

Evaluative and Taxonomic Encoding in Children's Memory¹

ROBERT V. KAIL, JR.
University of Michigan

AND

JOHN T. SCHROLL
Ohio Wesleyan University

The experiments were conducted to investigate the development of evaluative and taxonomic encoding in children's memory. The task used was a modification of the Wickens short-term memory task in which subjects' recall of words is tested following a distraction task. The first experiment found that 11-year-old children, but not 8-year-old children, encoded words using the evaluative dimension of the semantic differential. In the second experiment, both 7- and 11-year-olds demonstrated the ability to encode words according to taxonomic categories. These findings were related to other recent work on the development of encoding in memory.

Recent research in memory has indicated that the developing child increasingly utilizes cognitive and verbal skills, such as rehearsal (Hagen & Kail, 1973) and the use of cues (Ritter, Kaprove, Fitch & Flavell, 1973) in the storage and retrieval of information. One aspect of memory that has received relatively little attention is the development of encoding processes. That is, what features of stimuli in general, and words, in particular, are important in their categorization in memory?

Wickens (1970) has studied memory encoding extensively with adult subjects using a paradigm that is a modification of the Brown-Peterson short-term memory task. In this paradigm a trial consists of presenta-

¹The authors are indebted to Harold Stevenson, John Hagen, Robert Jongeward, and Laura Levine for their critical readings of this manuscript and to Roy Wittlinger for his advice throughout this research. The cooperation of the staffs and pupils of St. Mary's School, Delaware, OH and Spencer Elementary School, Whitmore Lake, MI is greatly appreciated. Portions of these results were reported at the annual meeting of the American Psychological Association, Montreal, Quebec, Canada, August 1973. Requests for reprints should be addressed to Robert V. Kail, Jr., Department of Psychology, 529 Thompson St., University of Michigan, Ann Arbor, MI 48104.

tion of the to-be-remembered items, then a brief distraction task, followed by a test for recall. If the items are selected from a common category, performance declines rapidly over trials, indicating proactive interference (PI) caused by similar encoding in memory. If the items presented on a later trial shift to a different category, an increase in recall or "release from PI" effect is evident. The improvement in recall or "release effect" is due to the fact that words on the shift trial are encoded differently in memory, thus resulting in less interference (Wickens, 1970). Wickens (1970) suggests that the degree of recall improvement on a shift trial indicates the psychological importance of a category type for encoding in memory. In experiments using this paradigm Wickens has investigated the dimensions adults use for encoding and has found that various semantic dimensions, such as taxonomic class and values on the semantic differential, are of prime importance, while physical and syntactic characteristics, such as grammatical class, are of less importance.

Both Pender (1969) and Cermak, Sagotsky, and Moshier (1972) studied children's ability to use the evaluative dimensions of the semantic differential in memory encoding. Pender (1969) found that both second- and sixth-grade subjects demonstrated improved recall on the shift trial. Conversely, in the Cermak experiment only the sixth graders demonstrated both an accumulation and release of proactive interference, indicating the ability to encode along the evaluative dimensions employed in this study. Because the recall of the second- and fourth-grade subjects did not decline systematically over trials, indicating the build up of interference, interpretation of these data is problematic.

There is a problem with the stimuli used in the Cermak experiment that may have produced these inconsistent results. The positive and negative words used by Cermak *et al.* were listed by Osgood, Suci, and Tannenbaum (1957) as end points of scales which measure the evaluative dimensions of the semantic space. While such word pairs as *end points of scales* may measure the evaluative dimension of the semantic space, this does not mean that when presented *individually* they necessarily load heavily on the evaluative dimension themselves. For example, "fair" may tap the positive end of the evaluative dimension when presented with "unfair," but it clearly takes on an entirely different meaning when presented alone, e.g., a carnival. Pender (1969) selected stimuli from scales compiled by Heise (1965). Heise (1965) presented 1000 common words individually and had subjects rate them on all three major dimensions of the semantic space. Nine of the ten positive words used as stimuli in the Cermak study were measured by Heise (1965). The mean standard score ratings of these words on the three dimensions were evaluative,

0.89; activity, 0.05; potency, -0.81 . Thus the positive words used in the Cermak study confounded moderately high values on the evaluative dimension with moderately low scores on the potency dimension. Consequently, it is difficult to make any clear statements about the development of evaluative encoding ability from this study.

