

Verbal Labeling, Rehearsal, and Short-term Memory¹

JOHN W. HAGEN, JOHN A. MEACHAM, AND GARY MESIBOV

University of Michigan

The effects of verbal labeling in a serial position short-term memory (STM) task were investigated in two studies. In the first, 32 children at each of the grades 4, 6, and 8 were tested. In the second, 40 college students were tested. The stimuli consisted of eight cards with a central and an incidental picture on each. Half of the subjects overtly labeled the central pictures as they were presented; half did not. In Study I there were eight trials; in Study II there were additional trials in which stimuli contained only the central pictures. In Study I labeling did not affect overall performance; in Study II labeling was found to decrease performance significantly. For both studies labeling had differential effects at the primacy and recency portions of the serial-position data. There was no effect of one versus two stimuli per card in Study II. The findings were discussed in terms of issues concerning verbal mediation and rehearsal strategies. A developmental model for the processes involved in serial-position STM was presented.

Performance on a variety of cognitive tasks has been demonstrated to be affected by induced verbal labels, but the relation is not a simple one (Kendler, 1963). Recent investigations of short-term memory (STM) have shown that verbal labeling has differential effects at different chronological ages (CA) in children and have also helped clarify the nature of the mediating mechanisms (Hagen & Kingsley, 1968; Flavell, Beach, & Chinsky, 1966). The mediational deficiency hypothesis (Reese, 1962) has often been invoked to explain why labeling did not facilitate performance in children at young ages. The present study pursues some of the questions left unanswered by this hypothesis.

The mediational deficiency hypothesis states that there is a stage in development during which verbal responses are present but do not serve as mediators. Flavell and his colleagues (Flavell, Beach, & Chinsky,

¹ This research was supported by Grant Number 1 PO1 HD 01368-04 from the National Institute of Child Health and Human Development. The second and third authors were United States Public Health Service trainees while the studies were conducted. Thanks are given to the faculty of the St. Thomas School, Ann Arbor, Michigan, for its friendly cooperation. Author Hagen's address: Department of Psychology, University of Michigan, Ann Arbor, Michigan 48104.

1966; Keeney, Cannizzo, & Flavell, 1967) have argued that there is a production deficiency in young children; if labels are produced, either spontaneously or under experimental inducement, then mediation occurs and performance is facilitated, at least in STM tasks. Hagen and Kingsley (1968) studied verbal labeling effects in serial STM tasks and found that at certain ages there is a facilitating effect, but at either younger or older ages no such effect occurs. However, serial-position analysis of the data demonstrated that there was a striking Serial Position \times Verbal Labeling \times Age interaction effect. At the recency portion of the curve, performance was consistent across the age range from six to ten years: labeling facilitated performance and there was no improvement with CA. However, the picture was different at the primacy portion: performance improved with CA, but by age ten verbal labeling resulted in significantly lower performance than no labeling.

We hypothesized that by age ten, children have developed strategies which consist of rehearsing the names of the items already exposed during the intertrial intervals. Labeling disrupts these strategies by interfering with rehearsal. Since rehearsal facilitates primacy learning (Postman, 1964), and labeling inhibits rehearsal, labeling reduces primacy performance. Labeling heightens the saliency of recency items, perhaps by placing the label in immediate memory store; hence, performance is facilitated for these items. Several studies by other investigators support the theory that a serial-position STM task involves two distinct memory processes: very short-term and longer term (Kausler, 1966; Ellis & Hope, 1968; Glanzer & Cunitz, 1966).

The two studies reported here tested the following hypotheses concerning labeling effects: (a) At older CA levels labeling should result in a deficit in STM performance, since rehearsal strategies which facilitate performance are hindered; (b) Serial-position data should show the deficit due to labeling on all but the most recently presented positions, which are facilitated by labeling.

In each study a measure of incidental recall was also included. Earlier studies indicated that such a measure provides an index of efficiency of task performance under various experimental conditions (Maccoby & Hagen, 1965; Hagen, 1967; Hagen & Sabo, 1967). The difference between the central and incidental STM measures has been shown to increase with increasing CA. Central task performance shows steady improvement, while incidental task performance either remains constant or declines. Further, the central measure correlates positively with standard measures of intelligence while the incidental measure shows a slightly negative correlation or no association with IQ.

