JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY 34, 490–509 (1982)

Learning the Functional Significance of Mnemonic Actions: A Microgenetic Study of Strategy Acquisition

SCOTT G. PARIS, RICHARD S. NEWMAN, AND KELLY A. MCVEY

University of Michigan

How children learn to use memory strategies in a microgenetic investigation of learning and metacognition is examined. Seven- and eight-year-olds were given two memory trials with 24 pictures on each of 5 consecutive days. Days 1 and 2 were baseline, practice trials; Day 3 included strategy training; and Days 4 and 5 were unprompted tests of strategy maintenance. All children were taught how to label, rehearse, and group the pictures as well as to self-test their own memories and use blocked recall. Half of the children were shown the actions and told to do them; the other half received elaborated feedback on the usefulness and appropriateness of the techniques for remembering. The elaborated instructional group exhibited significantly greater recall, clustering, strategic study behavior, and metamemory regarding the mnemonic techniques than the other group. Path analysis provided evidence of a causal role of training and metacognitive awareness that mediated the use of sorting and higher recall. The study illustrates how learning the significance and utility of mnemonic techniques can influence children's acquisition of the tactics as self-controlled strategies. In addition, the microgenetic method affords patterns of data across time and experience that permit richer interpretations of strategy learning and memory development.

One of the most notable characteristics of cognitive development during childhood is the acquisition of cognitive strategies. From approximately 5 to 12 years of age children learn to direct their own attention, apportion study time, devise plans for communicating, monitor their own understanding while reading and listening, and so forth (Brown & DeLoache, 1978; Paris & Lindauer, 1982). A common thread among these diverse strategies for tackling and solving problems is the self-management of one's cognitive resources. An important developmental issue is to determine how children profit from practice and instruction

The authors express their gratitude to Craig Barclay for thoughtful comments on an earlier draft. We also thank the principals, David Aberdeen and Howard Walker, and the teachers of Burns Park and Stone Elementary Schools in Ann Arbor. Reprint requests should be sent to Scott G. Paris, Department of Psychology, 3406 Mason Hall, University of Michigan, Ann Arbor, MI 48109. Richard Newman is now at the State University of New York at Stony Brook.

0022-0965/82/060490-20$02.00/0
Copyright © 1982 by Academic Press, Inc.
All rights of reproduction in any form reserved.

and subsequently learn to plan, select, and regulate their behavior. More simply, what is the ontogeny of strategic thinking?

Research on memory development has provided some provocative answers to this question and has been a good "window" for investigating information processing in children (Kail & Hagen, 1977). Two reliable findings have emerged concerning children's acquisition of memory strategies. First, when given a deliberate memorization task, children younger than 7 years of age often do not generate mnemonic strategies spontaneously (Flavell, 1970). When given a second trial with explicit directions to use a particular strategy, though, their memory scores usually improve. The perplexing problem is that performance often declines on a third test of memory when explicit directions are absent. This pattern of results is described as a production deficiency for mnemonic strategies. The second reliable finding is that young children are relatively unaware of their own memory abilities, how task parameters influence memory, and how mnemonic actions for encoding and retrieval are linked to remembering (Kreutzer, Leonard, & Flavell, 1975). The lack of metamemory, or more generally, impoverished metacognition, has been hypothesized to underlie children's failures to produce appropriate cognitive strategies (Brown, 1978; Flavell, 1978). The purpose of the present study is to explore the relationship between metacognition and strategy learning, that is, how children's understanding of mnemonic techniques influences their use of self-controlled strategies to aid recall.

Many studies have addressed the relationship between metamemory and behavior, some with less success than others (see Cavanaugh & Perlmutter, 1982). Part of the problem appears to be the selection and operationalization of good measures of metamemory. However, training studies that emphasize the connection between mnemonic strategies and recall have revealed positive effects of increased awareness. For example, providing feedback to children about the usefulness of rehearsal as a memory aid enhanced the maintenance of the strategy on a later undirected trial (Kennedy & Miller, 1976). Watching a model successfully use memory strategies can also lead children to adopt and use the techniques (Borkowski, Levers, & Gruenenfelder, 1976).

The critical link may be the child's realization that there is a means–goal connection between the use of the strategy and the consequences for recall (Paris, 1978). Ringel and Springer (1980), for example, trained first, third, and fifth graders to use physical sorting techniques to aid recall of pictures. Four training conditions were employed; no instructions to sort pictures, instructions to sort only, and two conditions combining instruction with feedback about the sorting techniques. All instructed groups performed better than noninstructed control groups although there was an advantage for feedback plus instruction only for third graders. Recall levels of first and third graders also dropped considerably on tests

following training, a result consistent with other demonstrations of production deficiencies. The lack of effectiveness of training over trials, and for the first graders in particular, may indicate that young children require more elaborated interventions. Certainly, long-term training studies that emphasize the necessity, utility, and generalizability of mnemonic strategies have been most successful in teaching durable skills (Borkowski, Cavanaugh, & Reichert, 1978; Brown, Campione, & Barclay, 1979).

Despite the encouraging results of these studies to teach children mnemonic strategies, the role of metamemory in the acquisition of mnemonic strategies remains elusive. We think that this is due to two factors. First, the conceptual relationship between metamemory and strategy use has not been specified in detail. Since metamemory and memory performance both increase with age (see Flavell & Wellman, 1977), it was thought that simply making children more aware of person, task, and strategy variables might improve remembering. But teaching children to predict their memory abilities more accurately, demonstrating strategy use, and providing information about task demands appear to have modest effects on children's strategy learning. Those studies that enhance children's metamemory by providing informed feedback about the task goal and the utility of the strategy for accomplishing the goal seem to be most successful. The common denominator between developmental and training studies on memory strategies is that the mnemonic actions acquire personal significance and functional value. Mnemonic actions become mnemonic strategies because they link *will* and *skill* together and provide a rationale for using the action based on expectations of successful recall (Smirnov & Zinchenko, 1969). The present study tests this utilitarian view of the relationship between strategy awareness and strategy use.

