The role of meaning in the learning of verbally presented paired associates (PAs) is described by means of an experiment with a replication. A printed PA list containing six different types of items was presented to Sixth Graders. Four of the item types consisted of sentence-embedded PAs on the study trials, followed by test trial cues varying in terms of their similarity in meaning to the study materials. The two remaining item types served as baseline measures of performance, i.e., controls. It was found that: (1) the greater the semantic similarity between the study and test trial contexts, the greater was the recall of PA items, while at the same time; (2) control item types presented in conjunctive phrases were better recalled than those presented in sentences. Implications of these findings are discussed in terms of ongoing research into semantic and imaginal components of children's learning. (Author)
Working Paper No. 50
Verbal and Visual Processes in Children's Learning: I. Meaning In Paired Associate Learning

Report From the Project on Variables And Processes in Cognitive Learning
WORKING PAPER NO. 50

VERBAL AND VISUAL PROCESSES IN CHILDREN'S LEARNING:
I. MEANING IN PAIRED-ASSOCIATE LEARNING

By Joel R. Levin and James M. Morvitz

Report from the Project on Variables and Processes in Cognitive Learning


Wisconsin Research and Development Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

October 1970

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE-5-10-154
STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Working Paper is from the Project on Variables and Processes in Cognitive Learning in Program 1. General objectives of the Program are to generate new knowledge about concept learning and cognitive skills, to synthesize existing knowledge, and to develop educational materials suggested by the prior activities. Contributing to these Program objectives, the Verbal and Visual Components of Children's Learning Project has the following three general objectives: to develop a test battery which will reliably assess the effectiveness of different types of presented materials in children's verbal learning; to evaluate such materials as a function of particular stimulus, procedural, task, and instructional variables; and to identify individual capabilities, as related to presentation mode, to examine systematically the performance of various learner types, and to diagnose individual learning profiles, culminating in tailor-made instructional sequences.
TABLE OF CONTENTS

Abstract ............................................. vii
I. Introduction ....................................... 1
II. Method ............................................. 3
   Design and Materials ............................... 3
   Subjects ......................................... 6
   Procedure ...................................... 6
III. Results .......................................... 8
    Replication Study ............................... 11
IV. Discussion ....................................... 13
    References ..................................... 18

LIST OF TABLES

Table                                      Page

1 Examples of Study- and Test-Trial Materials . . . . . . 4
2 Mean Number of Items Recalled on Each Test Trial
   (out of 2) as a Function of Item Type ........................ 9

LIST OF FIGURES

Figure                                      Page

1 Comparison of the Initial and Replication Studies:
   Mean Number of Correct Responses, by Item Type,
   Over Two Test Trials ................................ 12
In recent years the paired-associate (PA) paradigm has provided a framework for investigating the role that semantic and syntactic factors play in learning. For example, a great deal of empirical evidence suggests that the learning of a PA list involves more than mere passive participation on the learner’s part. Rather, subject reports obtained during and after learning, albeit introspective and less than perfectly reliable, indicate that proficient learners restructure or elaborate upon the to-be-associated items (by the addition of verbal or imaginal elements) in order to render the pairs more memorable (Martin, 1967; Paivio, 1969; Rohwer, 1967; Runquist and Farley, 1964).

Furthermore, it has been reported that experimenter-provided sentences or phrases, serving to relate PA items to one another, facilitate the task (Davidson, 1964; Reese, 1965; Rohwer, 1966). That the kind of auxiliary verbal material supplied is important has also been confirmed. Rohwer (1966) found syntactic structure to be an effective facilitator of PA learning only when the provided elaboration consisted of meaningful English words. Thus, sentences like “The running COW chases the bouncing BALL” were facilitative, while Lewis Carroll-like pseudosentences (“The ludding COW drases the spraking BALL”) were not, relative
to non-elaborated pairs ("COW--BALL").

In a subsequent experiment, Rohwer and Levin (1968) discovered that the properties: \{meaningful, English, words\} were necessary but not sufficient effectors of improved PA performance. In that study, the way in which stimulus and response members were related to one another was crucial. When the relationship was semantically appropriate (e.g., "CATS jump CATES") facilitation was produced; when semantically anomalous (e.g., "CATS jump SONGS"), it was not.

