This study was designed to investigate the functional similarity of the mental processes children use to learn verbal tasks and pictorial tasks. Children in grades 3 and 6 (n=144) and in grade 9 (n=112) were given four short paired-associate tasks entitled Pictures, Concrete Words, Abstract Words, and Japanese Characters. The tasks consisted of six stimulus-response pairs presented over two trials. Planned comparisons in mean level of performance involved the first three tasks. Performance on Pictures was found to be superior to Concrete Words, and Concrete Words was superior to Abstract Words, with the former effect reaching significance for grades 3 and 9 and the latter for grades 3 and 6. An analysis of all four tasks found that correlation between Pictures and Concrete Words increased across grade levels to a greater degree than correlation between any other pair of tasks. This last result paralleled data from an auxiliary experiment and suggests a developmental increase in children's use of verbal processes along with imagery to learn pictorial materials. (Author/NH)
INTERRELATIONS IN CHILDREN'S LEARNING OF VERBAL AND PICTORIAL PAIRED ASSOCIATES

Gordon A. Hale

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
January 1971
INTERRELATIONS IN CHILDREN'S LEARNING OF VERBAL AND PICTORIAL PAIRED ASSOCIATES

Gordon A. Hale
Educational Testing Service

Several recent studies have examined children's use of imagery in learning tasks by varying the degree to which the stimuli are capable of evoking mental images (see Paivio, 1969; Rohwer, 1970). This variation has been effected in two major ways, both generally in connection with a paired-associate learning format. In one series of investigations, the learning of problems employing concrete nouns has been compared with performance on tasks using pictorial representations of those nouns. Another set of studies has contrasted the learning of abstract and concrete verbal materials—that is, words which differ in the extent to which they can be represented pictorially. In general, this research has shown that children learn associations more easily when pictures rather than words are used (e.g., Rohwer, 1968; Rohwer, Lynch, Levin & Suzuki, 1967) and, in the case of verbal materials, when the words are concrete rather than abstract (e.g., Paivio & Yuille, 1966).

Although it is generally agreed that visual imagery plays at least some role in producing these effects, the nature of this role remains largely undetermined on the basis of the simple intertask comparisons which have been performed. The fact that pictures are easier to learn than words, for example, can be interpreted in either of two ways. It is possible that pictorial and verbal associations are learned by essentially different means, perhaps through imagery on the one hand and verbal memory on the other, and that the former
method is more effective. Alternatively, these two types of material may be learned in an essentially similar manner—e.g., through verbal association—but pictorial representation serves as an additional aid to learning. A major problem of interpretation, then, is to determine the degree of similarity in the mental processes employed to learn these various verbal and pictorial tasks. One means of addressing this issue is to examine correlations in performance on problems such as these. A basic assumption of this approach, as has been elaborated elsewhere (Stevenson, Hale, Klein & Miller, 1968), is that high correlations observed across different tasks indicate a similarity in the processes used to learn them, while lower correlations suggest the operation of relatively dissimilar processes.

The current study was designed to provide evidence bearing on the functional similarity for children of various verbal and pictorial tasks, through assessment of the degree of correlation in performance on these problems, along with intertask comparisons in mean level of performance. Children at three grade levels, third, sixth, and ninth, were given a series of short paired-associate tasks, which differed in the types of materials employed. Problems involving pictures, concrete words, and abstract words were included, and the major intertask comparisons of the study were contrasts between successive pairs of these problems, as in the research cited above. To provide additional information for an analysis of intercorrelations among tasks, a problem using Japanese characters was also included as a second measure of pictorial memory. With a sample thus consisting of more than a single measure from each of the categories of interest
(e.g., verbal, pictorial), a correlational analysis could provide information bearing on the influence of such categories in determining functional similarities among tasks. Specifically, this influence was expected to be manifested in relatively high correlations (a) between the two verbal measures, (b) between the two pictorial measures, and (c) between the two tasks involving representation of the same objects in two modalities (Pictures and Concrete Words). Differences in the patterns of correlations across age levels were also examined, to determine the extent to which the processes employed in learning these tasks undergo developmental changes.