The second impediment to investigating the relationship between metamemory and strategy learning has been the restricted methods employed. With the exception of some very good training studies, we have few long-term or longitudinal investigations of mnemonic strategy acquisition. Furthermore, there have been virtually no studies that provide repeated measurements of metamemory before, during, and after training. Repeated measurement of subjective awareness of the strategy is required in order to evaluate the effects of training on awareness directly. Further, the patterns of change and causal directions among awareness and action require multiple assessments. There is also little evidence on the reliability of metamemory reports or the extent to which children change their beliefs about strategy utility as a function of practice or instruction. Traditional one-shot, cross-sectional, static experimental designs may not be adequate to illuminate the evolving and dynamic relations between metamemory and strategy use.

In this study we demonstrate the value of a microgenetic analysis of strategy learning (cf. Vygotsky, 1978) for elucidating the role of meta-

memory in ameliorating production deficiencies. Our focus is on learning as a function of practice and teaching. In the present study, we obtained measures of children's perceptions of various mnemonically relevant actions in addition to actual memory performance. Measures were gathered over time and experience with the task to measure progressive patterns of change in learning. Our primary hypothesis was that a strategy is a constructed means–goals relationship (cf. Paris, 1978) that reflects understanding about the utility and significance of various optional actions. Children's beliefs about what they are doing and the value of their own behavior direct learning and supply the rationale for self-controlled use of strategies (see Paris & Lindauer, 1982).

In the present study we tested this hypothesis by asking children to study and recall 24 pictures twice each day for 5 days. On Day 3 all children were instructed to use mnemonic strategies of cumulative rehearsal, sorting, labeling, blocked recall, and self-testing. However, only half of the subjects (the *elaboration group*) received justification, rationale, and feedback about the utility and benefits of using the strategies while the other half (the *nonelaboration group*) received only demonstration and practice. We measured recall, clustering, strategic study behavior, and strategy understanding repeatedly during the 5 days of learning to test reciprocal patterns of behavioral and cognitive changes.

METHOD

Subjects

Thirty children from first- and second-grade classes in Ann Arbor, Michigan, participated in this study. The average chronological age was 87 months (range = 80–100). Four children were tested and subsequently replaced because they failed to meet the minimum recall criterion of at least five pictures recalled on each trial. The replaced subjects included two boys and two girls, one of each sex in each experimental group. The control group, which received nonelaborated instructions, included seven males and eight females. The group with elaborated instructions included eight males and seven females.

Materials

Five decks of 24 stimulus cards, one for each day of testing, were prepared by cutting out pictures of 120 common objects in magazines and pasting them on 10 × 15-cm index cards. The pictures were selected for their simplicity and representativeness of categories such as vegetables, wild animals, and furniture. There were four examples of six different categories per deck with different categories used in each of the five decks. The decks were arranged in seven random orders with four or five children receiving each order.

Two devices were used to assess children's metacognitive judgments about remembering. The first was a set of interview questions about children's intended use and perceived value of mnemonic strategies. Some questions were specifically relevant to the experimental task and others were more general. The data from three specific questions are reported in the results. The second device was a strategy-rating task modeled after Paris and Myers (1981) and Cox and Paris (Note 1). Each child was given a 5-point scale graphically represented by a serial ordering of different-sized rectangles along a line and asked to rate 10 mnemonic actions. The scale ranged from (1) "It really hurts remembering" through (3) "no effect" to (5) "It really helps remembering." Each of the 5 points on the scale was explained and all children readily understood the task. The 10 memory techniques were selected to include strategies that were potentially effective as well as noneffective, and trained as well as nontrained. They were sorting into taxonomic groups, rehearsal, self-testing, holding pictures to forehead, putting pictures under pillow at night, color cueing, making up a story about the pictures, alphabetizing, rhyming, and assigning a number to each picture. Each action was described and rephrased or demonstrated to ensure children's comprehension.

Procedure

Each child was seen on 5 consecutive days for approximately 30 min per day. On each day the child was asked to perform two study–recall trials with a deck of 24 randomly presented pictures. The children named each item before the study period to insure that they knew the pictures. They were then given 3 min to study the cards in any way they wanted. The cards were collected and a verbal recall task followed. The order of recalled items was recorded numerically on a score sheet. After a brief interval, this procedure was repeated a second time with the same set of cards. During the 3-min study period, the experimenter sat beside the child and recorded his or her study behavior on a checklist. The frequencies of five classes of overt behavior were recorded during each 15-sec interval as described in the Results section. The metacognitive questions and strategy ratings followed the recall tests on Days 2, 3, and 5.

The third day of testing began with a training session prior to the study–recall tasks. The training lasted approximately 8–10 min and consisted of giving the children information about five mnemonic strategies. Using the deck of cards for the third day's task, the experimenter demonstrated how to (1) put each picture into groups of similar items, (2) label each item, (3) cumulatively rehearse the pictures by group, (4) cover or close the eyes and self-test, and (5) recall the pictures by group. The strategies were presented as a set of different activities from which

the child could choose. No interdependence or joint operation of the strategies was taught. Half the children received only a demonstration of each strategy (nonelaborated condition) while the other group received the demonstration plus a brief explanation of the reasons why each strategy would aid in remembering the pictures (elaborated condition).

Thus, the main difference between groups was the inclusion of a brief *rationale,* specific to each of the five directed behaviors. Examples of the rationales for sorting and self-testing are

> "It's easier to remember things when you put them into smaller groups instead of one big group. That way you only have to remember a few things in each group instead of lots of things in one big group."

> "Covering your eyes and trying to remember the pictures will help you decide if you've studied enough and learned all the pictures. This will help you find out which pictures you already know and which ones you don't know. That will tell you if you should study some more."

