DOCUMENT RESUME

ED 091 430 95 TM 003 644

AUTHOR Steve, Michael H.; Tennyson, Robert D.
TITLE Prior Memorization of Definitions, Examples and Nonexamples with Conjunctive and Disjunctive Concept Learning Tasks.

SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C. Regional Research Program.
BUREAU NO BR-2-D-027
PUB DATE [Apr 74]
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Classification; *Concept Formation; Definitions; Junior High School Students; *Memorizing; Time Factors (Learning); *Verbal Learning

ABSTRACT

The effect on concept learning of requiring the memorization of either examples or nonexamples prior to going through a theoretically effective training program was compared to the performance of groups of seventh and eighth graders who either memorized nothing or memorized subconcepts of the concept definition. Correct classification scores and undergeneralization error scores were the primary dependent variables. With both a disjunctive and a conjunctive concept, no significant treatment differences were found with these variables. The three prior-memorization groups took significantly less time to reach criterion in the training program, but took significantly more total instructional time than did the no-prior-memorization groups. (Author)

PRIOR MEMORIZATION OF DEFINITIONS, EXAMPLES AND NONEXAMPLES WITH CONJUNCTIVE AND DISJUNCTIVE CONCEPT LEARNING TASKS

Michael H. Steve and Robert D. Tennyson

Florida State University

Presented at the

American Educational Research Association
Annual Meeting
1974

Chicago, Illinois

PRIOR MEMORIZATION OF DEFINITIONS, EXAMPLES AND NONEXAMPLES WITH CONJUNCTIVE AND DISJUNCTIVE CONCEPT LEARNING TASKS

Michael H. Steve and Robert D. Tennyson
Florida State University

Abstract

The effects of four instructional sequences on concept acquisition were compared. Memorization of different concept task components was required prior to a training program. Groups were required to memorize either examples, nonexamples, key words in concept definitions, or nothing. Correct classification scores, undergeneralization and overgeneralization error scores and latency times were the primary dependent variables. With both a disjunctive and conjunctive concept, the four treatments appeared to be equally effective. The three prior memorization groups spent less time to reach criterion in the training program, but were less efficient than the no prior memorization groups when total instructional time was considered.

PRIOR MEMORIZATION OF DEFINITIONS, EXAMPLES AND NONEXAMPLES WITH CONJUNCTIVE AND DISJUNCTIVE CONCEPT LEARNING TASKS[1]

Michael H. Steve[2] and Robert D. Tennyson
Florida State University

Recall of verbal information from concept tasks is a different type of capability than correctly classifying instances (Gagné, 1970; Merrill & Boutwell, 1973). Although both types of behaviors are important in instructional situations, learning a defined concept always entails the capability of the learner to correctly classify instances as either examples or nonexamples according to a definition. The objective of this study was to assess the effects of memorization of verbal information comprising concept tasks on the acquisition of correct classification skills. In this study the memorization of examples or nonexamples or the memorization of key words from a concept definition preceded a training program designed to teach correct classification skills.

Instructional science research has demonstrated that variables dealing with the critical and the irrelevant attributes of a concept can be important for the elicitation of correct classification behaviors in a teaching situation. Tennyson, Woolley, & Merrill (1972) demonstrated that displays of examples and nonexamples which contrast the critical attributes with the irrelevant attributes lead to fewer overgeneralization and undergeneralization errors. Markle and Tiemann have theoretically postulated (1969) and empirically demonstrated (1972) that by presenting sets of examples and nonexamples which represent the full range of example and nonexample possibilities, undergeneralization and overgeneralization errors can be minimized. Presentation of the concept definition along with the systematic assemblage of examples and nonexamples has provided additional increments of concept acquisition success (Merrill & Tennyson, 1971; Feldman & Klausmeier, 1973). The compatible nature of the instructional design variables researched above suggests that an effective concept teaching paradigm is available. Such a paradigm was used in this study, although its effictiveness was not tested. Instead, further instructional design modifications were introduced and evaluated.