STUDY I

The CA range of the earlier study (Hagen & Kingsley, 1968) was extended to include children up through 14 years. At each grade level, half the sample was required to label the stimuli in the serial-position STM task, and half did not label overtly. In addition, incidental stimulus items were included for both label and no label conditions.

Method

Subjects. The subjects were 96 boys and girls from grades 4, 6, and 8 of St. Thomas School (a parochial school in Ann Arbor, Michigan). CA ranges were: Grade 4, 9–10; Grade 6, 11–12; Grade 8, 13–14. Subjects were from upper middle-class families and were above average in intelligence. Mean scores on the California Test of Mental Maturity (CTMM) were as follows: Grade 4, 111.3; Grade 6, 115.1; and Grade 8, 118.2.

Stimuli and test materials. In both the no label and the label conditions, the subjects were presented a series of eight cards, one at a time. Each card, 3 × 5 in. in size, had a black line drawing of an animal and a household object, just touching. These were paired in the following manner: fish-telephone, cat-lamp, horse-chair, camel-television, monkey-book, bear-cup, dog-table, deer-clock. The cue for the central test was given by means of a board, 7 × 9 in., on which all eight of the animals appeared in a circular design.

For the measure of incidental recall an eight-page booklet was used. On each page appeared one of the cards, with the animal in the original position, but with the object removed. In addition, there was a second board with all eight household objects arranged in rows upon the board.

Procedure. The subjects were taken one at a time from the classroom to the testing room. The experimenter and the subject sat in two student desks facing each other. The experimenter showed the board with all the animals, and said:

We're going to play a game with these animals. I have these same animals cut out and pasted onto these cards—you don't have to pay attention to those other pictures (here, the deck of eight cards was shown). The way the game works is like this—I'll show you one of the cards, *and you tell me the name of the animal*, and then I'll lay the cards down like this in a row (the experimenter demonstrates by laying three or four cards face down in a row on the desk). Then I'll point to one of the animals (on the large board) and we'll see if you can find where that animal is in the row, OK?

The only difference between the two experimental conditions was the italicized phrase above: "and you tell me the name of the animal." All subjects appeared to understand readily what was required.

Presentation proceeded from the subject's left to his right. The experimenter held each card in view for approximately 2 sec., and then placed it face down. He then indicated, by pointing to an animal on the large board, which animal was to be found on the trial. On each trial, the subject was asked to find a different animal. The test stimulus appeared in a different position each time. The performance score on the central task was defined as the total number of animals correctly located on the first attempt on each of the eight trials.

After all eight trials on the central task were completed, the experimenter then said:

OK, here are those other pictures that were on the cards (the experimenter holds up the second board with all the household objects), and here's one of the animals (points to first page of the booklet). I want to see if you can remember which one of these objects was touching this animal.

Performance on the incidental task was defined as the number of correct pairings of the eight. After both measures were completed, the central task was discussed, and he was asked if he had used any special way of remembering where each animal had been. These responses were noted.

Results

The mean central and incidental scores for verbal and nonverbal condition at each grade level are shown in Figure 1 ($n = 16$ per cell). Analyses of variance show that there were no significant differences between the means for the two conditions and no significant age trends.

A two-way analysis for repeated measures on one factor (central versus incidental measures) with no label and label conditions combined showed a significant interaction between condition and age ($F = 4.04$, $df = 2/93$, $p < .025$), which replicates earlier findings (Maccoby & Hagen, 1965; Hagen, 1967). Further analyses showed that this interaction was significant for the label condition ($F = 3.77$, $df = 2/45$, $p < .025$), but not for the no label condition ($F = 1.34$, $df = 2/45$).