All subjects were encouraged to use the directed strategies on the first trial of Day 3. The second trial was presented without encouragement or directions to use the strategies. After the first trial, all children received general, positive feedback. Additionally, the children in the elaboration group received feedback tailored to the actual study behavior they exhibited:

> "You did very well when you studied the way I told you. What exactly did you do, anyway? . . . Yes, good, you (experimenter described child's behavior, e.g., put them into groups, labeled the pictures, said the names over and over again, tested yourself, remembered them one group at a time). I guess (fill in the specific strategies just mentioned) helped you remember the pictures better, right? You could even do this for other things you wanted to remember. I think if you keep studying the way I taught you, you'll probably be able to remember most of the pictures next time—maybe even all of them."

On Days 4 and 5 neither group received any training and the tasks were presented as on Days 1 and 2. Thus, Days 4 and 5 were tests of strategy learning and maintenance under the two instructional conditions.

RESULTS

In the first four parts of this section we will report on findings involving (a) number of items recalled, (b) clustering, (c) study behavior, and (d) metacognitive judgments. The fifth part will integrate these measures in a structural model. Least-squares analyses of variance involving the four different measures were conducted with repeated measures. Cognizant of the issues raised by Winer (1971), Keppel (1973), and McCall and Appelbaum (1973) concerning possible violations of homogeneity assumptions and consequent bias in such designs, we evaluated the F ratios

that involved repeated factors against critical values that were conservatively corrected by the lower bound of the ε method (Greenhouse & Geisser, 1959). Twenty-six significant F's reported in the Results section involve repeated factors; 20 of these remain significant at least at $p < .05$ with the Greenhouse–Geisser correction. The other six are sufficiently close to satisfy estimated ε criteria which are not at a lower bound (see Box, 1954; McCall & Applebaum, 1973). For consistency and clarity, all the probability values reported below are based upon unadjusted degrees of freedom and critical values of F. Analyses for (a), (b), and (c) above are reported according to 5 days instead of 10 trials because the experimental design was by days. ANOVAs were performed on both trials and days factors and yielded identical results.

Number of Pictures Recalled

Each child studied and recalled 24 pictures on two trials each day. The average numbers of pictures recalled on each day for each group are shown in Fig. 1. Both groups recalled the most pictures on Day 3 when the instructions were given regarding strategies for remembering. However, recall of the pictures remained high on Days 4 and 5 only for the group receiving elaborated instructions that included an explanation for using the mnemonic strategies. The data in Fig. 1 were subjected to a group (2) × sex (2) × days (5) ANOVA with repeated measures. The main effects of group, $F(1, 26) = 6.40$, $p < .05$, and days, $F(4, 104) = 27.82$, $p < .01$, were significant as was the group × days interaction, $F(4, 104) = 7.64$, $p < .01$. There were no other significant effects for the number of items recalled. Newman–Keuls tests on the differences between means were conducted for the group × days interaction. Within

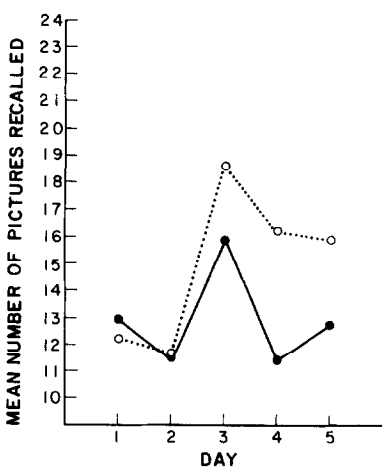

FIG. 1. Mean number of pictures recalled by group and days. ○, Elaboration group; ●, nonelaboration group.

the elaboration group, more pictures were recalled on Day 3 than any other day and recall was significantly higher on Days 4 and 5 than on Days 1 and 2 (p's $< .05$). In contrast, the group receiving no elaboration recalled significantly more items on Day 3 than any other day but the levels of recall did not differ among Days 1, 2, 4, and 5. Between-group differences were significant on Days 3, 4, and 5 but not before training.

Clustering

The degree to which children used taxonomic relationships to aid recall can be measured by several formulas. The adjusted ratio of clustering (ARC) was chosen as a good measure of children's ordered, organized recall because it is less confounded by the number of items recalled than other measures (Murphy, 1979). Furthermore, Murphy's (1979) simulation data with 24-item lists with six categories revealed ARC to be the best available measure. Certainly children's recall of pictures can be influenced by many factors, but one of our main interests in this study was to test the acquisition and maintenance of strategies, two of which (sorting and blocked recall) capitalized on the categorical relations among pictures. The equations for calculation of the ARC scores follow Roenker, Thompson, and Brown's (1971) original procedures, with the expected value of repetitions computed according to Frender and Doubilet (1974).

The mean ARC scores for each group, averaged over two recall trials each day, are shown in Fig. 2. Both groups clustered the most on Day 3 and it appears that ARCs remained higher on Days 4 and 5 for the elaborated group than the control group. A group (2) × sex (2) × days (5) ANOVA with repeated measures was performed on the data. There was a significant main effect of group, $F(1, 26) = 4.68$, $p < .05$, a significant effect due to days, $F(4, 104) = 24.19$, $p < .01$, and a significant

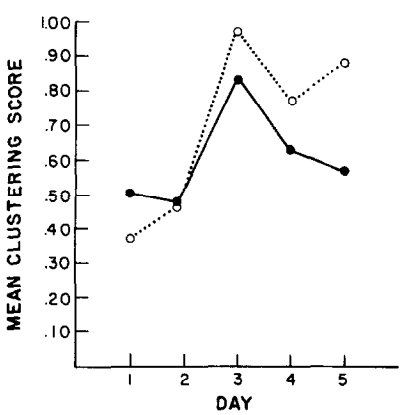

FIG. 2. Mean clustering scores (ARCs) by group and days. ○, Elaboration group; ●, nonelaboration group.

group × day interaction, $F(4, 104) = 4.54$, $p < .01$. Newman–Keuls tests revealed that there were no differences between groups on Days 1 or 2, but the elaboration group had significantly higher clustering scores on Days 3–5 (p's $< .05$). The nonelaboration group had ARC scores that were elevated significantly only on Day 3, while the elaboration group's ARC scores on Days 3–5 were all significantly higher than Days 1 and 2. Clearly the elaboration group increased and maintained high rates of categorical clustering among items. This finding precisely parallels the group × day interaction for numbers of items recalled.