Gagné (1970) classifies learning defined concepts as a specific type of rule learning. The definition of the concept, which can be considered a rule statement, is used for teaching and communication purposes. Just as correct rule application skills are not guaranteed by the recall of rule statements or of problem solutions, so correct classification behavior is not guaranteed by the recall of the concept definition or of instances. However, memorization of the concept definition and of examples and nonexamples of the concept may influence the amount of instructional time needed to teach such capabilities and may influence the resultant classification capabilities themselves.

Such hypotheses were tested in this experiment. The independent variable was the type of information required of students to memorize prior to classification training and testing. The dependent variables were correct classification scores, overgeneralization and undergeneralization error scores, and latency times for different instructional segments. Overgeneralization scores refer to the number of nonexamples erroneously classified as examples, while undergeneralization scores refer to the number of examples classified as nonexamples.

Method

Subjects

A total of 92 subjects from the Florida State University Developmental Research School were used in this study. Data from one subject was discarded because she answered indiscriminately on the posttest. Of the 91 remaining, data from 17 subjects were not analyzed because these subjects failed to reach criterion on one or both of the concepts. Of the 74 subjects from whom complete data were collected, 33 were seventh graders, 41 were eighth graders. Twenty eight subjects were males, 46 were females. Because of the large number of subjects dropped from the study, a selection bias may have resulted. This possibility is addressed later.

Learning Task

The experimental session consisted of three main phases:

(a) memorization of concept task components, (b) training of correct classification behaviors, and (c) testing of correct classification behaviors. Except for the information memorized in the memorization phase, the experimental presentation was the same for all subjects. The instructional objective in the training phase was: Given an unfamiliar instance, the subject will correctly identify it as either an example or a nonexample. In the testing phase, subjects were required to classify previously unencountered instances.

Two concept definitions were constructed for this experiment. They appeared in the following format throughout the experiment:

1. A <u>Skeethand</u> is a hand of five cards which:
   a. has no card appearing more than once
   b. has all cards lower than 10
   c. contains a 2, 5, and a 9
2. A <u>Derf</u> is a series of letters which has <u>either</u>:
   a. no vowels
   b. no consonants
   c. one or more letters occurring twice

These concept definitions were chosen for a number of reasons. First, both definitions allowed for the construction of an infinite number of instances. No instances would appear more than once. Second, the definition would be new to all subjects. Third, each concept is governed by a different conceptual rule, i.e., Skeethand is a conjunctive concept, while Derf is a disjunctive concept. Use of two types of concepts should increase the generalization of results. Fourth, a standard dictionary format for the definition was followed for both concepts. Both definitions described the general class to which the concept belonged, and then how the defined instances differed from other members in the general class, i.e., definition by genus and difference (Copi, 1972). Thus, a hand of five cards is the genus, and the critical attributes differentiate Skeethand from other kinds of hands containing five cards. Fifth, it was assumed that all subconcepts of the definition were familiar to the subjects and that all critical and irrelevant attributes were easily identifiable in the instances.

A standard teaching display was used throughout the training phase. It consisted of the concept definition and six instances (three examples and three nonexamples). The concept teaching paradigm was a result of extending the empirical work of Tennyson, et al. (1972), Tennyson (1973), and Markle and Tiemann (1972). Their research on the effects of different stimulus similarity variables in deductive concept teaching situations was incorporated into the following instructional design algorithm:

When teaching conjunctive concepts:

1. Select K examples (K refers to the number of critical attributes in the concept definition; K=3 for both concepts in this investigation), such that together they exhibit the fullest range of irrelevant attributes.

2. Select K nonexamples, each having all critical attributes except one and each lacking a different critical attribute.
3. Select the K nonexamples such that when each is paired with one of the K examples, the example-nonexample pair shares the same irrelevant attributes.

When teaching disjunctive concepts:

1. Select K examples, each having only one of the critical attribute and each having a different critical attribute.
2. Select K nonexamples such that together they exhibit the fullest range of irrelevant attributes possible.
3. Select the K nonexamples such that when each is paired with one of the K examples, the example-nonexample pair shares the same irrelevant attributes.

This algorithm was followed for the construction of all teaching displays.

