the observational data are included only as indicative of generalization. In the absence of reliability scores the observational data should be interpreted with caution.

General Discussion

This study represents our first successful attempt at inculcating a generalized cognitive skill in EMR children. Students (MA = 8) trained to use
common mnemonics that embody self-testing routines were found to perform much better on the list-learning, recall-readiness task that was the vehicle of the original training. The durable effects of training were still present even a year later. Those students who successfully maintained adequate recall-readiness for a list learning task also revealed the benefits of this training on a quite dissimilar recall-readiness task involving prose learning. Students trained in the task-appropriate strategies of anticipation and rehearsal out-performed control subjects on several indices of recall-readiness monitoring, including amount of gist recall, patterns of recall as a function of the importance of textual segments, amount of time spent studying and overt signs of strategy maintenance. We attribute the success of this training program, to the redirection in our thinking concerning what skills to train. Having failed to effect generalization when attempting to inculcate specific mnemonics we turned to a more general problem-solving routine involving self-testing of the effects of these mnemonics. The monitoring routine necessary for adequate recall-readiness estimation does appear to be (a) susceptible to training and (b) generalizable across quite distinct tasks.

As a result of both our success and failure in attempting to train EMR children to perform more effectively on common memorization tasks we suggest that the types of cognitive activities which are most suitable for intensive intervention should have certain properties, (a) they should have wide transsituational applicability, (b) they should readily be seen by the child to be reasonable activities that work, (c) they should have some counterpart in real-life experiences, and (d) their component processes should be well understood so that effective training techniques can be devised. This bias directs us to a concentration on a subset of general metacognitive activities (Brown, 1978a, 1978b; Brown & DeLoache, 1978) which we feel admirably fit this
prescription. These include checking, monitoring, and reality testing. This is, of course, still too ambitious and, therefore, we have chosen to concentrate on routines that can be subsumed under the heading self-interrogation.

The eventual aim is to train the child to think dialectically, in the sense of the Socratic teaching method. In the Socratic method, the teacher constantly questions the students' basic assumptions and premises, plays the devil's advocate, and probes weak areas, using such techniques as invidious generalizations and counter-example (Anderson, 1977; Brown, 1977; Collins, 1977). The desired end-product is that the student will come to perform the teacher's functions for himself via self-interrogation. Although the sophisticated skills described by Collins (1977) are obviously not directly applicable to young slow-learning children, the basic principles underlying the approach are. We have begun at the very simple level of teaching the child to self-interrogate when faced with a certain class of problems (instructions, math problems, a laboratory task, etc.). The type of self-interrogation which we think might work is to provide the child with a routine set of questions to ask himself before proceeding, e.g., (a) stop and think! (b) do I know what to do (i.e., understand the instruction, both explicit and implicit)? (c) is there anything more I need to know before I can begin? and (d) is there anything I already know that will help me (i.e., is this problem in any way like one I have done before)?

We are currently attempting to train educable retarded children to follow instructions both verbal and written and to perform a variety of simple prose comprehension tasks, all in the context of a meaningful activity, like assembling a toy or following a recipe. In the course of these activities, they must deliberately and overtly pass through a self-interrogation routine like the one described above. We believe that devising simple systems for
elicit self-awareness and conscious control over one's own activities is an important form of training because the end-product is desirable in its own right, it should have transsituational applicability and it should improve both the child's cognitive and metacognitive skills and his feeling of personal competence and control.
References


Brown, A. L. Knowing when, where, and how to remember: A problem of meta-
cognition. To appear in R. Glaser (Ed.), Advances in instructional
psychology. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978,
in press.

Brown, A. L. Metacognitive development and reading. In R. J. Spiro, B.
Bruce & Wm. F. Brewer (Eds.), Theoretical issues in reading and comprehen-
press (b).

Brown, A. L. & Barclay, C. R. The effects of training specific mnemonics on
the metamnemonic efficiency of retarded children. Child Development,
1976, 47, 70-80.

Brown, A. L. & Campione, J. C. Training strategic study time apportionment
in educable retarded children. Intelligence, 1977, in press.

Brown, A. L. & Campione, J. C. Memory strategies in learning: Training
children to study strategically. In H. Pick, H. Leibowitz, J. Singer,
A. Steinschneider, & H. Stevenson (Eds.), Application of basic research

Brown, A. L., Campione, J. C., Bray, N. W. & Wilcox, B. L. Keeping track of
changing variables: Effects of rehearsal training and rehearsal prevention
in normal and retarded adolescents. Journal of Experimental Psychology,

Brown, A. L., Campione, J. C. & Murphy, M. D. Keeping track of changing
variables: Long-term retention of a trained rehearsal strategy by retarded
adolescents. American Journal of Mental Deficiency, 1974, 78, 446-453.

Brown, A. L., Campione, J. C. & Murphy, M. D. Maintenance and generalization
of trained metamnemonic awareness in educable retarded children. Journal


Campione, J. C. & Brown, A. L. Memory and metamemory development in educable retarded children. In R. V. Kail, Jr., & J. W. Hagen (Eds.),


Kennedy, B. A. & Miller, D. J. Persistent use of verbal rehearsal as a function of information about its value. Child Development, 1976, 47, 566-569.


Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C. & Brown, A. L. Recall of thematically relevant material by adolescent good and poor
Footnote

Copies of the stories with the corresponding idea units and their rated value are available from the first author upon request.
Table 1

Table 1. Mean proportion of correct recall as a function of Groups, Conditions, and Test Days: Maintenance Phase

<table>
<thead>
<tr>
<th>Test Days</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Condition</td>
<td>unprompted</td>
<td>unprompted</td>
<td>prompted</td>
</tr>
<tr>
<td>Young</td>
<td>Anticipation</td>
<td>.50</td>
<td>.48</td>
<td>.81</td>
</tr>
<tr>
<td></td>
<td>Rehearsal</td>
<td>.46</td>
<td>.50</td>
<td>.90</td>
</tr>
<tr>
<td></td>
<td>Label</td>
<td>.46</td>
<td>.58</td>
<td>.78</td>
</tr>
<tr>
<td>Old</td>
<td>Anticipation</td>
<td>.80</td>
<td>.72</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>Rehearsal</td>
<td>.74</td>
<td>.73</td>
<td>.84</td>
</tr>
<tr>
<td></td>
<td>Label</td>
<td>.60</td>
<td>.61</td>
<td>.67</td>
</tr>
</tbody>
</table>
Figure Captions

Figure 1. The proportion of items recalled as a function of Mental Age, Training Conditions and Test Phase. The data from both the pretest and posttests 1-3 are from Brown & Barclay (1976).

Figure 2. The proportion of idea units recalled as a function of Mental Age, Training Conditions and the Rated Importance of the Idea Units.
No. 1: Durkin, D. Comprehension Instruction--Where Are You?, October 1977. (ERIC Document Reproduction Service No. ED 146 566, 14p., HC-$1.67, MF-$0.83)

No. 2: Asher, S. R. Sex Differences in Reading Achievement, October 1977. (ERIC Document Reproduction Service No. ED 146 567, 30p., HC-$2.06, MF-$0.83)


No. 4: Jenkins, J. R., & Pany, D. Teaching Reading Comprehension in the Middle Grades, January 1978.

*Available only through ERIC*

**No. 1:** Halff, H. M. *Graphical Evaluation of Hierarchical Clustering Schemes,* October 1975. (ERIC Document Reproduction Service No. ED 134 926, 11p., HC-$1.67, MF-$0.83)

**No. 2:** Spiro, R. J. *Inferential Reconstruction in Memory for Connected Discourse,* October 1975. (ERIC Document Reproduction Service No. ED 136 187, 81p., HC-$4.67, MF-$0.83)

**No. 3:** Goetz, E. T. *Sentences in Lists and in Connected Discourse,* November 1975. (ERIC Document Reproduction Service No. ED 134 927, 75p., HC-$3.50, MF-$0.83)

**No. 4:** Alessi, S. M., Anderson, T. H., & Biddle, W. B. *Hardware and Software Considerations in Computer Based Course Management,* November 1975. (ERIC Document Reproduction Service No. ED 134 928, 21p., HC-$1.67, MF-$0.83)

**No. 5:** Schallert, D. L. *Improving Memory for Prose: The Relationship Between Depth of Processing and Context,* November 1975. (ERIC Document Reproduction Service No. ED 134 929, 37p., HC-$2.06, MF-$0.83)

**No. 6:** Anderson, R. C., Goetz, E. T., Pichert, J. W., & Halff, H. M. *Two Faces of the Conceptual Peg Hypothesis,* January 1976. (ERIC Document Reproduction Service No. ED 134 930, 29p., HC-$2.06, MF-$0.83)

**No. 7:** Ortony, A. *Names, Descriptions, and Pragmatics,* February 1976. (ERIC Document Reproduction Service No. ED 134 931, 25p., HC-$1.67, MF-$0.83)

**No. 8:** Mason, J. M. *Questioning the Notion of Independent Processing Stages in Reading,* February 1976. (Journal of Educational Psychology, 1977, 69, 288-297)

**No. 9:** Siegel, M. A. *Teacher Behaviors and Curriculum Packages: Implications for Research and Teacher Education,* April 1976. (ERIC Document Reproduction Service No. ED 134 932, 42p., HC-$2.06, MF-$0.83)


No. 16: Jenkins, J. R., & Pany, D. Curriculum Biases in Reading Achievement Tests, November 1976. (ERIC Document Reproduction Service No. ED 134 938, 24p., HC-$1.67, MF-$0.83)


No. 20: Kleiman, G. M. The Effect of Previous Context on Reading Individual Words, February 1977. (ERIC Document Reproduction Service No. ED 134 941, 76p., HC-$4.67, MF-$0.83)


No. 34: Bruce, B. C. *Plans and Social Actions*, April 1977.


No. 59: Mason, J. M. Reading Readiness: A Definition and Skills Hierarchy from Preschoolers' Developing Conceptions of Print, September 1977.


No. 61: Spiro, R. J., & Smith, D. Distinguishing Sub-Types of Poor Comprehenders: Overreliance on Conceptual vs. Data-Driven Processes, April 1978.


No. 77: Nash-Webber, B. L. Inference in an Approach to Discourse Anaphora, January 1978.


No. 83: Reynolds, R. E., Standiford, S. N., & Anderson, R. C. Distribution of Reading Time when Questions are Asked about a Restricted Category of Text Information, April 1978.