Study Behavior

On each trial, subjects studied the pictures for 3 min. During this time the experimenter recorded the frequency of five classes of behavior during each 15-sec interval. The five types of behavior were (1) labeling and/or rehearsing by naming at least one of the pictures at least once, (2) physical sorting by putting pictures into discrete groups other than a single pile of 24 pictures, (3) self-testing by covering the picture, closing one's eyes, or turning away, (4) distracted or off-task behavior, and (5) passive looking or visual inspection. These types of study behavior are commonly recorded with high reliability (cf. Wellman, Ritter, & Flavell, 1975). The first four activities were not mutually exclusive of one another. For instance, a child could engage in any combination, including all four, in any one time interval. Multiple occurrences of the same activity in a 15-sec interval were given only one check so that the range of activity frequency was 0–12 for each. The fifth behavior, mere looking, was checked only if the child engaged in it for a full 15-sec period of time to the exclusion of all other behaviors. Separate analyses were performed on the first four behaviors taken together and the looking behavior by itself to maintain statistical independence.

The observed frequencies of each of the four independent types of study behavior were entered into a group (2) × sex (2) × study behavior (4) × days (5) ANOVA with repeated measures. Group, study behavior, and days had significant main effects, $F(1, 26) = 5.98$, $p < .05$, $F(3, 78) = 20.30$, $p < .01$, and $F(4, 104) = 30.48$, $p < .01$, respectively. Group × days, study behavior × days, and group × days × study behavior were significant interactions, $F(4, 104) = 6.31$, $p < .01$, $F(12, 312) = 6.48$, $p < .01$, and $F(12, 312) = 2.85$, $p < .001$, respectively.

To determine the sources of differences within the triple interaction, two-way ANOVAs (group × day) were conducted for each of the four study behaviors. No significant main effects or interactions were found for distracted behavior. However, there were significant effects in each of the other three types of study behaviors which can be seen clearly in Fig. 3. Within the labeling and rehearsal category, there was only a significant main effect of days, $F(4, 112) = 12.80$, $p < .01$. There was

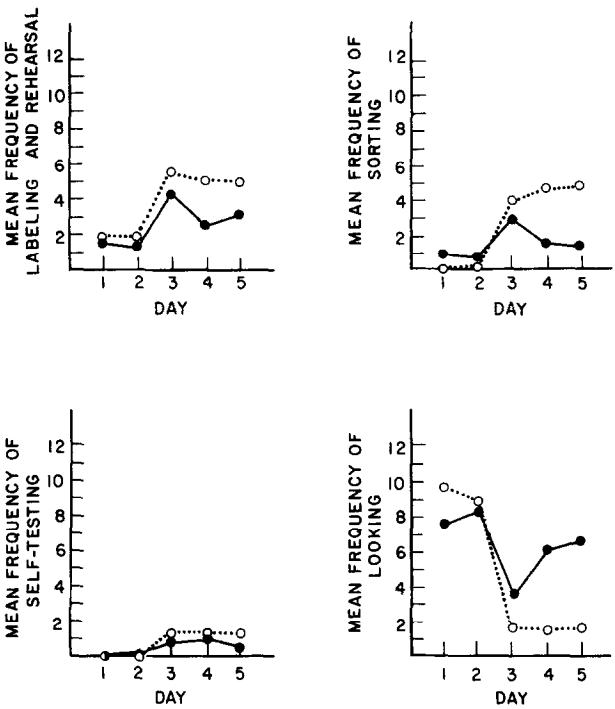

FIG. 3. Mean frequencies of strategic study behavior by group and days. ○, Elaboration group; ●, nonelaboration group.

significantly more labeling and rehearsal on Days 3–5 than on either Days 1 or 2. Self-testing behaviors also showed only an increase in frequency by days, $F(4, 112) = 6.66$, $p < .01$ (again due to more self-testing on Days 3–5 than either Day 1 or 2). Sorting behavior revealed main effects due to group, $F(1, 28) = 6.62$, $p < .05$, and days, $F(4, 112) = 16.51$, $p < .01$, as well as a significant group × days interaction, $F(4, 112) = 7.59$, $p < .01$. As Fig. 3 illustrates, only the elaboration group continued to use physical sorting as a study behavior on Days 4 and 5 to the same degree as on Day 3. Newman–Keuls tests revealed that subjects in the elaboration group used physical sorting as a study behavior significantly more often than the nonelaboration group on days 3–5 (p's $< .05$). Although not statistically significant, the data for labeling and rehearsal reveal a similar interactional trend as sorting.

As an interesting converse to the increase in strategic sorting, subjects in the elaboration group *decreased* the amount of mere looking during the study period (see Fig. 3). A group (2) × sex (2) × days (5) ANOVA with repeated measures on looking behavior showed a main effect of days, $F(4, 104) = 34.98$, $p < .01$, and a group × days interaction, $F(4,$

104) = 10.62, $p < .01$. Newman–Keuls tests revealed that subjects in the elaboration group engaged in significantly less looking on Days 3–5 than the other group despite a significantly higher incidence on Day 1. Whereas the children in the elaboration group maintained their lower levels of mere looking on Days 4 and 5, the other children reverted to their pretraining levels. In summary, children in the elaboration group maintained significantly higher levels of strategic sorting behavior, with lower levels of passive behavior, following training than nonelaboration children.

Physical sorting, then, is the one strategy whose increased use was maintained as a result of the elaborated training. In an effort to examine differential effectiveness of study behavior on actual recall, multiple regression analyses were performed. Recall was regressed on (a) sorting, (b) labeling and rehearsal, and (c) self-testing for each of the 5 days. On Days 1–3 the study behaviors did not explain a significant amount of the variance in recall. This was as expected, including Day 3 when *both* groups of children significantly increased their study behavior, clustering scores, and recall. On Days 4 and 5, however, when maintenance of strategies was examined, multiple regressions produced significant overall F's. On Day 4, $F(3, 26) = 6.94$, $p < .005$, $R^2 = .445$, and there were significant contributions from sorting (beta weight = .393, $p < .05$) and labeling and rehearsal (beta weight = .425, $p < .01$). On Day 5, $F(3, 26) = 4.97$, $p < .01$, $R^2 = .365$, and there were significant contributions from sorting (beta weight = .386, $p < .05$) and labeling and rehearsal (beta weight = .335, $p < .05$). These findings indicate a relationship after training between two types of study behavior and recall that did not exist before or during training.