Method

Subjects

The subjects were members of eight classrooms at each of grades 3 and 6 and four classrooms at grade 9 in public schools in a middle-class area. After subjects were randomly excluded to yield an equal representation from each classroom in a grade, the final sample consisted of 144 subjects at each of grades 3 and 6, and 112 subjects at grade 9 (mean CAs = 8.7, 11.7, and 14.6 years, respectively).

Materials

Stimulus and response elements. Four short paired-associate tasks were constructed, each of which consisted of six stimulus-response pairs presented over two trials. The stimulus elements in each case were nonsense syllables having a high association value for adults (83% - 86%) according to the Archer (1960) norms, with a different but equivalent set of stimulus elements being used in each
of the four tasks. Response elements were selected from four classes of material, which varied along the critical dimensions of the study. In the task entitled Pictures the response elements were line drawings of the following objects: bird, chair, flower, girl, hand, and sheep. In the task entitled Concrete Words, the responses were the words represented in the Pictures task, while in the problem entitled Abstract Words, the responses were words that were less amenable to pictorial representation—dark, health, hungry, joy, trouble, and wish. The words for both tasks were selected from those listed in Thorndike and Lorge (1944) which occur 100 or more times per million, to ensure that the two sets of words would be of approximately equal familiarity. In the problem entitled Japanese Characters, the response elements were simple Japanese characters of roughly equal complexity, each containing five strokes in a different configuration.

Booklets. Response booklets were used to record the subjects' answers. Each booklet contained nine pages, one for an initial practice test and the remainder devoted to two test trials for each of the four tasks. On each page of the booklet, six stimulus elements appeared in a column along the left side, with six response elements in a row to the right of each stimulus. The orders of stimulus and response elements differed across successive pages, and these orders bore no correspondence to the temporal order in which the pairs were presented on the study trials.

Procedure

The tasks were administered in a single group-testing session in the subjects' classrooms. A study trial-test trial procedure
was used, in which the materials to be learned were projected from
slides onto a screen, after which the subjects responded in booklets.
On each study trial, the stimulus-response pairs were projected one
at a time for a period of six seconds each, with one second between
pairs. Following each study trial, the subjects turned to a given
page in the booklet for the test trial and drew a circle around the
picture (or the word) which they remembered to be correct for each
"made up word." Instructions concerning the nature of the task and the
use of the task booklet were given in connection with a practice test
involving six pairs of nonsense syllables, and care was taken to
ensure that the children at all age levels understood the procedur.
The presence of two monitors served to minimize the incidence of
cheating; following an initial admonition to the classroom, a monitor
noted any cheating that persisted, and booklets for these subjects
were later discarded.

Experimental Design

The two trials for any given task were presented in succession,
and the order in which the four tasks were presented was counter-
balanced across classrooms. Four different orders were constructed,
such that each task appeared once in each of the four temporal positions,
and each problem was followed once by every other problem. In addition,
two forms of each task were constructed with identical stimulus and
response elements but different stimulus-response pairings. Each of the
eight resulting Form x Order combinations was assigned to one of the
eight classes at each of grades 3 and 6. The four classes at grade 9
were assigned the four orders, but only one form of each task. As
the major analyses were to be based on a combination of all classrooms with each grade, subjects were randomly discarded to yield an equal number from each classroom in a grade and thus an equal representation of all orders in which the problems appeared.

Results and Discussion

Table 1 presents the mean number of items correct for each task and grade level, and these data are depicted graphically in Figure 1.

Insert Table 1 and Figure 1 about here

Differences in performance on these problems were examined by means of simple contrasts rather than an overall analysis, as two comparisons were of particular theoretical interest. These were (a) Pictures versus Concrete Words and (b) Concrete Words versus Abstract Words. The Japanese Characters task was designed primarily for inclusion in the correlational analysis, although mean performance data have also been tabulated for this problem to give a perspective on the difficulty of this task relative to the others.