Experiment I further investigated the development of evaluative encoding in children. Children at two age levels, 8 and 11 years, were tested on the Wickens task, with shifts occurring between words valued positively and negatively on the evaluative dimension. Words were selected from the scales compiled by Heise (1965). It was assumed that both age groups would show an improvement in recall on the shift trial, replicating Pender's findings, because the stimulus confounding that occurred in the Cermak study had been eliminated.

EXPERIMENT I

Method

Subjects. Subjects were 50 second and third graders, and 50 fifth and sixth graders from Spencer Elementary School, Whitmore Lake, MI. All subjects were white and came from lower-middle- to middle-class homes. Twenty-five children at each age level were arbitrarily assigned to each of two experimental groups: younger shift (median CA = 8 years, 1 month); younger nonshift (8 years, 9 months); older shift (11 years, 7 months); older nonshift (11 years, 4 months). There was approximately an equal number of boys and girls in each group.

Materials. The words selected for this study from the scales compiled by Heise (1965) are presented in Table 1. Words were chosen to maximize evaluative values and minimize scores on the other scales; in general, the words chosen had a standard score of one or more on the evaluative dimension, and less than one on the other two dimensions. In addition, attempts were made to minimize reading difficulty for the younger children, as well as acoustic and associative similarity. The word dyads were printed in 1-in. block letters on 4×6 -in. white index cards. For the distractor task, 2×3 -in. colored patches were mounted on 4×6 -in. index cards. Presentation rate was paced with a continuously running stopwatch.

Procedure. The subject sat across a small table from the experimenter. The subject first saw a ready signal (a card with an asterisk) for 2 sec, followed by a word dyad for 2 sec. The subject read the words aloud as the card was shown, and was prompted if he failed to do so. Then the subject named the colored patches as they were shown to him. Colors

TABLE 1
 RATINGS OF WORD DYADS USED IN EXPT I ON EVALUATIVE, ACTIVITY, AND
 POTENCY DIMENSIONS OF THE SEMANTIC SPACE (FROM HEISE, 1965)^a

	Evaluative	Activity	Potency
<u>Positive words</u>			
fresh	1.60	-.38	-.68
save	1.11	.87	-.19

true	1.23	.41	-.35
home ^b	1.80	1.29	-1.81

good ^b	1.03	-.27	-.51
live ^b	1.22	.83	.14

open	1.15	-.27	.25
nice	1.76	-1.28	-1.42

church ^b	2.21	-.77	-.66
milk	1.38	-.96	.15
	$\bar{X} =$ 1.45	-.05	-.51
	$\bar{X}^b =$ 39.24	45.34	40.14
<u>Negative words</u>			
bad ^b	-3.35	-.79	.50
hate ^b	-3.11	.11	-.61

kill	-3.29	.98	-.27
cut ^b	-1.88	.95	-.67

fall	-2.17	-.04	-.65
worry	-2.20	.25	-1.64

burn	-2.47	.74	-.52
no ^b	-2.09	.27	.37

fear ^b	-2.25	-.05	-.99
cold ^b	-2.41	.05	1.39
	$\bar{X} =$ -2.52	.25	-.31
	$\bar{X}^b =$ 67.37	49.22	47.06

^a Dashed lines separate the word dyads used in this experiment.

^b Words included in the DiVesta-Walls (1970) monograph and their respective means.

were presented at the rate of one patch per second for 15 sec. Following the color-naming task, the subject had 10 sec in which to recall the two words presented on that trial.

The subject was first given a practice trial. In order to minimize interference with the word dyads, the stimuli were two digits. Five test trials followed the practice trial, with a 6-sec interval between trials. Color cards were randomly reordered during this period. For the subjects in the shift groups, the word class was shifted on the fifth trial. Approximately half of the subjects in the shift groups were tested with positive words on the first four trials, then shifted to negative words on the fifth trial; the remaining subjects shifted from negative to positive words. For the nonshift groups, approximately half of the subjects were tested with positive word dyads exclusively, while the remainder were tested with the negative word dyads. Word dyads were counterbalanced so that each dyad appeared equally often in each trial position.