Metacognitive Judgments

Children's responses to interview questions and strategy evaluations were analyzed as measures of personal understanding about strategy usefulness. Indices of strategy awareness were constructed from children's interview data that were collected immediately after the recall trials on Days 2, 3, and 5. Answers to the following three questions were analyzed: (1) "How did you try to remember the pictures today?" (2) "What plan worked best?" (3) "What's the very best thing to do to help you remember?" On each of the three questions, answers were categorized three times, according to whether or not the child mentioned (a) sorting, (b) labeling or rehearsal, or (c) self-testing. There were other answers given; however, it was these three categories that were most frequently represented. The number 1 was assigned if the child did not mention a strategy; the number 2 if he or she did. Three ordinal indexes were formed by summing over the three questions; scores ranged from 3 to 6. They provide a measure of children's likelihood of spontaneously

mentioning the strategy in response to questions of what they considered to be the most effective way to remember the pictures.

The group means and standard deviations on each of these indexes across the three days are shown in Table 1. A group (2) × sex (2) × strategy (3) × days (3) repeated measures ANOVA was performed on the strategy awareness data. There are main effects of group, $F(1, 26) = 11.66$, $p < .01$; strategy, $F(2, 52) = 4.07$, $p < .05$; and days, $F(2, 52) = 14.19$, $p < .01$. The following interactions were also significant: strategy × group, $F(2, 52) = 4.17$, $p < .05$; days × strategy, $F(4, 104) = 4.11$, $p < .01$; and days × group × strategy, $F(4, 104) = 2.97$, $p < .05$). To investigate the three-way interaction, a separate day × group ANOVA was performed for each strategy. Only sorting revealed significant results. There were main effects of group, $F(1, 28) = 13.35$, $p < .01$; day, $F(2, 56) = 15.07$, $p < .01$; and a group × day interaction, $F(2, 56) = 6.32$, $p < .01$. This latter interaction was examined with Newman–Keuls comparisons, and it was observed that the elaborated group mentioned sorting on Days 3 and 5 more often than the nonelaborated group. Additionally, only the elaborated group exhibited a change over days, with an increase in awareness from Day 2 to both Days 3 and 5.

Other data concerning metacognitive judgments were gathered on Days 2, 3, and 5 in the form of strategy ratings that followed recall trials. These evaluations were done on a 5-point scale for 10 different strategies. The means and standard deviations for each of the strategy ratings over the 3 days are displayed in Table 2.

A sex (2) × group (2) × strategy (10) × days (3) repeated measures ANOVA was performed on the strategy ratings. There were main effects of both strategy, $F(9, 234) = 47.47$ and days, $F(2, 52) = 6.75$ (both p's $< .01$). Significant interactions were found between days and strategy,

TABLE 1
MEAN SCORES OF STRATEGY AWARENESS BY GROUP AND DAYS

	Strategy								
	Sorting			Labeling or rehearsal			Self-testing		
Group	Day: 2	3	5	2	3	5	2	3	5
Elaboration	3.13	4.80	4.87	4.07	4.13	3.93	3.33	3.33	3.47
	(.38)	(1.24)	(1.22)	(1.20)	(1.01)	(.96)	(.60)	(.95)	(.94)
Nonelaboration	3.20	3.60	3.53	3.60	4.06	3.33	3.20	3.80	3.47
	(.42)	(.87)	(.71)	(.92)	(.85)	(.39)	(.28)	(1.14)	(.92)

Note. Scores are from three questions about specific strategy use. Scores range from 3, "relatively unaware," to 6, "relatively aware." Numbers in parentheses are standard deviations.

TABLE 2
MEAN STRATEGY RATINGS BY GROUP AND DAYS

	Elaboration group			Nonelaboration group		
Strategy	Day 2	Day 3	Day 5	Day 2	Day 3	Day 5
Sorting	2.7	4.9	4.7	3.7	4.7	4.4
	(1.4)	(.3)	(.6)	(1.3)	(.5)	(.9)
Rehearsal	4.7	4.7	4.5	4.3	4.4	4.4
	(.7)	(.6)	(.7)	(1.3)	(.8)	(1.2)
Self-testing	3.1	4.3	4.0	3.6	4.4	3.4
	(1.4)	(.6)	(.9)	(1.3)	(.6)	(1.7)
Pictures to forehead	2.0	1.5	2.2	1.9	1.7	1.3
	(1.4)	(1.1)	(1.7)	(1.4)	(1.2)	(1.0)
Pictures under pillow	1.4	1.4	1.4	1.9	1.7	1.3
	(.7)	(.8)	(1.1)	(1.1)	(1.2)	(.5)
Color cueing	2.9	3.8	3.0	2.8	3.4	3.4
	(1.4)	(1.1)	(.9)	(1.4)	(1.1)	(1.2)
Story-telling	3.7	3.9	2.8	2.7	2.9	2.9
	(1.2)	(1.4)	(1.5)	(1.2)	(1.3)	(1.4)
Alphabetizing	4.3	3.7	3.5	3.5	3.9	3.9
	(.8)	(1.1)	(1.3)	(1.1)	(1.3)	(1.0)
Rhyming	3.6	3.3	2.8	3.5	3.7	3.7
	(1.1)	(1.0)	(1.4)	(1.4)	(1.1)	(1.3)
Numbering	3.9	3.8	3.4	3.5	3.3	3.3
	(1.1)	(1.1)	(1.1)	(.9)	(1.2)	(1.0)

Note. Scores range from 1, "very poor strategy," to 5, "very good strategy." Numbers in parentheses are standard deviations.