The serial-position data for the central STM task were analyzed. For each subject a primacy score, defined as the sum of his scores on the first two positions, and a recency score, the sum for the last two positions, were computed. For all three grades, the no label condition resulted in a greater primacy score, and for the two younger grades the label group had a greater recency score. The interaction between primacy and recency scores was significant ($F = 4.03$, $df = 2/93$, $p < .025$). Grade levels were then combined. The primacy score for the no label group was greater than for the label group ($t = 3.52$, $df = 94$, $p < .001$). The recency score,

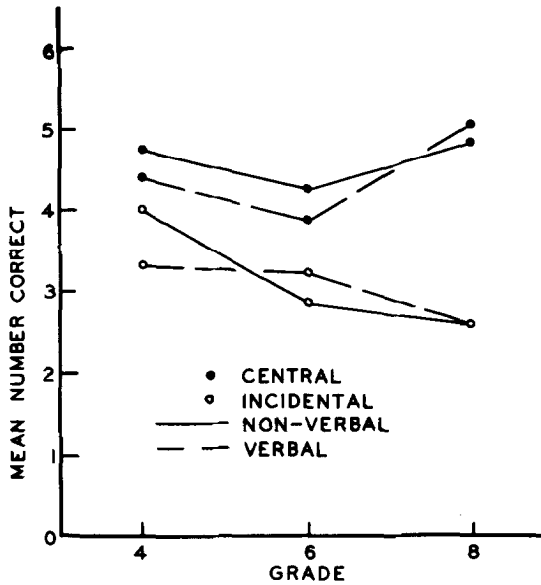


FIG. 1. Mean number correct of central and incidental scores in label and no label conditions as a function of grade level.

on the other hand, was higher for the label group ($t = 4.16, df = 94, p < .001$).

Table 1 presents the correlations between central and incidental scores with intelligence scores as measured by the CTMM. Since no differences between the no label and the label conditions were found, these have been combined. Correlations with central scores at grades 6 and 8 were positive and significant, but those with incidental scores were small and not significant.

The decrement due to labeling at the primacy portion of the serial-

TABLE 1
Correlations Between IQ and Scores, Label and No Label Groups
Combined ($n = 32$ in each Group)

	Grade		
	4	6	8
Central	-.22	.48**	.31*
Incidental	.22	.00	.05

** $p < .01$.

* $p < .10$.

position data was found as well as the facilitation due to labeling at the recency portion. However, for these children the two effects cancelled and no overall decrement in STM was found. The predicted interaction in central versus incidental recall performance was also found, but was significant only for the verbal condition. Intelligence measures correlated as expected.

STUDY II

This study differed from the first in several respects. First, only college-age subjects were used. It was expected that the predicted overall labeling deficit would be found with this population, since they should be highly skilled in the use of rehearsal strategies. Second, an additional condition was employed in which only the central picture was presented to determine the effects of the incidental stimuli on the central STM task. Third, the Scholastic Aptitude Test was used as the intelligence measure.

Method

Subjects. The subjects were 40 females from introductory psychology courses at The University of Michigan. They were assigned at random into either the label or no label condition. Each subject was tested in two conditions: one stimulus item per card and two stimulus items per card. Scores on the Scholastic Aptitude Test (SAT) were available for 35 of the subjects.

Stimulus materials. The stimulus materials were the same as in Study I. In addition, a set of cards was used which pictured only the animals.

Procedure. The procedure was the same in Study I except that the wording was changed to be appropriate for the age level of subjects in this study. In addition, all subjects were tested on a series of eight trials with the one picture cards. Half the subjects in each condition received the one-picture-per-card trials first and half received the two-picture-per-card trials first.

Results

Table 2 presents the mean scores for each condition. A two-way analysis of variance with repeated measures on one factor (one stimulus versus two stimuli) was performed. There was a significant decremental effect of labeling ($F = 5.16$, $df = 1/38$, $p < .05$); but there was no significant effect of one versus two stimuli ($F < 1$) and no significant interaction ($F < 1$).

A t test was performed to compare the difference due to one versus two stimuli in the label condition to the differences in the no label condition. At each serial position within each group a difference score was obtained

TABLE 2
Mean Task Score for Each Condition

	1 Stimulus	2 Stimuli	Total	Incidental
Label	5.05	4.95	5.00	2.35
No label	5.90	5.65	5.78	2.30

which represented the absolute value of the difference of one versus two stimuli. These difference scores were summed across position in each condition and compared. The no label condition was found to have greater difference scores than the label condition ($t = 2.27$, $df = 38$, $p < .05$).

The incidental recall scores were examined (Table 2). There was no significant difference in incidental recall for label versus no label conditions ($t < 1$). The incidental scores were therefore combined across groups. Incidental recall was higher than would be expected by chance ($t = 4.65$, $df = 39$, $p < .01$).