All subjects were tested individually in a room within the school by a white male experimenter. Each experimental session lasted approximately 10 min.

Results

Figure 1 depicts the mean percentage of correct responses as a function of trials for the two age groups and shift and nonshift groups. Each subject scored one point for each word correctly recalled; order of recall was not included in the scoring. A preliminary analysis conducted on the data from trials 1 to 4 only indicated that there were differences due to trials ($F(3,264) = 17.57, p < .001$). There were no differences in recall of positive and negative words, nor did this variable interact with the trial factor. However, there was an Age \times Word-type interaction ($F(1,88) = 4.66, p < .05$). Simple effects tests revealed that the younger children recalled significantly more positive than negative words ($F(1,88) = 4.84, p < .05$), while the older subjects did not differ in their recall of the two types of words ($F(1,88) = 1.86$). There were no other significant differences.

A three-way analysis of variance for the Age, Condition (shift vs nonshift), and Trials variables on data from all five trials yielded only a significant trial effect ($F(4,384) = 13.00, p < .001$). The main effect for age approached significance ($F(1,96) = 3.82, p < .10$). There were no other significant main effects or significant interactions.

Comparisons between the means for shift and nonshift groups on trial 5 indicated no differences for the 8-year-olds ($t < 1$), while the shift group recalled more than the nonshift group at the older age level ($t(48) = 1.63, p < .1$). Furthermore, t tests measuring the difference

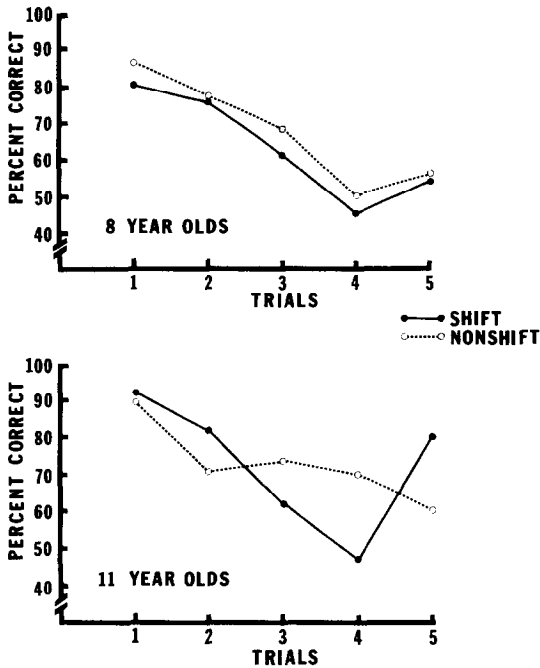


FIG. 1. Percent correct for shift and nonshift subjects at both levels in Expt I.

between performance on trials 4 and 5 yielded a significant improvement for the shift group ($t(24) = 2.47, p < .05$), and no change for the non-shift group ($t < 1$). Thus only the recall of the older children improved on the shift trial.

To test if the release effect for the older children was symmetric, the recall of the two subgroups in the shift condition was examined separately. On the fourth trial, both the group shifting from the positive to negative words and the group shifting from negative to positive words recalled 45% of the items correctly. On the shift trial, the recall of the positive-negative group improved to 75% and the negative-positive improved to 85%. In both groups a strong release effect obtained; thus the effect appears to be symmetrical with regard to the direction of the shift.

Discussion

The finding that younger children recalled more positive than negative words was unexpected but is consistent with previous research demonstrating such a "polyanna effect." Rhine (1965) found that adults learn positive cases more readily than negative ones in a concept-learning task. Klatzky, Clark, and Macken (1973) demonstrated that 4- and 5-year-

olds learn the positive instance of an adjective pair more rapidly than the negative case. Why this effect did not appear with the older children is unclear.

