DOCUMENT RESUME

ED 214 126 CS 006 543

AUTHOR McGivern, Julie E.; Levin, Joel R.
TITLE The Keyword Method and Children's Vocabulary Learning: An Interaction with Vocabulary Knowledge.
PUB DATE Apr 82
NOTE 20p.; Paper presented at the Annual Meeting of the International Reading Association (27th, Chicago, IL, April 26-30, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Associative Learning; Grade 5; Intermediate Grades; Learning Theories; *Mnemonics; *Prior Learning; Reading Instruction; *Reading Research; Recall (Psychology); *Vocabulary Development; *Vocabulary Skills; Word Recognition
IDENTIFIERS *Keywords

ABSTRACT

A study explored a potential aptitude-by-treatment interaction associated with the keyword method of vocabulary acquisition. This method is a two-stage mnemonic process whereby an unfamiliar term is first transformed into a familiar concrete stimulus and then a thematic relationship is created between the transformed stimulus and the information associated with the original term. Subjects were 144 fifth grade students with either high or low levels of vocabulary knowledge who were randomly assigned either to one of three keyword instructional conditions that varied in the degree of structure provided or to a no-strategy control condition. All subjects were asked to learn 16 new vocabulary words. Results showed that all three variations of the keyword method facilitated students' vocabulary learning. However, aptitude-by-treatment interactions involving vocabulary knowledge materialized in such a way that the degree of keyword structure made far less difference for the high knowledge students than it did for the low knowledge subjects. In particular, when the students had to execute the dual components of the keyword method entirely on their own, low vocabulary knowledge students experienced considerably more difficulty than did the high knowledge students. (FL)

The Keyword Method and Children's Vocabulary Learning: An Interaction With Vocabulary Knowledge

Julie E. McGivern and Joel R. Levin

University of Wisconsin

Running Head: Keyword Method Interaction

The Keyword Method and Children's Vocabulary Learning: An Interaction with Vocabulary Knowledge

The purpose of the present study was to explore a potential aptitude-by-treatment interaction (ATI) associated with Atkinson and Raugh's (1975) "keyword" method. Proposed originally as a technique for improving college students' foreign vocabulary learning, the keyword method has recently been applied to children's acquisition of new vocabulary and other factual information in their own native language (Levin, 1981; Pressley, Levin, & Delaney, in press). The method itself is a two-stage mnemonic process, whereby: (1) the unfamiliar term is first transformed into a familiar concrete stimulus; and (2) a thematic relationship is then created between the transformed stimulus and the information associated with the original stimulus. Thus, for example, to remember that the English word dogbane means a type of plant, one could use the keyword dog and imagine a dog chewing a plant. Or, to remember that a dorado is a fish, one could picture a fish knocking on someone's door. It is not unusual for students learning via the keyword method to experience memory increments of 100% or more.

To date, the bulk of keyword research has focussed on several materials-and-procedural variations associated with the method. In contrast, little attention has been paid to the question of whether the keyword method is equally effective for students of differing cognitive abilities. A few studies have addressed this question by defining individual differences in terms of a group characteristic, such as age (e.g., Pressley & Levin, 1978). Moreover, a few other studies have examined

selected individual differences variables using secondary and college students (e.g., Delaney, 1978; Pressley, Levin, Nakamura, Hope, Bispo, & Toye, 1980; Mullis, Note 1). What was of interest here was whether the keyword method--or more specifically, different variations of the method--would work equally well for younger students (fifth graders) who differed widely in ability.

The ability of interest was general verbal ability. Because such a construct cannot be easily defined or measured, a vocabulary knowledge subtest of a standardized instrument served as a proxy. The major reason for selecting a verbal ability measure as the target individual differences dimension was that the first (stimulus-transformation) stage of the keyword method certainly appears to involve a component of verbal facility. In the present study, variations of the method, as well as the specific vocabulary items selected, exploited this presumed verbal facility either to a greater or lesser degree. It was anticipated that the more the task was fashioned to require verbal manipulations on the part of the child, the more highly related would performance be to the measured individual differences variable.

The type of ATI anticipated relative to a no-strategy control condition was not so obvious. On the one hand, students of higher ability might exhibit greater facilitation from an ability-demanding variation of the keyword method, in comparison to their lower ability peers. On the other hand, the higher ability students might be expected to perform at a high level when left to their own devices (i.e., in a control condition), which would serve to work in opposition to the potential

interaction just noted. Indeed, some would expect an interaction to materialize in which differences between higher and lower ability students were larger under control than under keyword instructions (see Levin, 1976).