Correlations were computed between central recall scores and SAT scores. SAT scores were available for 17 of the 20 subjects in the no label group and for 18 of the 20 subjects in the label group. The central recall scores for the correlation were a combined total of subjects' scores on the one stimulus and two stimuli conditions. The correlation between central scores and SAT verbal scores was $-.02$ for the label condition. The correlation between central scores and SAT verbal scores for the no label condition was $.38$ ($t = 1.58$, $df = 15$, $p < .05$). The difference in the verbal SAT central task correlation for the label versus the no label condition was not significant ($z = 1.13$). The correlation between central scores and SAT mathematical scores for the label condition was $-.06$. For the no label group the correlation was $.51$ ($t = 2.27$, $df = 15$, $p < .05$). The difference between these was significant ($z = 1.66$, $p < .05$, 1 tail). Thus the predicted positive correlation was found in the no label condition only.

The correlation between two pictures per card central and incidental performance scores was $-.34$ ($t = 2.33$, $df = 38$, $p < .05$). Correlations between central and incidental performance for each condition were as follows: Label, $-.22$ (n.s.); no label group, $-.46$ ($t = 2.19$, $df = 18$, $p < .05$). Thus the relation was greater for the no label condition.

RESULTS OF STUDIES I AND II

The serial-position data for the two studies are compared in Figure 2. Since no differences were found as a function of grade level in Study I, the serial-position data were combined and labeled Children. Position 1

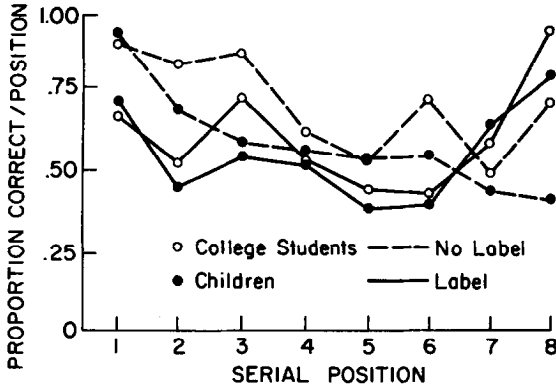


FIG. 2. Proportion of correct responses at each serial position in label and no label conditions for children and for college students.

represents the first card the subject saw on his left (primacy), and position 8 represents the most recently presented card (recency). The college students performed better than the children at 12 of the 16 positions for the two conditions. It is also clear that the decrement found for labeling appeared in all of the first six positions, for both groups of subjects, but only for the college group was it large enough to result in a significant overall decrement.

Figure 3 presents primacy and recency scores at each grade level.

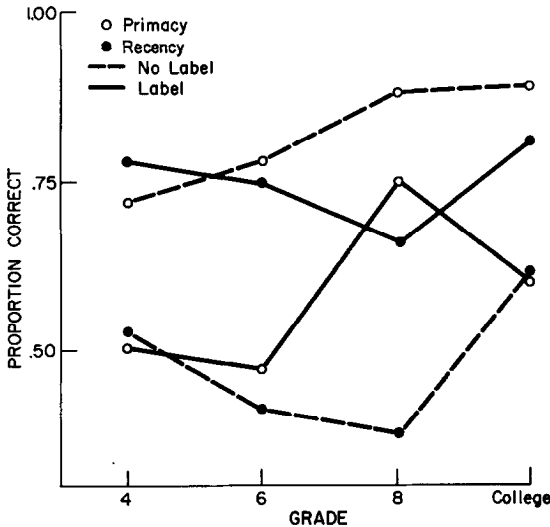


FIG. 3. Proportion of correct responses at primacy and at recency positions as a function of grade level.

Again, the first two positions were utilized to compute primacy scores and the last two, 7 and 8, to compute recency scores. As expected, primacy performance in the no label condition showed regular improvement as a function of CA. Recency performance showed no such trend, but did improve for college subjects in both conditions. It should also be noted that the two highest curves are: primacy, no label; and recency, label.