The present study follows from these and other recent experiments (e.g. Davidson and Dollinger, 1969; Davidson, Schwenn and Adams, 1970; Ehri and Rohwer, 1969) in which semantic characteristics of verbs have been manipulated in order to examine the relationship between sentence meaning and PA recall.

---

1 This is interesting in light of Epstein's (1961) finding that non-meaningful verbal material (e.g., "maff vlem ooth glox nerf") is more easily recalled when syntactic markers are inserted in generally agreed-upon places (e.g., "The maff vlems oothly the glox nerfs"). Perhaps the distinction lies in the interpretability of the original learning materials: when they are essentially meaningless (cf. Epstein, 1961) the addition of verbal elements is facilitative; when they are basically meaningful (cf. Rohwer, 1966) the addition of verbal elements is beneficial only if they too are meaningful.
II

METHOD

Design and Materials

The learning materials, which were typed onto a memory drum tape, consisted of twelve pairs of familiar nouns embedded in five-word sentences or conjunctive phrases. The twelve items were randomly distributed throughout the list, appearing as one of six types, with two items apiece representing each type. The six item types were defined by the form in which they were presented to Ss on study and test trials. Four of these were included to test hypotheses concerned with semantic factors in PA learning, while the remaining two served as controls.

Examples of the four experimental and two control item types may be found in Table 1. Beginning with E₁ and continuing through E₄, one will perceive a decrease in the semantic similarity between the sentences studied and the frames supplied as test cues. In E₁, for example, the verb initially studied is reinstated as a cue for recall on the test trial, while in E₂, a synonymous verb (or one which implies similar meaning) is provided. The extent to which the test verb evokes denotative and connotative responses different from those of the study might be manifested by performance differences on item types E₁ and E₂.

However, this difference should be smaller in magnitude than that anticipated when E₁ is compared with item types E₃ and E₄, where more
Table 1. Examples of Study- and Test-Trial Materials

<table>
<thead>
<tr>
<th>Experimental Item Types:</th>
<th>Study Trial</th>
<th>Test Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_1$</td>
<td>The GIRL grabs the BALL</td>
<td>The GIRL grabs the ___</td>
</tr>
<tr>
<td>$E_2$</td>
<td>The GIRL takes the BALL</td>
<td>The GIRL grabs the ___</td>
</tr>
<tr>
<td>$E_3$</td>
<td>The GIRL throws the BALL</td>
<td>The GIRL grabs the ___</td>
</tr>
<tr>
<td>$E_4$</td>
<td>The GIRL attends the BALL</td>
<td>The GIRL grabs the ___</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control Item Types:</th>
<th>Study Trial</th>
<th>Test Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>The GIRL grabs the BALL</td>
<td>The GIRL or the ___</td>
</tr>
<tr>
<td>$C_2$</td>
<td>The GIRL or the BALL</td>
<td>The GIRL or the ___</td>
</tr>
</tbody>
</table>
than orthographical changes in the test-trial verb are involved. In E3, the test verb signals an action or activity on the part of the stimulus noun different from what was communicated on the study trial. In E4, on the other hand, the original meaning of the response noun has been completely altered by the verb switch: in the example in Table 1, what may have been encoded as "girl present at gala event" is now retrievable only from the cues "girl reaching or grabbing for something" (in this case, a small spherical object).

If semantic factors are involved in the learning of such PA items, a decrement in performance would be predicted as a function of semantic dissimilarity (between study and test contexts). That is to say, in terms of recall the four item types would be rank-ordered as follows: $E_1 > E_2 > E_3 > E_4$, where at one extreme the study- and test-trial contexts are identical ($E_1$) while at the other they convey dissimilar, or even competing, sentence meaning ($E_4$). The characteristics of the predicted function are intentionally not hypothesized but will be empirically evaluated in the present study.

The two item-type controls, examples of which are also presented in Table 1, were incorporated to serve a two-fold purpose: (a) to provide sentence- and nonsentence-baselines against which to assess the contribution to learning of the hypothesized semantic factor; and (b) to yield an estimate of the previously discussed sentence facilitation effect for the present task, by comparing $C_1$ (sentence control) with $C_2$ (nonsentence control).

Each S received one of six different combinations ("arrangements") of item types and item content. This was counterbalanced across Ss in
order that a partial test for possible item-content/type differences could be conducted.