Method

Subjects

The subjects were 144 fifth-grade students from three elementary schools in a midwestern community. Initially, 254 students were administered a pretest--the Level C, Form 1 vocabulary subtest of the Cognitive Abilities Test (Thorndike & Hagen, 1971)--to determine their level of vocabulary knowledge. This pretest included 25 vocabulary items. For each item, students were to choose the best definition from five possible choices.

Students' scores on the vocabulary-knowledge pretest ranged from 8 to 25 correct responses. The reliability of the pretest was determined for both the initial sample of 254 students and for the 144 students selected for participation in the study. Cronbach's alpha was .70 for the initial sample and .82 for the subsample. There were no significant sex differences in performance on the pretest among subjects participating in the study ($p > .20$). Subjects chosen for this study as "higher ability" subjects were students scoring 21 and above (75 students, from whom 72 subjects were randomly chosen). These subjects were randomly assigned in equal numbers ($n = 18$) to four experimental conditions, to be described shortly. "Lower ability" subjects were students scoring 16 and below on the vocabulary test (74 students, from whom 72 subjects were randomly chosen). These subjects were also randomly assigned in

equal numbers (n = 18) to the four experimental conditions. The resulting two ability groups represent approximately the top and bottom 28% of the original sample, and the mean pretest difference between the two groups (about 8 points) represents more than five within-group standard deviations.

Design and Materials

Each subject was asked to learn the definitions of 16 low-frequency English vocabulary words (e.g., penna, meaning feather; and pyrene, meaning stone). A list length of 16 items was chosen to avoid both ceiling and floor effects (see Pressley, 1977a). Each vocabulary word contained three or fewer syllables and a concrete one-syllable keyword. All keywords were located in the first syllable of the vocabulary item. Half of the vocabulary items contained keywords that could be extracted directly from the vocabulary word itself (e.g., pen for penna), and half contained keywords that could be extracted only after at least one orthographic transformation of the vocabulary word (e.g., pie for pyrene). The two types of items will be referred to hereafter as direct keyword and indirect keyword, respectively. This item-type manipulation was examined in relation to the ATIs of interest here. (Note, however, that performance on the two item types cannot be compared directly, in that the constituent items were neither randomly selected from some well-defined population nor matched on all other potentially relevant orthographic and semantic dimensions.)

Three keyword conditions and a control condition were included. The three keyword-condition variations were: fully structured, where subjects were taught the keyword method, were provided with keywords, and were shown a relational picture of

the keyword and the definition for each vocabulary item; semi-structured, where subjects were taught the keyword method, were provided with keywords, but were required to generate their own relational images for all vocabulary items; and unstructured, where subjects were taught the keyword method, but were required to generate both relational images and keywords for all vocabulary items. In the control condition, subjects were given motivating instructions to use their own learning strategies. The general ATI expectation was that as more self-generation was required on the part of keyword subjects, larger differences between higher and lower ability students would emerge.

The examples and vocabulary items learned were the same in all four conditions. The materials used for presenting the examples were essentially the same in the three keyword conditions. These included two 8.5" x 11" (21.6 x 27.9 cm) cards per item, one containing the sample vocabulary item, its keyword, and its definition, and the other containing a relational picture of the keyword and the definition. In the unstructured condition, for each example there was an additional card that contained only the vocabulary item and its definition.

Materials used during the study phase of the experiment in the fully structured condition included a relational picture of the keyword and definition for each of the 16 vocabulary items. On the bottom of each picture was printed the vocabulary item with its keyword and its definition. For the semi-structured condition, study materials included one card for each vocabulary item, on which was printed the vocabulary item, its keyword, and

its definition. In the unstructured condition, materials included one card per item, which contained only the vocabulary item and its definition. Materials for the control condition included one card for each of the three examples and each of the 16 study items. Each card contained the vocabulary item and its definition.

Procedure

The vocabulary subtest of the Cognitive Abilities Test was group-administered by one experimenter to all potential subjects one week before the start of the experiment, to minimize the effects of pretesting. Subjects were not aware that they were selected or grouped according to the results of the pretest. All keyword and control treatments were individually administered to each subject by three experimenters who were "blind" to the vocabulary-knowledge level of the students. The amount of instructional time varied slightly from subject to subject in all conditions, but generally lasted approximately three to four minutes.