DISCUSSION

The effects of verbal labeling in a serial-position STM task have now been studied across a wide developmental spectrum. From an earlier study (Hagen & Kingsley, 1968) it was established that labeling did not affect performance of nursery school four- and five-year-old children, facilitated performance of six- and eight-year-old children, and did not affect performance of ten-year-olds. The absence of an effect for older children was replicated and found to apply through age 14 in Study I. A detrimental effect of labeling was found for college-age adults in Study II. Verbal labeling of the stimuli resulted in a decrement for the primacy portion of the serial-position curve. However, labeling facilitated recall of the most recently presented items across the whole age span studied.

The authors hypothesized that strategies which involve covert rehearsal in memory tasks develop with increasing CA. Induced labeling interferes with these more efficient strategies and thus becomes increasingly detrimental. The two studies presented here further support this hypothesis. In Study I no overall labeling effect was found, but the serial-position analyses were consistent. The serial-position effects of this study (Figure 2) were similar to those of Hagen and Kingsley (1968) for the label condition. In the no label condition, however, the primacy portion of the curve was higher in this study, as expected since the children were older. The lowest performance was obtained at the two most recently presented positions, which was not predicted. No explanation can be offered for this result at present.

Observations of the experimenter lend further support to the rehearsal hypothesis. In the no label condition, children were observed to whisper the animal names softly. Some children looked at and pointed to the cards that were already face down while waiting for another one to be presented. During the testing, some children counted up to the correct card. After testing, they were asked if they had any special way of remembering where the animals were in the row. Many replied that they had said the names repeatedly to themselves as they were presented. Some described strategies, such as recalling only first letters of the names. In a recent study, Kingsley and Hagen (1969) have demonstrated that

nursery-school children can be induced to use a rehearsal strategy and when they do so, their primacy performance is facilitated.

The serial-position performance of the verbal labeling condition is revealing. For the first six positions it was similar in shape to the no label condition although performance was lower. At positions seven and eight labeling facilitated recall. Since this effect occurred only at the two most recently presented positions, the increased saliency of the labeled stimulus is apparently easily interfered with by new information, or quickly fades with time.

The serial-position data for adults in Study II (Figure 2) were very similar to those of the younger subjects. Labeling did have an overall detrimental effect, however, due to the large difference between conditions at the primacy positions. Three types of evidence indicate that labeling inhibited the use of individual strategies. First, in the no label conditions for one versus two stimuli there was more variability than in the corresponding label conditions. Performance thus seems to be more under experimenter control and less under the subjects' control in the label condition.

Second, the correlations between central and incidental performance further support the interference hypothesis. Hagen (1967) found increasingly negative correlations between central and incidental recall as a function of CA. It was argued that as the child becomes older, he is a more efficient information processor and thereby is more able to ignore incidental or irrelevant information. In the present study the correlation between central and incidental recall was significant ($-.34$). For the no label condition it was $-.46$, and for the label condition it was $-.22$. In the earlier study (Hagen, 1967) it was found that when a distractor task was added the correlations disappeared. The verbal labeling thus had an effect similar to that of an imposed distractor.

Third, in Study I the correlation between central performance and intelligence (CTMM) was positive and significant at the older grade levels. The intelligence measures (SAT) employed in Study II were also correlated with central task performance in the no label condition (verbal, $.38$; mathematical, $.51$). In the label condition the correlations were near zero (verbal, $-.02$; mathematical, $-.06$). High scorers on the SAT lose whatever advantage they have in performing this task when required to label overtly. These findings are not, of course, conclusive, but together suggest that labeling interferes with the use of rehearsal strategies which are used in the performance of certain STM tasks by older children and adults.

These results call for new interpretations of the results of Atkinson, Hansen, and Bernbach (1964) and Bernbach (1967). Atkinson *et al.*

found no primacy effect in serial-position recall for four- and five-year-olds in a paradigm very similar to that used in this study. Bernbach attempted to show that requiring four- and five-year-olds to label overtly increased their primacy recall scores. He found that induced labeling in young children resulted in serial-position curves which looked more like the serial-position curves of older children and of adults which show primacy effects, although he did not actually test for significance. Bernbach concluded that labeling makes rehearsal possible and rehearsal increases primacy recall. Because adults label automatically, they show high primacy performance. While labeling does facilitate short-term memory in children at certain ages, the present studies should alert investigators that the processes are more complex than Bernbach suggests. Adults are doing more than merely saying names; they are actively repeating them in cumulative rehearsal fashion. Bernbach's conclusion that the difference "between the short-term memory performance of children and adults that was observed by Atkinson *et al.* may be traced to the effects of *labeling* [italics ours] on short-term memory, rather than to any inherent difference in the memory processes of children and adults" does not seem warranted (p. 150).