$F(18, 468) = 3.94$, $p < .01$ and among days, strategy, and group, $F(18, 468) = 2.26$, $p < .01$. To investigate the three-way interaction, a separate day × group ANOVA was performed for each strategy. Self-testing showed a day effect, $F(2, 56) = 6.46$, $p < .01$ with a significant increase in strategy rating from Day 2 to Day 3 and a decrease from Day 3 to Day 5, both p's $< .05$ with Newman–Keuls comparisons. Sorting showed a day effect, $F(2, 56) = 28.07$, $p < .01$ and a day × group interaction, $F(2, 56) = 5.08$, $p < .01$. Newman–Keuls comparisons here revealed that the elaborated group rated sorting less highly on Day 2 than the nonelaborated group. Both groups significantly increased their ratings of sorting from Day 2 to both Days 3 and 5. Color cueing was the only other strategy that contributed to the three-way interaction. The rating of color cueing showed a main effect of days, $F(2, 56) = 3.78$, $p < .05$ due to improvement between Days 2 and 3.

Training seems, then, to have had an effect on most children's awareness of sorting as an effective study behavior, but especially among those in the elaborated training condition. Did this belief that was fostered through training have an effect on the children's subsequent recall? Table

TABLE 3
RECALL AS A FUNCTION OF AWARENESS AND TRAINING

	Pretraining recall		Posttraining recall	
Awareness of sorting	Low	High	Low	High
Absent	12	14	11	2
Present	3	1	4	13

Note. Numbers of subjects are shown in each cell. Pretraining recall is averaged over Days 1 and 2 and posttraining recall is averaged over Days 4 and 5. High and low recall refers to scores above or below the median. Awareness measures are derived from interview questions on Days 2 and 5.

3 presents the number of children above and below the median number of recalled pictures in relation to their awareness of sorting. Awareness was measured by the interview questions on Day 2 (pretraining) and Day 5 (posttraining). Pretraining recall was averaged from Days 1 and 2; posttraining from Days 4 and 5.

It can be seen that 17 of the 30 children did express some belief in the goodness of sorting following training. Of these 17, 13 did well at recall. Of the 13 children who did not express an awareness of sorting as an effective strategy, 11 did poorly at recall (Fisher's exact $p = .0013$). In other words, there was a significant relationship between belief and recall after training. Furthermore, it should be pointed out that of the 17 children who expressed a belief in sorting on the last day of the experiment, 12 were in the elaborated training group, whereas 10 of the 13 "unaware" children were in the nonelaborated group (Fisher's exact $p = .0127$). In other words, there was a significant relationship between training condition and sorting awareness that occurred only after training.

Structural Modeling

The microgenetic design of the study provided multiple and repeated measures over the 5-day period. Analyses of number of items recalled, clustering, study behavior, and metacognitive judgments indicate that during the period after training, children in the elaboration group behaved differently than the other children in several ways. The common pattern of these findings supports our hypotheses about the links among metacognition, strategy usage, and memory performance and invites further analysis with structural modeling techniques. Preliminary factor analysis did not provide evidence of general, across-strategy factors of metacognitive judgment or study behavior. Given the sample size and experimental, time-ordered design, structural modeling with hierarchical path analysis is a heuristic method for exploring the interrelationships among variables over time (see Kenny, 1979; Sanders, 1980).

We have conceptualized the role of metamemory in strategy development as belief leading to action and recall. In the path analytic model we propose that the training condition predicts the child's metacognitive judgments, which in turn predict study behavior and actual memory performance. The model takes into account the time sequence of the experimental procedure. That is, posttraining metacognitive judgments were represented both by children's responses to interview questions and the strategy ratings from the end of Day 3. Study behavior and recall performance were from Days 4 and 5. It should be noted that the results were consistent across these two measures of metacognitive judgments and across the two posttraining days. For simplicity, we report path coefficients from a model employing measures derived from the interview on Day 3, along with study behavior and recall on Day 4.

The path analysis was performed with eight ordinary least-squares regressions. The three metacognitive judgments were first regressed on training condition; these simply produced zero-order correlations. Next, the three study behaviors were regressed on the previous day's metacognitive judgments and the child's training condition. Clustering was regressed on the three study behaviors and training condition. Finally, recall was regressed on clustering, the three study behaviors, and training condition. We report path coefficients, which are beta weights from the various regression equations, and R^2 values for the overall prediction of recall.

Figure 4 contains our hypothesized model and the initial path coefficients. The $R^2 = .654$ for the five-predictor regression of the average number of items recalled on Day 4. In other words, 65.4% of the variance in recall is explained by training condition, clustering, and the three study behaviors. Nonsignificant coefficients were "trimmed" from the model at a $p < .05$ level, and the regressions were rerun. The new path coefficients are shown in Fig. 5. The R^2 drops slightly to .650 for this three-predictor regression of recall. Here, 65.0% of the variance in recall is explained by training condition, clustering, and the labeling and rehearsal study behavior.

Ordinarily, a small sample size, along with a relatively large number of predictors, would cause concern in interpreting the results of multiple regression, and hence path analysis. In the present case, the very slight drop in R^2 (from .654 to .650) that occurs by selectively eliminating predictors in the model implies that the amount of variance we are accounting for is not greatly inflated. Additionally, the largest number of predictors in any one regression in the final model, that is, three, satisfies the guideline of having at least a 10:1 ratio of number of subjects per predictor (see Cohen & Cohen, 1975).