In all keyword conditions, the procedure for explaining the keyword method and presenting the first example was constant. The subjects were shown a card containing a vocabulary word (tarn), a keyword (tar) and the definition (lake). The experimenter read each of these to the subject. During the explanation of the method, a relational picture of a lake and some tar was shown to the subject. Presentation of the second and third examples varied according to keyword condition. In the fully structured and semi-structured conditions, subjects were shown the vocabulary items (piggin and sopor), keywords (pig and soap), and definitions (bucket and sleep). Subjects in

the fully structured condition were shown a relational illustration for each example while the experimenter described the illustration. Subjects in the semi-structured condition were first asked to generate and describe their own relational image for each example, and then were shown the experimenter's illustration. Subjects in the unstructured keyword condition were shown only each vocabulary item and definition, and were asked to generate and describe both a possible keyword and a relational image for each. After doing this, subjects were shown (as an alternative possibility) the same keyword and illustration used in the two other keyword conditions. Subjects in all keyword conditions were asked after each example to explain how they were using the keyword method to remember the item definition. The experimenter then provided feedback to the subjects regarding their use of the method.

In the control condition, subjects were shown the three examples on cards that contained only the vocabulary items and definitions, both of which the experimenter read. With the presentation of the first example, subjects were given motivating instructions to try as hard as possible to remember the definitions and to use any special method that would help them remember the vocabulary items. After presentation of each of the examples, subjects were asked to recall the definition and to explain anything that helped them remember it. The experimenter provided feedback to help subjects evaluate their own learning strategies.

In the study phase of the experiment, all subjects were allowed to study each vocabulary item and definition for 15 seconds. The study cards, containing eight direct and eight

indirect keyword items, were shuffled prior to each use so that all items were presented in a randomly ordered mixed list to each subject. In all keyword conditions, subjects were reminded to try to use the keyword method to remember the definitions. Control subjects were reminded to use their own best method to remember the definitions. In the fully structured keyword condition, subjects were shown a relational picture for each vocabulary item. At the bottom of the picture, the vocabulary item, its keyword, and its definition were printed. Subjects were asked to read the item, keyword, and definition to themselves, and to study the picture. In the semi-structured keyword condition, the subjects were shown cards, each containing a vocabulary item, its keyword, and its definition. Students were asked to read these to themselves and make up a picture to help them remember the meaning. In both the unstructured keyword and control conditions, subjects were shown cards, each containing a vocabulary item and its definition. Unstructured keyword subjects were asked to read each item and its definition to themselves, make up a good keyword ("word clue"), and make up a picture to help them remember the definition. Control subjects were asked to read each item and definition to themselves and study it carefully. This phase of the experiment lasted approximately five minutes.

All subjects were given the same oral recall test. Subjects were shown each vocabulary item on a 5" x 8" (12.7 x 20.3 cm) card and were allowed a maximum of 15 seconds per item to respond orally with the item definition. The experimenter showed each item to each subject and recorded all responses. The 16 test items were presented in the same random order to all

subjects.

Results

A subject's response was scored as completely correct if it represented either the exact definition previously provided, a synonym, or the essential meaning of the vocabulary item (e.g., "coat" for cape). Responses were given half-point credit if they were partial definitions of the item, examples of the item, or broader or narrower essences of the definition. One experimenter scored a random sample of twenty recall tests twice to determine intra-rater reliability of the scoring procedure. No scoring errors or discrepancies were found. Two experimenters independently scored a random sample of twenty recall tests to determine the inter-rater reliability of the scoring procedure. Again, no scoring errors or discrepancies were discovered.

Analyses were conducted on both item-type sums (i.e., number correct on the eight direct keyword items plus number correct on the eight indirect keyword items) and item-type differences (i.e., number correct on the direct keyword items minus number correct on the indirect keyword items) to permit an assessment of effects across item types and in interaction with item types, respectively. All analyses followed Dunn's method of multiple comparisons (Kirk, 1968), with specific per-comparison and familywise Type I error probabilities selected on the basis of a priori statistical power considerations (Levin, 1975).