Intertask Comparisons

Pictures versus Concrete Words. The results for this comparison are in agreement with evidence from earlier research in that pictorial presentation of the materials generally appears to have been more effective than verbal presentation. The data also suggest rather
complex developmental changes in this effect, as indicated by $t$ tests (for correlated means) on the difference between tasks at each grade level. The superiority of pictures over words was significant at grade 3 ($t = 2.82$, $df = 143$, $p < .01$), was only slight at grade 6 ($t = .42$, $df = 143$, $p > .50$), and was significant again at grade 9 ($t = 2.73$, $df = 111$, $p < .01$). An additional analysis was performed to determine the significance of these apparent developmental changes.

For each subject, a difference score was derived by subtracting performance on Concrete Words from performance on Pictures, and changes in this score across successive grade levels were examined. Of the two comparisons involved, only the drop in difference score from grades 3 to 6 was found to be significant ($t = 1.97$, $df = 286$, $p < .05$).

The last result is consistent with that obtained in the study by Rohwer, et al. (1967), in which it was found that pairs of pictures were more easily learned than pairs of printed words, and this superiority tended to decrease from grade 3 to grade 6. Prior to grade 3, on the other hand, Rohwer (1968) has found an age-related increase in the superiority of pictures over words presented auditorially. While several differences exist between the present study and this last study, including the mode in which the verbal stimuli were presented (i.e., visual vs. auditory) and the method of assessing learning (recognition vs. recall), it might be hypothesized that the primary basis for the difference in results is the age of the subjects involved. If this is true, a complex developmental trend is suggested, with the superiority of pictorial over verbal presentation of materials increasing from preschool age to third grade, remaining relatively constant from
third to ninth grade, but possibly decreasing for a period between these ages. It would be of value to assess developmental changes in the pictorial-verbal distinction across this entire age range using identical conditions throughout, such as auditory presentation of the verbal materials, to provide a direct test of this interpretation.

It should be mentioned parenthetically that the superiority of pictorial over verbal presentation of materials has been found in some research to occur primarily when this variation is effected on the stimulus side of pairs rather than the response side (e.g., Dilley & Paivio, 1968). However, the present data indicate that the superiority of pictorial presentation may also be observed on the response side under appropriate conditions. Although these conditions are not readily identifiable, an important factor may be the means by which retention is assessed. The study by Dilley and Paivio, for example, used a recall method, while the present study employed a recognition method, in which the response elements were visible during the test of retention as well as during learning. Thus, the response elements were presented here under conditions more similar to those in which stimulus elements are typically presented, and this similarity could account for the present obtainment of effects, via response variation, that have usually been produced by stimulus variation.

Concrete Words versus Abstract Words. The level of performance on Concrete Words was found to be significantly higher than that on Abstract Words, as expected, for grades 3 and 6 (t = 2.97 and 2.84, respectively, df = 143, p < .01), but not for grade 9 (t = .99, df = 111, p > .20). The score representing the difference between these
two tasks did not change significantly across grade levels; nevertheless, the apparent reduction in superiority of Concrete Words to nonsignificance at grade 9 presents a developmental picture that is difficult to explain. Some related studies have found that adults also learn concrete words more easily than abstract words (Kusyszyn & Paivio, 1966; Paivio, 1963), hence the present data cannot be interpreted to indicate that this effect disappears by adolescence. Yet the present comparison is apparently not subject to a ceiling effect, as the means for both verbal tasks at grade 9 were considerably lower than that for the Pictures task.

Perhaps the data are viewed most appropriately in terms of relative performance on all three of the tasks involved in the above comparisons. It is apparent in Figure 1 that, at each grade level, these tasks form a hierarchy from Pictures to Concrete Words to Abstract Words. While there is some fluctuation in the degree of difference between tasks, the most striking aspect of the data is the consistency with which the tasks rank order in this manner. Pending replication to clarify the apparent interactions, then, the present evidence might generally be interpreted as indicating a uniformity, from middle childhood to adolescence, in the relative difficulty of these three types of material.