It can be seen from Fig. 5 that the training condition was directly associated with posttraining awareness of the efficacy of sorting as a

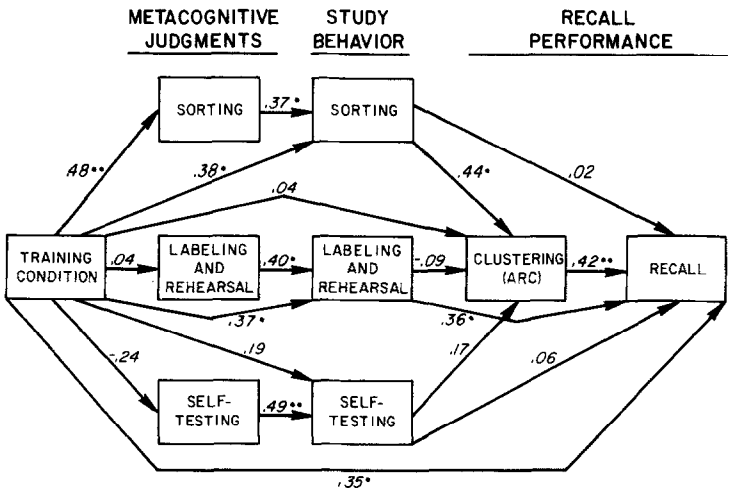

FIG. 4. Full causal model of metacognitive judgments, study behavior, and recall performance. Path coefficients are for the initial Day 3–Day 4 posttraining model. $*p < .05$, $**p < .01$.

mnemonic strategy. Use of sorting and labeling and rehearsal were predicted by training condition and strategy awareness, jointly. For all three strategies those children who showed *awareness* on Day 3 employed that particular strategy *behavior* on Day 4, regardless of training condition. Those children who sorted on Day 4 also tended to cluster their recall regardless of other study behaviors or training condition. Their clustering was positively associated with recall. Those children who engaged in

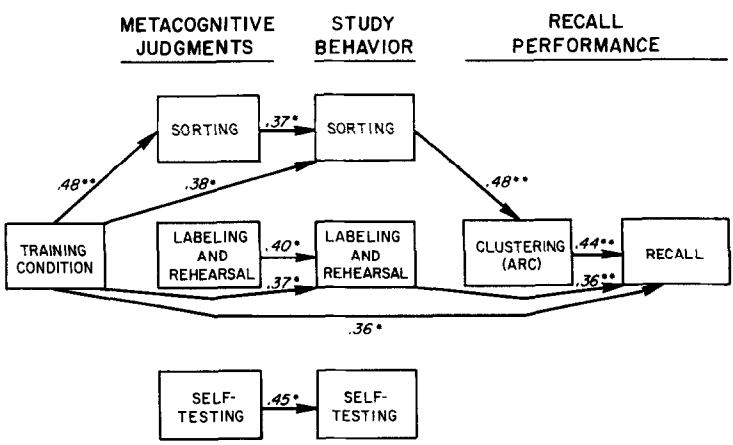

FIG. 5. Reduced causal model of metacognitive judgments, study behavior, and recall performance. Path coefficients are for the final Day 3–Day 4 posttraining model. $*p < .05$, $**p < .01$.

rehearsal also tended to do well in their recall, regardless of their clustering scores, their other study behaviors, or the training condition. Given the natural relatedness between physical sorting during study and taxonomic clustering in free recall, it is not surprising that the effect of sorting on recall was mediated through clustering in the model, whereas the effect of labeling and rehearsal was direct. Finally, the training had a direct effect on final recall, over and above the effect mediated through clustering and the use of the three strategies.

The path analysis reveals that the elaborated training had multiple and direct effects upon the children; on their awareness of sorting, on their use of sorting and rehearsal strategies, and on their actual recall. Metacognitive judgments predicted strategy use, regardless of which training condition the child was in. The elaborated training seemed to convince the child to both sort and rehearse, although it did not have much impact on self-testing. However, those who reported a belief in the strategy as a good means of remembering did, in fact, employ it.

DISCUSSION

Two training conditions were embedded in this microgenetic study of children's learning about memory strategies. One condition included demonstrations and directions to use particular mnemonic strategies like sorting, self-testing, and labeling and rehearsal. The other condition elaborated upon the directions by providing explanations and feedback regarding the mnemonic value of these actions. The empirical question was whether learning and maintenance of the strategies would be different for children in the two conditions. The results revealed a distinct advantage to children in the elaborated condition. They recalled more, clustered more, and sorted more strategically during training and both days afterwards. Clearly, better learning resulted when children understood the utility and significance of using mnemonic strategies. We hasten to add, though, that one strategy was primarily responsible for better recall. Physically grouping pictures into related piles was the most effective strategy. This strategy may have been more novel, motorically distinct, and compatible with the taxonomic nature of the stimuli than either rehearsal, labeling, or self-testing. As the path analyses illustrated, increased evaluation of the utility of *each* strategy led to more frequent use of the mnemonic actions on subsequent trials. Recall levels, however, were associated most strongly with sorting (as mediated through clustering) as well as labeling and rehearsal. A task that relied on different strategies for successful recall might lead to a different pattern of results. The important demonstration here is that children's evaluations of the functional utility of mnemonic strategies changed with instruction and predicted subsequent study behavior and recall patterns.

It seems reasonable to surmise that children's beliefs about the usefulness of the mnemonic tactics provided rationales for the subsequent use (or nonuse) of the strategies. Demonstrations of the mnemonic actions may not be sufficient to persuade children that the tactics are worthwhile or effective. Explanations of the means–goals connection and informative feedback about the linkage may be necessary for young children to adopt the strategies as better plans than their customary habits. The evidence for the mediating role of personal beliefs that coordinate means and goals comes from children's answers to interview questions and their subjective ratings of strategy utility. These metacognitive judgments, although susceptible to distortion, corroborate the behavioral data and converge on the same conclusion: beliefs influence learning. This statement may seem intuitive or trivial but past theories and methods of memory development have often omitted the roles of subjective perceptions of one's own behavior in learning. The constructive coordination of how, when, and why to use memory strategies requires more than a cognitive–developmental competence; it also requires a motivational disposition to behave accordingly. This motivation may be due in part to the subject's metacognitive evaluations of strategy usefulness. Subsequent plans to employ the strategy appear to be predicated on the subject's *understanding of the value of the skill* and not solely on prior success or behavioral compliance.