Insert Table 1 about here

The mean percentage of total items correct, by conditions and ability level is summarized in Table 1. Across ability levels, each keyword condition mean was significantly higher than the control mean, $t > 4.08$, $p < .001$, in all cases. Moreover, comparisons among the three keyword conditions revealed that the fully structured variation was statistically superior to the semi-structured variation, $t = 3.41$, $p < .001$. Neither the fully structured versus unstructured difference ($t = 1.39$) nor the semi-structured versus unstructured difference ($t = -2.02$) reached the needed Dunn critical value of 2.62.

In Table 1, evidence may be found to support the notion that different variations of the keyword method do interact with children's abilities. First, considering just the three keyword conditions, the difference between higher and lower ability students' mean performance decreases with increases in structure (mean differences of 40.6, 20.0, and 8.0 in the unstructured, semi-structured, and fully structured conditions, respectively). The interaction comparison involving the unstructured and fully structured conditions was statistically significant, $t = 3.63$, $p < .001$. Neither the unstructured versus semi-structured comparison ($t = 2.29$) nor the semi-structured versus fully structured comparison ($t = 1.33$) reached the needed Dunn critical value of 2.36. However, the ability difference in the unstructured keyword condition (40.6) was statistically greater than that in the control condition (18.9), $t = 2.41$, $p < .02$. Neither of the two remaining keyword versus control comparisons was statistically significant.

Finally, concerning differences between direct and indirect keyword item types, performance was consistently higher on the

former. The item-type difference was statistically greater in the fully structured condition (a mean difference of 17%) than in the control condition (4%), $t = 2.62$, $p < .01$; and although a comparable t of 2.52 was found for the semi-structured versus control comparison, it did not reach the needed Dunn critical value of 2.53. None of the remaining item-type interaction comparisons--including the three-way interactions involving conditions, ability, and item types--was statistically significant.

Discussion

The most prominent outcome of the present study was the interaction between variations of the keyword method of vocabulary learning and a measure of children's verbal ability (viz., vocabulary knowledge). Whereas higher and lower ability students differed considerably in performance when they were required to generate both keywords and relational images on their own, the difference all but disappeared when the experimenter provided the students with appropriate keywords and illustrations. Such a result cannot be attributed to a ceiling effect in the fully structured condition, in that: (a) the mean performance of higher ability students in that condition (64.2%) was nowhere near perfection; and (b) the variability associated with such subjects' performance was not at all reduced (in fact, it was the second largest of the eight cells).

The unstructured keyword variation also contributed to a complementary interaction involving the no-strategy control condition. In this case, the benefits of the keyword method, relative to control instructions, were much more pronounced for higher ability subjects. Similar mean-and-variability arguments

can be offered to dismiss a floor-effect interpretation of this interaction.

The present findings are basically consistent with previous conclusions about the efficacy of associative-learning strategies in children (e.g., Levin, 1976; Pressley, 1977b; Rohwer, 1973). In particular, two generalizations that seem to be warranted on the basis of the data summarized in Table 1 are: (1) higher ability students would be expected to benefit more than lower ability students from a relatively complex (unstructured) strategy; but (2) such differences would be expected to diminish considerably when the strategy is simplified (structured). Some additional features of the present data are worth mentioning as well.

First, it is obvious that the ATIs uncovered here are ordinal in nature. That is, the performance of the higher ability students surpasses that of the lower ability students in all conditions, but by different amounts. In some cases the difference is considerable (viz., when an unstructured strategy is used), and in other cases the difference is negligible (viz., when a structured strategy is used). Note, however, that no claims can be made that use of one keyword-strategy variation or another will enable lower ability students to outperform higher ability students. (Of course, such an outcome becomes more plausible when treatments are carefully matched or mismatched with aptitudes, or when aptitudes are defined by specific skills, preferences, or interests, rather than by general cognitive ability.)