More important, these data were contrary to the predicted results. The general findings of Cermak *et al.* were replicated. There were no overall age differences in recall, yet there were developmental differences in the ability to use the evaluative dimension as an encoding category in memory. A possible explanation lies in an additional critical difference between the Pender (1969) study that found both older and younger children encoding along evaluative dimensions, and the Cermak and present studies. In the Pender experiment the stimuli were presented auditorily, while visual presentation of stimuli was used in the Cermak and present studies.

EXPERIMENT II

In research with adults, Wickens (1970) found that taxonomic features were the most widely used encoding categories. This finding has been replicated with first graders (Libby & Kroes, 1971), and with second and third graders in two unpublished doctoral dissertations (Pender, 1969; Wagner, 1970). However the procedures in each experiment involved auditory presentation of stimuli. Because of the ambiguous findings obtained on evaluative encoding in children, in Expt II visual presentation of stimuli was used to investigate the development of taxonomic encoding in 7- and 11-year-old children.

Method

Subjects. Subjects were 34 second and 34 fifth graders from St. Mary's School in Delaware, OH. Most subjects were white and came from lower-middle- to middle-class homes. Seventeen children at each age, with approximately the same number of boys and girls, were assigned to each of the two groups: younger shift (median CA = 7 years, 5 months); younger nonshift (7 years, 6 months); older shift (10 years, 6 months); older nonshift (10 years, 7 months).

Materials. The words used in this experiment were selected from *My Self Help Dictionary*, commonly used by first and second graders. Eight words were chosen from each of two categories, body parts and four-legged animals, with efforts to minimize acoustic similarity and associations between word pairs. Words were printed in 1-in. block letters on 5 × 7-in. white index cards. For the distractor task, colored patches were centered on 5 × 7-in. index cards. All time intervals were measured with a metronome.

Procedure. With a few modifications, the procedure from Expt I was

also used in this study. There were four rather than five trials. Thus the fourth trial was the shift trial in this experiment. Shifts were made in both directions, with approximately half the children shifting from animal words to body part words, the remainder from body parts to animals. Subjects in the nonshift groups were tested with either all animal or body part dyads.

Results

The mean percentage of words recalled as a function of trials for both age groups and shift conditions is presented in Fig. 2. Possible scores ranged from 0 to 2; order of recall was not included in the scoring. A preliminary analysis was computed on data from trials 1-3 only. Recall declined over trials ($F(2,128) = 11.49, p < .001$), and the older children's recall was higher than the younger children's ($F(1,64) = 4.82, p < .05$). Furthermore, there was a significant difference between the shift and nonshift groups ($F(1,64) = 10.03, p < .01$). This difference is inexplicable as all necessary counterbalancing procedures were performed. Because the shift and nonshift groups were treated identically on the preshift trials, this difference apparently reflects sampling error.

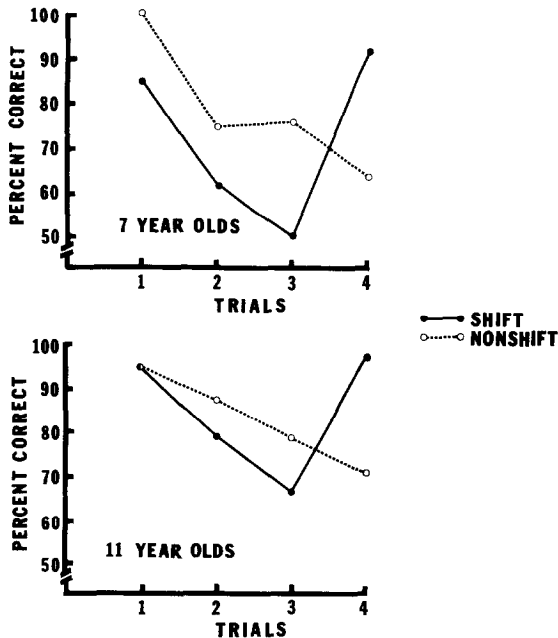


FIG. 2. Percent correct for shift and nonshift groups at both age levels in Expt II.

There was no difference in recall of the two word types, nor did this variable interact with any other variable.