Correlational Analysis

In order to discount the possibility that differences in the intertask correlations to be reported are attributable to differences in reliabilities of the present tasks, some attention was given to the issue of reliability assessment. While it is difficult, if not impossible, to obtain accurate estimates of reliability for problems that involve
learning (Melton, 1936), a rough indicator of internal consistency for the present tasks is the correlation between trial 1 and trial 2 performance. These correlations were expected to be only of moderate magnitude but should permit the detection of gross differences among the tasks in internal consistency. The intertrial correlations for Pictures, Concrete Words, Abstract Words, and Japanese Characters, respectively, were: for grade 3, .44, .55, .46, .44; grade 6, .58, .61, .55, .52; and grade 9, .45, .64, .56, .61. No consistent ordering of tasks is apparent, except that the correlation for Concrete Words was slightly higher than the others at each grade. The correlations increased somewhat across grade levels, which may indicate an age-related increase in internal consistency of these measures. However, these differences may also reflect the possibility that the older children, being closer to asymptotic performance in the two trials of each problem, are more likely to use similar mental processes on both trials.

Table 2 presents the intercorrelations among tasks. Two general features of these correlations will be noted initially. First, as would be expected, all of the correlations were significant ($p < .001$), reflecting the operation of general abilities in performance on all of these tasks. Second, the correlations appeared to increase across grade levels. This result may be due to factors similar to those affecting the correlation between trials of a single task—for example, the fact that the older children were closer to an asymptotic level of
performance. The assumption in this case is that higher intertask correlations are obtained in later stages than in early stages of learning, an assumption which is supported by the following analysis.

Two additional sets of intertask correlations were computed, one based on the first trial of each task and the other based on the second trial. A comparison of these two sets indicated the former correlations to be generally lower than the latter; the medians across all pairs of tasks for these two sets, respectively, were: for grade 3, .25 and .36; for grade 6, .33 and .49; and for grade 9, .43 and .41.

Most important for the purposes of this study are the patterns of correlations observed—i.e., differences between pairs of tasks in the magnitudes of the correlations obtained. In this regard, it had been expected that the data would reflect the dimensions along which the materials had been designed to differ. That is, it was anticipated that the correlations (a) between the two verbal tasks, (b) between the two pictorial tasks, and (c) between the tasks involving representation of concrete objects in two modalities (Pictures and Concrete Words) would be higher than the correlations across these categories. Such did not prove to be the case, however, as differences among the correlations within each grade followed neither the expected pattern nor any pattern that was identifiable. Perhaps functional similarities and differences among the present tasks may be determined to a large extent by factors orthogonal to the dimensions on which these tasks were intended to vary.

Although examination of the patterns of correlations within grade levels suggested no consistencies, comparison of these patterns across grade levels yielded an interesting result. The correlation
between Pictures and Concrete Words was the second lowest of the six intertask correlations at grade 3 but was the highest in the matrix at both grades 6 and 9. Furthermore, the differences between grades 3 and 6 and between grades 3 and 9 in the correlation between these two tasks marked the only significant age differences in the magnitude of correlations obtained ($z = 2.21$ and $2.43$, respectively, $p < .05$, using standard significance levels). Additional evidence for a developmental increase in correlation between these two tasks was provided by an auxiliary experiment performed in connection with this study.

In this experiment, eight third- and eight fifth-grade classrooms were given a series of complete eight-trial learning tasks, each of which involved one of the four sets of material described above. The procedure was identical to that which has been described, except that the stimulus-response pairs were presented, via film, in the following manner: stimulus, three seconds; stimulus plus response, three seconds; intertrial interval, two seconds. Each classroom received three of the four tasks, with a one-week separation between tasks. Each of the four possible combinations of three tasks was assigned to two classrooms at each grade; the two classrooms given a particular combination differed in (a) the stimulus-response pairings used in each task and (b) the order in which the three tasks were presented. The total number correct on each task was tabulated for every subject and correlations between all possible pairs of tasks were calculated for each classroom, which ranged in size from 20 to 32 pupils. For each pair of tasks, a median correlation was then derived from the four classrooms in each
grade that had received that pair of tasks. These median correlations, in the order of pairs listed in Table 2, were: for grade 3, .29, .49, .49, .28, .41, .57; and for grade 5, .61, .48, .51, .37, .47, .47.