Second, the pattern of results involving the semi-structured and unstructured conditions might be taken as

evidence that generating one's own keywords are better than using someone else's--especially for higher ability students. This conclusion would be a mistake, we think, and appears premature for a couple of reasons. Foremost, a number of recent studies have examined the keyword generated/keyword provided question, with no consistency of findings (see Pressley et al., in press). In addition, it must be remembered that in the present study, the students had to read the vocabulary items on their own and, at the same time, either read or generate keywords. Because of this format, keyword-provided (semi-structured) subjects would be expected to be at a disadvantage, relative to keyword-generated (unstructured) subjects, to the extent that vocabulary/keyword reading mismatches occurred and/or the vocabulary items and keywords were not well integrated. In previous keyword research with children, care has been taken to avoid these problems by having the experimenter read the vocabulary items and keywords, as well as provide an initial separate phase for keyword mastery (e.g., Levin, McCormick, Miller, Berry, & Pressley, in press). Because of the constraints created by the specific conditions and interaction questions of interest here, however, it was thought best that students read all items on their own and that the keyword-learning and definition-learning phases be combined. (Combining the two phases has recently been found to have an adverse effect with children of this age--see Levin, Berry, Miller, & Bartell, in press.)

Finally, a few words concerning the educational significance of these findings are in order. It is not premature to conclude that the keyword method is a highly

effective strategy for remembering new vocabulary, as well as other factual information with an associative component (see Pressley et al., in press). This is true for subject populations of all kinds, ranging in age from preschoolers to adults. What is added by the present study, however, is that how the keyword method is implemented will likely make a difference for students of different ability. That is, just as Pressley and Levin (1978) have found that younger students cannot benefit as much as older students from less structured variations of the keyword method, the same can now be said of upper elementary school children who differ in one aspect of verbal ability (namely, vocabulary knowledge). Surely other individual differences dimensions are relevant as well, which ought to be investigated in conjunction wih the two separate requisites of the unstructured variation of the method--keyword generation and imagery generation.

Reference Note

1. Mullis, L. L. Second-language vocabulary learning: An interaction of the keyword method, mediation, and level of divergent thinking. Unpublished doctoral dissertation, Mississippi State University, 1976.

References

Atkinson, R. C., & Raugh, M. R. An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, 126-133.

Delaney, H. D. Interaction of individual differences with visual and verbal elaboration instructions. Journal of Educational Psychology, 1978, 70, 306-318.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, Calif.: Brooks/Cole, 1968.

Levin, J. R. Determining sample size for planned and post hoc analysis of variance comparisons. Journal of Educational Measurement, 1975, 12, 99-108.

Levin, J. R. What have we learned about maximizing what children learn? In J. R. Levin & V. L. Allen (Eds.), Cognitive learning in children: Theories and strategies. New York: Academic Press, 1976.

Levin, J. R. The mnemonic '80s: Keywords in the classroom. Educational Psychologist, 1981, 16, 65-82.

Levin, J. R., Berry, J. K., Miller, G. E., & Bartell, N. P. More on how (and how not) to remember the states and their capitals. Elementary School Journal, in press.

Levin, J. R., McCormick, C. B., Miller, G. E., Berry, J. K., & Pressley, M. Mnemonic versus nonmnemonic vocabulary-learning strategies for children. American Educational Research Journal, in press.

Pressley, M. Children's use of the keyword method to learn simple Spanish vocabulary words. Journal of Educational Psychology, 1977, 69, 465-472. (a)

Pressley, M. Imagery and children's learning: Putting the picture in developmental perspective. Review of Educational Research, 1977, 47, 585-622. (b)

Pressley, M., & Levin, J. R. Developmental constraints associated with children's use of the keyword method of foreign language vocabulary learning. Journal of Experimental Child Psychology, 1978, 26, 359-372.

Pressley, M., Levin, J. R., & Delaney, H. D. The mnemonic keyword method. Review of Educational Research, in press.

Pressley, M., Levin, J. R., Nakamura, G. V., Hope, D. J., Bispo, J. G., & Toye, A. R. The keyword method of foreign vocabulary learning: An investigation of its generalizability. Journal of Applied Psychology, 1980, 65, 635-642.

Rohwer, W. D., Jr. Elaboration and learning in childhood and adolescence. In H. W. Reese (Ed.), Advances in child development and behavior (Vol. 8). New York: Academic Press, 1973.

Thorndike, R. L., & Hagen, E. Cognitive abilities test. Boston: Houghton-Mifflin, 1971.

Table 1

Mean Percent Correct Across Item Types, by Condition and Ability Level

		Condition			
		Control	Unstructured	Semi-Structured	Fully Structured
Ability Level	Higher	35.9	74.3	54.9	64.2
	Lower	17.0	33.7	34.9	56.2
	Across Ability	26.5	54.0	44.9	60.2

Note: MSE(136) = 364.7