A three-way analysis of variance was performed on the data from trials 1-4. The older children recalled significantly more words than did the younger children ($F(1,64) = 6.73, p < .05$). As predicted, performance declined over trials ($F(3,192) = 8.75, p < .01$), and there was a significant Shift Condition \times Trials interaction ($F(3,192) = 10.19, p < .01$). It is evident in Fig. 1 that the recall of the shift groups at both age levels improved on the shift trial, while the nonshift groups' performance continued to decline. Comparisons of recall on the fourth trial indicated a significant difference between the shift and nonshift groups at both age levels (younger: $t(32) = 3.69, p < .01$; older: $t(32) = 2.92, p < .01$), thus indicating taxonomic encoding by both younger and older children. There were no other significant differences.

Discussion

The general findings of Cermak *et al.* (1972) were replicated in Expt I. There were no overall age differences in recall, but only the older children encoded within the evaluative dimensions. One might argue that these data are artifactual in that the scales presented in Heise's monograph were obtained from adults. DiVesta and Walls (1970) presented children's ratings of words along the semantic differential. According to these norms also, the stimuli used in Expt I tap the evaluative dimension with little confounding of the other dimensions (see Table 1). Thus it cannot be argued that the experiment only investigated how closely children's evaluative encoding approximates the evaluative encoding of adult subjects.

Apparently this ability to use evaluative classes in memory develops sometime after the ability to place words correctly in evaluative categories. DiVesta (1966) has demonstrated that 7-year-olds can accurately classify words into positive and negative categories along the evaluative dimension, yet only the older and not the younger children utilized these categories to encode words. It is not clear from these data why only the older children can use these dimensions. They may have a richer semantic network in which to encode items, or their improvement in recall on the shift trial may reflect more sophisticated retrieval strategies.

Smothergill (1973) commented that the Wickens task may not be sensitive to the encoding categories used by children, and thus it would be of dubious utility for the study of memory development. It is clear from Expt II that the task is indeed capable of tapping encoding processes of children as young as 7 years. At each age level, a comparison of shift vs nonshift subjects did identify an encoding dimension—taxonomic

categories—used by the younger as well as the older children. However, it is certainly *not* possible to use the degree of release from PI as an index of the importance of a category in encoding (as has been done with adults) across different age levels, because various task dimensions, such as the difficulty of the distraction task for the subject, also vary with age.

The results of both experiments, plus the previous findings using the Wickens task to study encoding in children's memory, suggest the possibility of an Age \times Presentation Modality \times Encoding Dimension interaction. That is, it appears that, as the child develops, he encodes a stimulus into an ever increasing number of categories. At an early age the child may encode along a particular category in a single modality and only later use both the auditory and visual modalities. Specifically, it is suggested that the 7-year-old child can encode using taxonomic categories with both auditory and visual presentation, but can utilize only the auditory modality for evaluative encoding. On the other hand, the 11-year-old child is equally adept at using both modalities to encode items within evaluative dimensions.

That such differences in presentation modality could influence recall on the shift trial is well documented. Recent experiments (Hopkins, Edwards & Cook, 1973; Hopkins, Edwards & Gavelek, 1972) have demonstrated that auditory and visual presentation of stimuli yield different forms of storage in memory.

Also, several experiments (e.g., Hawkins, 1897; Murray & Roberts, 1968; Siegel & Allik, 1973) have shown that different presentation modalities lead to differential recall by children.

Why presentation modality should interact with encoding dimension at the younger age level is a question posed but not answered by these data. One possibility is that auditory presentation of stimuli may produce a release effect because the stimuli are in some way encoded more completely or consistently and thus are more sensitive to a change in the encoding of subsequent stimuli. Because stimuli are encoded more completely, there should be a greater amount of interference in the auditory than visual presentation mode. A comparison of Pender's data (1969, Fig. 2) and the results of Expt I provides some support for this explanation. In both experiments recall is approximately 87% correct on trial 1. However, in the Pender experiment, recall declines to 39% on the second trial, but only decreases to 76% in Expt I. Thus there is some evidence of greater interference with auditory presentation, indicating that processing of stimulus items in the auditory modality is somehow more complete or consistent.