As will be noted, the median correlation between Pictures and Concrete Words increased 32 points from .29 at grade 3 to .61 at grade 5, while the change in median correlation between all other pairs of tasks was 10 points or less.

In general, then, the data appear to indicate a developmental increase in the relation between performance on a task involving printed words and performance on a problem employing the pictorial representations of those words. This result implies that relatively different processes may be used in performing these problems at the younger age levels but, as a child grows older, the mental processes that come into play in learning these two tasks become more similar. Perhaps with increasing age children begin to use verbal processes in learning both problems, so that the older child tends to employ a primarily verbal type of memory to recall information whether it is presented verbally or pictorially. Alternatively, the data may be interpreted in terms of a developmental increase in the use of visual imagery in recalling either type of information. A third possibility is that, as children mature, they tend increasingly to employ both imagery and verbal memory simultaneously in recalling information presented through either verbal or visual means.

In a synthesis that is related to the last of these interpretations, Rohwer (1970) states, "pictures evoke imagery at all age levels assessed, but the ability to profit from the stored images is contingent upon
the subject's ability to store an appropriate verbal representation of the object along with its image [p. 395]." Thus, Rohwer hypothesizes that a child becomes better able to transform pictorial into verbal information with increasing age and to store both types of information together in memory. This analysis was formulated to account for an observed increase prior to grade 3 in superiority of pictorial over auditory presentation of materials, but it is also consistent with the present data. The rise in relative correlation between the pictorial and verbal tasks observed in the present study may reflect a continued increase beyond grade 3 in children's tendency to add verbal processes to imaginal ones in learning pictorial materials. That some form of imagery maintains an important function in learning pictorial materials, however, is indicated by the continued superiority of performance on the pictorial task over the verbal task through grade 9.

Order Effects

An additional set of analyses was performed, involving the effects of order of task presentation, to aid in interpreting the results discussed above. First, examination of the overall effects of temporal order indicated that tasks presented later in the sequence were generally learned more readily than those presented earlier. This effect was significant for grades 6 and 9, as indicated by one-way analyses of variance with temporal position as a repeated-measures factor ($F(3, 429) = 4.25, p < .01$, and $F(3, 333) = 3.31, p < .05$, respectively). In a second analysis, subjects who received the Pictures task before Concrete Words were separated from those who received the tasks in the reverse order; for each of these two groups a score indicating the difference
between Pictures and Concrete Words was derived, and the inequality between groups in this difference score was assessed by means of a t test. Similar comparisons were performed to determine the effects of the order in which Concrete Words and Abstract Words were presented. The results of these analyses showed the difference between Pictures and Concrete Words to be relatively unaffected by the order in which these tasks were presented (p > .10 at all grade levels). However, the difference between Concrete Words and Abstract Words was significantly greater, at both grades 6 and 9, for subjects receiving the former rather than the latter task first (t = 2.45, df = 142, p < .05, and t = 2.48, df = 110, p < .05, respectively); a similar though nonsignificant difference was observed for grade 3. These results cannot be attributed to the general order effect described above, as they are in a direction opposite to that effect. Possibly a type of proactive interference was involved in the case of the two verbal tasks, causing relative performance on either of these tasks to be poorer when preceded by the other.