The finding that the presence of the incidental stimulus pictures did not affect central STM performance in Study II is not in agreement with an earlier study (Hagen, 1967) in which it was found that performance was lower in the two picture per card condition in the CA range 6-13. However, the incidental scores of the adult subjects of the present study were significantly above chance; thus they did not completely ignore these stimuli. Since *labeling* did affect their central STM performance, it can be argued that there is an interaction between the effects of experimental constraints and the developmental level of the subjects. Irrelevant stimuli have a detrimental effect at younger CA levels but no effect at older; imposed labeling facilitates at younger but inhibits performance at older CA levels. It seems that at younger ages the child is more dependent upon immediate stimuli in his environment, and he has not developed strategies for coping with specific task demands. The older child becomes more dependent upon his own information processing strategies; he can ignore stimuli that are irrelevant unless they conflict with his strategies. Verbal labels which are imposed externally are irrelevant, and even distracting, for the individual who does not utilize them in task performance.

The effects of experimentally induced labeling on serial-position STM studied by Hagen and Kingsley (1968) have been verified and extended developmentally by these two studies. It has been more convincingly demonstrated that active rehearsal strategies play an increasingly important role in STM processes as development occurs. Recently there has

been much research devoted to the broader issues involved. A recent chapter by Flavell (in press) summarizes his research on mediated memory and presents a theoretical account for the development of the use of mnemonic mediators for storing and retrieving information. The research of Flavell and the findings presented here are consistent with this theoretical argument.

REFERENCES

- ATKINSON, R. C., HANSEN, D. N., & BERNBACH, H. A. Short-term memory with young children. *Psychonomic Science*, 1964, **1**, 255-256.
- BERNBACH, H. A. The effect of labels on short-term memory for colors with nursery school children. *Psychonomic Science*, 1967, **7**, 149-150.
- ELLIS, N. R., & HOPE, R. Memory processes and the serial position curve. *Journal of Experimental Psychology*, 1968, **77**, 613-619.
- FLAVELL, J. H. Developmental studies of mediated memory. In H. W. Reese and L. P. Lipsitt (Eds.), *Advances in child development and behavior*. Vol. 5, New York: Academic Press, (in press).
- FLAVELL, J. H., BEACH, D. R., & CHINSKY, J. M. Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development*, 1966, **37**, 283-299.
- GLANZER, M., & CUNITZ, A. R. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 351-360.
- HAGEN, J. W. A developmental study of task-relevant and task-irrelevant information processing under distraction and non-distraction conditions. *Child Development*, 1967, **38**, 685-694.
- HAGEN, J. W., & KINGSLEY, P. R. Labeling effects in short-term memory. *Child Development*, 1968, **39**, 113-121.
- HAGEN, J. W., & SABO, R. A. A developmental study of selective attention. *Merrill-Palmer Quarterly*, 1967, **13**, 159-172.
- KAUSLER, D. H. Long-term retention. In D. H. Kausler (Ed.), *Readings in verbal learning*. New York: Wiley, 1966.
- KEENEY, T. J., CANNIZZO, S. R., & FLAVELL, J. H. Spontaneous and induced verbal rehearsal in a recall task. *Child Development*, 1967, **38**, 953-966.
- KENDLER, T. S. Development of mediating responses in children. In J. C. Wright and J. Kagan (Eds.), *Basic cognitive processes in children*. *Monographs of the Society for Research in Child Development*, 1963, Serial No. 86.
- KINGSLEY, P. R., & HAGEN, J. W. Induced versus spontaneous rehearsal in short-term memory in nursery school children. *Developmental Psychology*, 1969, **1**, 40-46.
- MACCOBY, E., & HAGEN, J. W. Effects of distraction upon central versus incidental recall: Developmental trends. *Journal of Experimental Child Psychology*, 1965, **2**, 280-289.
- POSTMAN, L. Short-term memory and incidental learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964.
- REESE, H. W. Verbal mediation as a function of age level. *Psychological Bulletin*, 1962, **59**, 502-509.

(Accepted June 10, 1969)