There are two suggested ways in which auditory presentation could result in this more advanced level of processing of stimuli. Murray and

Roberts (1968) found that recall of auditorily presented stimuli was superior to recall of stimuli presented visually in 7-year-olds. In that study, auditory presentation of stimuli apparently facilitated rehearsal, a strategy that 7-year-olds do not ordinarily use with any degree of proficiency (Hagen & Kail, 1973). Thus auditory presentation may produce a release effect because the stimuli are encoded more completely due to enhanced rehearsal by the younger child.

Alternatively, a further difference between the procedures used for auditory and visual presentation may be critical. In most experiments that have used visual presentation, the stimuli for each trial have been presented simultaneously on the same card or slide, while auditory presentation is always a sequential process. The sequential presentation of stimuli, regardless of modality, may result in increased interference because the items on a given trial interfere with each other more than stimuli presented together simultaneously as a unit. Because the sequential presentation sequence generates more interference, it may be more responsive to a different encoding of incoming stimuli. Clearly, future research should focus on the relationships between encoding modality, encoding dimensions, and stimulus presentation format, in order to better understand encoding processes in children and adults.

REFERENCES

- CERMAK, L. S., SAGOTSKY, G., & MOSHIER, C. Development of the ability to encode within evaluative dimensions. *Journal of Experimental Child Psychology*, 1972, **13**, 210-219.
- DIVESTA, F. J. A developmental study of the semantic structures of children. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 249-259.
- DIVESTA, F. J., & WALLS, R. T. Factor analysis of the semantic attributes of 487 words and some relationships to the conceptual behavior of fifth grade children. *Journal of Educational Psychology Monographs*, 1970, **61**, 1-40.
- HAGEN, J. W., & KAIL, R. V., Jr. Facilitation and distraction in short-term memory. *Child Development*, 1973, **44**, 831-836.
- HAWKINS, C. J. Experiments on memory types. *Psychological Review*, 1897, **4**, 289-294.
- HEISE, D. R. Semantic differential profiles for 1000 most frequent English words. *Psychological Monographs*, 1965, **79**, (8, whole No. 601).
- HOPKINS, R. L., EDWARDS, R. E., & COOK, C. L. Presentation modality and proactive interference in short-term memory. *Journal of Experimental Psychology*, 1973, **98**, 362-367.
- HOPKINS, R. L., EDWARDS, R. E., & GAVELEK, J. R. Presentation modality as an encoding variable in short-term memory. *Journal of Experimental Psychology*, 1971, **90**, 319-325.
- KLATZKY, R. L., CLARK, E. V., & MACKEN, M. Asymmetries in the acquisition of polar adjectives: Linguistic or conceptual? *Journal of Experimental Child Psychology*, 1973, **16**, 32-46.

- LIBBY, W. L., JR., & KROES, W. H. Conceptual encoding and concept recall-recovery in children. *Child Development*, 1971, **42**, 2089-2093.
- MURRAY, D. J., & ROBERTS, B. Visual and auditory presentation, presentation rate, and short-term memory in children. *British Journal of Psychology*, 1968, **59**, 119-125.
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana, IL: Univ. Illinois Press, 1957.
- PENDER, N. J. A developmental study of conceptual, semantic differential and acoustical dimensions as encoding categories in short-term memory. Unpublished Ph.D. dissertation. Northwestern University, 1969.
- RHINE, R. J. Preference for a positive evaluative response in concept learning. *Journal of Experimental Psychology*, 1965, **70**, 632-635.
- RITTER, K., KAPROVE, B. H., FITCH, J. P., & FLAVELL, J. H. The development of retrieval strategies in children. *Cognitive Psychology*, 1973, **5**, 310-321.
- SIEGEL, A. W., & ALLIK, J. P. A developmental study of visual and auditory short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 409-418.
- SMOTHERGILL, D. W. Does the Wickens STM technique tap encoding in young children? *Journal of Experimental Child Psychology*, 1973, **15**, 360-361.
- WAGNER, J. F. A developmental study of categorical organization in short-term memory. Unpublished Ph.D. dissertation. University of Connecticut, 1970.
- WICKENS, D. D. Encoding categories of words: An empirical approach to meaning. *Psychological Review*, 1970, **77**, 1-15.