Subjects

Sixty Sixth-Grade students from an elementary school serving a semi-rural community in the Midwest participated in the experiment. In addition, five other students were pilot-tested in order that E might adapt to the experimental procedure. The names of all participants were randomly selected from the roster of Sixth Graders in the school.

Procedure

All Ss were individually tested. Upon entering the experimental room, S was seated in front of a Lafayette memory drum. The instructions were then read by E, followed by the presentation of three sample items on the memory drum, one at a time.

The sample items were given to familiarize S with the to-be-performed task, as well as to impress upon him the fact that, for some of the items, study and test contexts would not be the same. This was done by means of two examples in which the connective was changed (verb to conjunction, and verb to preposition) from study to test trial, and one in which it was not (conjunction to conjunction). The order in which the sample items were presented was randomly predetermined for each S.

2 The authors wish to acknowledge the cooperation of the staff and students of the Central Grade School in Stoughton, Wisconsin.
Upon ascertaining that $S$ understood the nature of the task, $E$ presented the twelve study items to $S$ one at a time, at a four-second rate, while reading them aloud concurrently. Following the last item there was an eight-second intertrial interval, during which $E$ readied $S$ for the test trial by reminding him of the task. The test frames then appeared at a four-second rate while read aloud by $E$, $S$'s job being to supply (vocally) the missing response nouns. After the last test item was exposed, $S$ was given another eight-second "reminder" interval, followed by another study trial-test trial cycle. On each of the two study and two test trials, the twelve PA items appeared in a different (random) serial order, in order to remove position cues as a possible source of variance.
III
RESULTS

Learning was measured in terms of the number of correct responses given on each of the two test trials. A one-way multivariate analysis of variance with repeated measurements was performed, in which the between-subjects variation included the particular arrangement of item types and content provided for S, while twelve within-subject measures were produced by the combination of the six item types and two test trials.

The twelve dependent measures were initially transformed into eleven new variables in order that hypotheses regarding item-type differences, trials differences, and their interaction might be tested. In addition, the sum of the original twelve dependent measures provided a test for between-group (arrangements) differences in overall performance (Morrison, 1967).

All hypotheses in the present study were tested with the probability of a Type I error set equal to .05. The mean number of items correctly recalled on each trial, summed across the arrangements factor (which proved statistically nonsignificant), is presented for each item type in Table 2.

The multivariate test for item-type differences detected a statistically significant effect ($F = 21.10$ with 5 and 50 d/f). The "sum"
Table 2. Mean Number of Items Recalled on Each Test Trial (out of 2) as a Function of Item Type

<table>
<thead>
<tr>
<th>Experimental Item Types:</th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Sum</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E₁</td>
<td>1.15</td>
<td>1.75</td>
<td>2.90</td>
<td>0.60</td>
</tr>
<tr>
<td>E₂</td>
<td>0.93</td>
<td>1.52</td>
<td>2.45</td>
<td>0.59</td>
</tr>
<tr>
<td>E₃</td>
<td>0.70</td>
<td>1.30</td>
<td>2.00</td>
<td>0.60</td>
</tr>
<tr>
<td>E₄</td>
<td>0.43</td>
<td>1.13</td>
<td>1.56</td>
<td>0.70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control Item Types:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C₁</td>
<td>0.33</td>
<td>0.97</td>
<td>1.30</td>
<td>0.64</td>
</tr>
<tr>
<td>C₂</td>
<td>0.70</td>
<td>1.40</td>
<td>2.10</td>
<td>0.70</td>
</tr>
</tbody>
</table>

| Trial Totals:           | 4.24    | 8.07    |     |            |
Column in Table 2 reveals that the experimental item types conformed to the predicted pattern; that is, with decreasing semantic similarity between study and test contexts (going from $E_1$ to $E_4$) there is a corresponding decrement in response noun recall. Appropriate post hoc comparisons among the four experimental item-type means (Morrison, 1967) indicated that although differences between adjacent pairs were not statistically significant, $E_1$ was reliably different from $E_3$ and $E_4$, as was $E_2$ from $E_4$. More succinctly, as might be inferred from the item-type sums in Table 2 (and statistically corroborated by means of post hoc analysis), a negative linear trend accounts for virtually all of the variance in the difference among experimental item types.\(^3\)

With respect to the control item types, it was found through post hoc comparisons that: (a) each of the experimental item types except $E_4$ (where a complete meaning change occurred between study and test contexts) was statistically different from $C_1$, the sentence control, and in addition (b) $C_2$, the nonsentence control, was significantly higher than $C_1$.