There are direct theoretical and methodological implications of this study for research in many domains. We need to study children's beliefs about their own cognitive states and abilities in a continuous fashion to chart how learning and development merge in self-controlled cognitive strategies. Only in this way can we measure the functional, reciprocal, and dynamic relationships between behavior and metacognition as they relate directly to learning and development. It seems to us that the current emphasis on skill theory (e.g., Fischer, 1980) reaffirms three historically important yet neglected concepts in children's learning—habit, purpose, and belief. In order to understand the ontogeny of cognitive strategies we need to assess the child's history of habits and actions, the task's meaningfulness and familiarity, and the significance of the instructed actions to a naive subject.

This study also demonstrates how a short-term, longitudinal, experimental design can be used to illuminate causally linked changes during learning. Brown (1982), drawing on Vygotsky's work, has reemphasized the advantages of a microgenetic approach for "defossilizing" skills and analyzing detailed changes during learning. Our method also borrows heavily from Vygotsky's principles of (a) inter-to-intrapsychological planes of functioning and (b) mediated tools for mnemonic goals. Researchers must devise new paradigms that include practice, time, repeated measurements, and converging evidence for the use of mnemonic

actions under varying task demands. A developmental focus is required for conceptualization and methodology. We offer this study as initial evidence of the power of a microgenetic method and the usefulness of assessing children's purposes, beliefs, and strategy evaluations as mediators of skill learning and strategy acquisition.

REFERENCES

Borkowski, J., Cavanaugh, J., & Reichart, G. Maintenance of children's rehearsal strategies: Effects of amount of training and strategy form. *Journal of Experimental Child Psychology*, 1978, **26**, 288–298.

Borkowski, J. G., Levers, S. R., & Gruenenfelder, T. A. Transfer of mediational strategies in children: The role of activity and awareness during strategy acquisition. *Child Development*, 1976, **47**, 779–786.

Box, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems II. Effects of inequality of variance and of correlation between errors in the two-way classification. *Annals of Mathematical Statistics*, 1954, **25**, 484–498.

Brown, A. L., & DeLoache, J. S. Skills, plans, and self-regulation. In R. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Erlbaum, 1978.

Brown, A. L. Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology*. Hillsdale, N.J.: Erlbaum, 1978.

Brown, A. L. Learning and development: The problems of compatibility, access, and induction. *Human Development*, 1982, **25**, 89–115.

Brown, A. L., Campione, J. C., & Barclay, C. R. Training self-checking routines for estimating test readiness: Generalization from list learning to prose recall. *Child Development*, 1979, **50**, 501–512.

Cavanaugh, J. C., & Perlmutter, M. *Metamemory: A critical examination. Child Development*, 1982, **53**, 11–28.

Cohen, J., & Cohen, P. *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, N.J.: Erlbaum, 1975.

Cook, T. D., & Campbell, D. T. *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally, 1979.

Fischer, K. W. A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 1980, **87**, 477–531.

Flavell, J. H. Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior*, (Vol. 5). New York: Academic Press, 1970.

Flavell, J. H. Metacognitive development. In J. M. Scandura & C. J. Brainerd (Eds.), *Structural/process theories of complex human behavior*. The Netherlands: Sijthoff & Noordoff, 1978.

Flavell, J. H., & Wellman, H. M. Metamemory. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition*. Hillsdale, N.J.: Erlbaum, 1977.

Frender, R., & Doubilet, P. More on measures of category clustering in free recall—although probably not the last word. *Psychological Bulletin*, 1974, **81**, 64–66.

Greenhouse, S. W., & Geisser, S. On methods in the analysis of profile data. *Psychometrika*, 1959, **24**, 95–112.

Kail, R. V., & Hagen, J. W. (Eds.), *Perspectives on the development of memory and cognition*. Hillsdale, N.J.: Erlbaum, 1977.

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.

Kenny, D. A. *Correlation and causality.* New York: Wiley, 1979.

Keppel, G. *Design and analysis: A researcher's handbook.* Englewood Cliffs, N.J.: Prentice–Hall, 1973.

Kreutzer, M. A., Leonard, C., & Flavell, J. H. An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development,* 1975, **40** (1, Serial No. 159).

McCall, R. B., & Appelbaum, M. I. Bias in the analysis of repeated-measures designs: Some alternative approaches. *Child Development,* 1973, **44,** 401–415.

Murphy, M. D. Measurement of category clustering in free recall. In C. R. Puff (Ed.), *Memory, organization, and structure.* New York: Academic Press, 1979.

Paris, S. G. Coordination of means and goals in the development of mnemonic skills. In P. A. Ornstein (Ed.), *Memory development in children.* Hillsdale, N.J.: Erlbaum, 1978.

Paris, S. G., & Lindauer, B. K. The development of cognitive skills during childhood. In B. Wolman (Ed.), *Handbook of Developmental Psychology.* Englewood Cliffs, N.J.: Prentice–Hall, 1982.

Paris, S. G., & Myers, M. Comprehension monitoring in good and poor readers. *Journal of Reading Behavior,* 1981, **13,** 5–22.

Ringel, B. A., & Springer, C. J. On knowing how well one is remembering: The persistence of strategy use during transfer. *Journal of Experimental Child Psychology,* 1980, **29,** 322–333.

Roenker, D. L., Thompson, C. P., & Brown, S. C. Comparison of measures for the estimation of clustering on free recall. *Psychological Bulletin,* 1971, **76,** 45–48.

Sanders, D. Path analysis/causal modeling. *Quality and Quantity,* 1980, **14,** 181–204.

Smirnov, A. A., & Zinchenko, P. I. Problems in the psychology of memory. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology.* New York: Basic Books, 1969.

Vygotsky, L. S. *Mind in society.* Cambridge, Mass.: Harvard Univ. Press, 1978.

Wellman, H. M., Ritter, K., & Flavell, J. H. Deliberate memory behavior in the delayed reactions of very young children. *Developmental Psychology,* 1975, **11,** 780–787.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw–Hill, 1971.

REFERENCES NOTES

1. Cox, G. L., & Paris, S. G. *The nature of mnemonic production deficiencies: A lifespan analysis.* Paper presented at the biennial meeting of the Society for Research in Child Development, San Francisco, 1979.

RECEIVED: August 17, 1981; REVISED: December 15, 1981; February 16, 1982.