In a final analysis, correlations between Pictures and Concrete Words, and between Concrete and Abstract Words, were computed separately for each order of presentation of the tasks in these respective pairs. The results indicated little difference between orders in correlations for the two verbal measures but a significant difference for Pictures and Concrete Words at grades 3 and 9. At both of these grade levels, the intertask correlation was higher for subjects receiving Concrete Words before Pictures than for subjects receiving the reverse order (.21 vs. .57, z = 2.55, p < .05; and .48 vs. .72, z = 1.98, p < .05,
respectively); the results for grade 6, although nonsignificant, were in a similar direction (.54 vs. .62). Apparently, presentation of Concrete Words prior to Pictures increased the functional similarity of these tasks; that is, this order appears to have induced the employment of similar mental processes in learning these two types of material. It is reasonable to hypothesize further that the difference in functional similarity in this case related to employment of verbal processes. Assuming that presentation of the names of objects before their pictures fosters the use of these names when the pictures are later encountered, presentation of Concrete Words prior to Pictures in the present case likely increased the probability that the children would use verbal processes to store information presented pictorially. Thus, changes in the functional similarity of the verbal and pictorial tasks with variation in the order of presentation may be viewed as somewhat analogous to the changes occurring across grade levels; both effects appear to involve the addition of verbal processes to imaginal ones in learning materials presented pictorially.

The preceding interpretations are still highly speculative at this point, and many questions regarding children's use of imagery in these tasks remain to be answered. At the least, it is hoped that the present research has contributed suggestions for further work on this topic. First, the study has pointed out the value of correlational analysis as an aid to interpreting effects observed via intertask comparisons. Second, the data that have been obtained highlight the necessity for studying developmental trends across a wide range of ages in order to understand fluctuations in results which may occur
within a limited age range. Third, the need to examine performance on all of these verbal and pictorial tasks in a single research effort is indicated, as developmental changes in abilities related to imagery may be reflected simultaneously in the pictorial-verbal and concrete-abstract comparisons.
References


Footnote

This investigation was supported in part by the National Institute of Child Health and Human Development, under Research Grant No. 1 FO1 HD01762. The author wishes to express his gratitude to the Edison Township and Bridgewater-Raritan school systems of New Jersey for their generous cooperation in this research.
Table 1
Mean Number of Correct Responses and Standard Deviation
(in Parentheses) for Each Task at Each Grade Level

<table>
<thead>
<tr>
<th>Task</th>
<th>N</th>
<th>Pictures</th>
<th>Concrete Words</th>
<th>Abstract Words</th>
<th>Japanese Characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 3</td>
<td>144</td>
<td>6.20 (3.22)</td>
<td>5.37 (3.13)</td>
<td>4.60 (2.82)</td>
<td>4.66 (2.73)</td>
</tr>
<tr>
<td>Grade 6</td>
<td>144</td>
<td>9.02 (3.04)</td>
<td>8.92 (3.20)</td>
<td>8.12 (3.18)</td>
<td>7.08 (3.23)</td>
</tr>
<tr>
<td>Grade 9</td>
<td>112</td>
<td>10.16 (2.66)</td>
<td>9.49 (3.12)</td>
<td>9.23 (2.77)</td>
<td>8.45 (3.32)</td>
</tr>
</tbody>
</table>
Table 2
Correlations in Performance on All Pairs of Tasks, with Each Task Abbreviated by Its Initial Letter

<table>
<thead>
<tr>
<th>Pairs of Tasks</th>
<th>N</th>
<th>P-C</th>
<th>P-A</th>
<th>P-J</th>
<th>C-A</th>
<th>C-J</th>
<th>A-J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 3</td>
<td>144</td>
<td>.38</td>
<td>.51</td>
<td>.40</td>
<td>.46</td>
<td>.40</td>
<td>.31</td>
</tr>
<tr>
<td>Grade 6</td>
<td>144</td>
<td>.58</td>
<td>.50</td>
<td>.47</td>
<td>.44</td>
<td>.48</td>
<td>.43</td>
</tr>
<tr>
<td>Grade 9</td>
<td>112</td>
<td>.61</td>
<td>.53</td>
<td>.56</td>
<td>.56</td>
<td>.53</td>
<td>.50</td>
</tr>
</tbody>
</table>
Figure Caption

Fig. 1. Mean number of correct responses for each task at each grade level.
FIGURE 1

Mean Correct Responses

GRADE 3
GRADE 6
GRADE 9

Pictures
Concrete Words
Abstract Words
Japanese Characters