As in most learning experiments of this type, the trials effect was substantial ($F = 327.98$ with 1 and 54 d.f.), with an average of 4.2 items out of twelve correctly recalled on Trial 1 and 8.1 on Trial 2. At the same time, the nonsignificant interaction of trials and item types

\(^3\)Although the nature of the item-types variable (i.e., undefined, and perhaps unknown, size of the interval between adjacent types) would not permit an a priori application of trend analysis (see, for example, Winer, 1962), the present "linear trend" conclusion is simply based on the significant set of (linear) coefficients which best describes the data in a post hoc sense.
(F < 1) indicates that despite marked improvement from Trial 1 to Trial 2, the relative ordering of item-type effectiveness remained quite constant. This is reflected in the "difference" column (Trial 2 minus Trial 1) in Table 2.

**Replication Study**

Additional data which were collected, in part to replicate the experiment just reported, are currently being analyzed as part of the second author's Master's thesis. The task and procedures were essentially the same as those already reported with minor modifications of the materials employed, and different random arrangements of item types and item content. As before, Sixth Graders constituted the target population.

A graphical comparison of the two studies may be found in Figure 1. Although the replication and initial samples are performing at different absolute levels, it may be seen that the two item-type profiles are essentially parallel.

The multivariate test for item-type differences was once again statistically significant, with post hoc comparisons suggesting that a set of negative linear trend coefficients best describes the nature of the effect among the four experimental item types. The difference between the two control item types, though in the same direction as previously, was not significant (according to multivariate post hoc procedures at the .05 significance level). 4

Finally, the F-ratio for trials was again found to be significant, while the trials by item types effect was not (F < 1).

---

4 This is more than likely attributable to the smaller sample size (N = 30 as compared with N = 60) of the replication experiment, since the size of the C₂ - C₁ difference is even larger than before.
Figure 1. Comparison of the Initial and Replication Studies: Mean Number of Correct Responses, By Item Type, Over Two Test Trials

<table>
<thead>
<tr>
<th>ITEM TYPES</th>
<th>Initial Study</th>
<th>Replication Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>E₁</td>
<td>Experimental</td>
<td>Control</td>
</tr>
<tr>
<td>E₂</td>
<td>Experimental</td>
<td>Control</td>
</tr>
<tr>
<td>E₃</td>
<td>Experimental</td>
<td>Control</td>
</tr>
<tr>
<td>E₄</td>
<td>Experimental</td>
<td>Control</td>
</tr>
<tr>
<td>C₁</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C₂</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Typical of most behavioral research findings, the results of the present study generate a host of questions still in need of resolution. A few of these are intimated in the paragraphs which follow.

First, the data obtained from the four experimental item types attest to the importance of semantic factors in learning. What is more, the significant linear trend in recall as a function of the semantic similarity between study and test contents emerged despite the fact that (i) there was an extremely restricted range of values for the number of correct items possible by item type (0-2 for each trial, 0-4 across trials); and (ii) two trials were given, between which Ss were able to restructure inappropriate sets or better attune themselves to the "switch game" being played. Yet, the item-type profile observed on Trial 2 was virtually identical to that of Trial 1, with Ss exhibiting a consistent improvement of between .6 and .7 items (out of 2.00) on every item type (cf. Table 2).

In the present experiments a "mixed" list was incorporated to test the semantic hypothesis. That is, a single list containing each type of item was provided for every S. In this way, one is able to assess the effect of the semantic manipulation within Ss, since each S serves
as his own control. Additional merits of the mixed list design have been suggested elsewhere (Levin, 1970). Whether the results described here would obtain in an independent groups experiment, where Ss are randomly assigned to groups containing different types of items and where group means are subsequently compared, needs to be investigated. However, there is reason to believe that an independent groups experiment of this kind would produce profiles similar to those reported here (Levin, in press; Rohwer, 1967; Rohwer and Frederiksen, 1969; Wilder, 1970).

Next, although the present finding that the nonsentence control item type (C2) tended to produce better recall than the sentence control (C1) is not consistent with the previous literature about the efficacy of experimenter-provided sentence elaboration (outlined earlier), it should be re-examined here for a few reasons. In the first place, the similarity of the contexts presented on study and test trials was greater for the nonentence (conjunction-conjunction) than for the sentence (verb-conjunction) item types. In most of the experiments where the sentence effect has been obtained, the materials studied were generally sentences, conjunctive phrases, or noun-noun pairs followed by a test-trial cue of the stimulus noun by itself. In the present experiment, the switch from a sentence to a conjunctive phrase frame may have interfered with sentence/response noun retrieval, producing an effect equivalent to that observed when the sentence meaning was changed between study and test contexts. The fact that item types E1 and C1 were not statistically different lends support to this notion.

Of greater concern is that some recent studies employing printed materials have reported not detecting the now-famous sentence facilitation...
effect (Davidson, Schwenn, and Adams, 1970; Levin, in press; Yuillie and Pritchard, 1969), which was substantiated (or worse, detected in the opposite direction) in the present study. Since the phenomenon has been found to be most pronounced when either pictorial, aural, or subject-initiated elaboration is employed (e.g., Bower, 1968; Davidson, 1964; Paivio, 1969; Reese, 1965; Rohwer, et al., 1967), it may be that provided printed sentence material does not facilitate, or may even interfere with, learning relative to nonsentence-provided material. Why this is so has been suggested in a recently completed study (Levin, Horvitz, and Kaplan, in press) in which the way of presenting verbal elaboration to learners, either in printed or in aural form, was factorially manipulated. In that experiment, it was found that elaboration effects are indeed more pronounced when the verbalization is heard but not seen.

It seems inappropriate to conclude without relating the data reported here to the exciting work of Allan Paivio and his associates, which has reawakened a dormant interest in the importance of imaginal activity during learning. [See Paivio, 1969, for a review of some of this research.] It is easy to demonstrate, for example, that by inducing Ss to employ visual imagery on a learning task by means of pre-learning instructions followed by examples or special training, tremendous gains in recall are produced (e.g., Bower, 1967; Jugelski, Kidd, & Segmen, 1965; Craig & Raser, 1969; Spiker, 1960; Taylor & Black,

5 Although inferences of this kind frequently lend themselves to empirical validation (through an examination of Ss' incorrect overt responses), in the present study a determination of the nature of subject errors was not enlightening, since nearly three-quarters of the errors made under each item type were of the "omission" variety.
Whether or not such procedures are conducive to implementation among young children is an empirical question, since studies employing imagery instructions have focused largely on college-aged Ss, with only a few dropping down as far as the Sixth Grade.

The possible differential ability to profit from visual imagery as a function of age provides an interesting testing ground for theories of cognitive development. In the Master's thesis research already described, preliminary analysis indicates that Sixth Grade Ss were able to benefit from imagery instructions while Third Graders were not.

In the context of the present study, by relating the effects of imagery instructions at different age levels to the item-type differences reported here, inferences regarding developmental differences in the spontaneous utilization of visual imagery might be drawn. That is, if subjects do in fact engage in imaginal activity when attempting to encode the study pairs, the elicitation of the same image by means of an identical or similar context on the test trial might be assumed to facilitate response-noun retrieval relative to test contexts which are dissimilar or conflicting. Of course, this is what was observed among the experimental item types and interpreted in terms of semantic similarity.

The study-to-test trial changes in meaning of the response nouns, as employed in the research reported here, is but one indirect method of investigating the role of semantic factors in learning, as is the provision of instructional sets in studying the use of visual imagery. Professor Davidson's research at the Wisconsin Research and Development
Center has incorporated traditional transfer designs, among others, to answer similar questions. Through techniques such as those suggested here, psychologists and educators have been able to "get more of a handle on" the nature of covert verbal and visual processes commanded by children when they learn.
REFERENCES


Levin, J. R. Note on mixed lists in the analysis of repeated measurements. Technical Memo from the Wisconsin Research and Development Center for Cognitive Learning, Madison, No. 5-70-7, 1970.

Levin, J. R. Factors related to the sentence facilitation of paired-associate learning: some characteristics of verb. Journal of Educational Psychology, in press.