While the teaching display for the definition and the six instances was visible, explanations were offered as to why each of these instances were classified as either examples or nonexamples. These explanations appeared one at a time at the bottom of the cathode ray terminal (CRT) screen. An attempt was made to to reference these explanations as much as possible to the critical attributes in the concept definition. For example, an explanation for a nonexample of the Skeethand concept was: "Hand #4 does not meet requirement #1 (Note the two 3s)." Critical attributes were referred to as requirements in this experiment. The complete experimental program was presented on CRTs by an IBM 1500 computer system.

Experimental Design

The independent variable involved four conditions in the memorization phase. Subjects in the EX and NEX groups were required to memorize examples and nonexamples, respectively. Those in the DEF group were required to memorize selected key words (subconcepts) from the concept definitions. Those in the NULL group were not required to memorize anything and were passed directly to the training phase. Sex was crossed with the four memorization conditions, resulting in a 4 x 2 factorial design. Because males participated in experimental sessions at the beginning of the week and females in the latter part of the week, the sex variable is confounded with a time variable. This sex-time variable was used only as a blocking variable in the analysis, and the statistical significance of amounts of variance it accounted for, by itself or in interaction with the treatment variable, was not tested. The significance level of $p < .05$ was used for all statistical tests.

Treatment Programs

Each group except the NULL group passed through the memorization phase once for each concept. The EX and NEX groups memorized a total of 4 instances; 3 of one concept and 1 of the other. The tasks for these groups were to type the example(s) or nonexample(s) from memory. Each time the subject did not answer correctly, he was shown the example(s) or nonexample(s) and then asked to type them again from memory.

The initial randomization procedure determined the concepts for which subjects would memorize one and three instances. Such a procedure made it possible to adjudge the effects of memorization of different numbers of instances. Such a comparison was considered ancillary to the contrasts implied in the experimental design. Examples and nonexamples were chosen from the first teaching display of the training program. Therefore, a full range of examples and nonexamples were represented in the three-instance cases. The one-instance cases were randomly selected from the three-instance cases prior to the experiment

The DEF group memorized a set of key words (subconcepts) in each definition. Memorized in the Skeethand concept were: five, no, once, all, 10, 2, 5, and 9. For the Derf concept, the words memorized were: letters, either, no, more, and twice. This group was given an incomplete definition and asked to type the missing words from memory into the incomplete definition. Incorrect responses required subjects to study the complete definition and to try the incomplete definition task again.

Procedure

A number table randomly assigned subjects to one of the four experimental conditions and to a CRT booth. The program was individualized and each student was instructed to proceed at his own pace until completed. The CRT presented instructions for operating the terminals. Sample definitions of familiar concepts were displayed in the format that was used in the experimental tasks. It was explained how these definitions could be used to divide instances into example and nonexample groups and how this was their task in the experiment. Subjects were then familiarized with what was meant by "a hand of cards," "suit," and "rank," concepts prerequisite to the Skeethand task.

At this point, one concept was randomly assigned. A teaching display was presented for one minute. During this time the subject could familiarize himself with the concept and six instances. All groups except the NULL group then entered the memorization phase. Subjects were looped through the memorization phase until they could recall their respective task components with 100% accuracy.

The first teaching display of the training phase was then presented. After studying the display and the six explanations, subjects were tested on four unencountered instances. If they correctly classified all four, they were passed on to the second concept or to the testing phase. If they did not reach the four-for-four criterion, they were again passed through the training phase with a series of displays containing the same definition and six new instances and six new explanations. If any subject failed to reach 100% criterion on his or her fourth attempt through the training phase for either concept, he or she was dropped from the study and his data were not included in the analyses.

After successfully passing through the memorization and training phases for both concepts, subjects in the EX, NEX, and DEF groups were shown the task components they had memorized earlier in the program. The EX and NEX groups studied four instances. The DEF group studied the two definitions, the words they had memorized were underlined. After these three groups had studied their respective displays for one minute, they were administered a posttest designed to assess classification competency. The NULL group was administered the posttest directly after reaching criterion on the second concept in the training phase. After completing the posttest, subjects were ushered into an adjoining room.

Tests

Training phase test items and posttest items were parallel in form. All instances used were members of the genus. Therefore, Skeethand test items were always made up of five cards and Derf test items always were made up of only letters. Critical in concept acquisition research is the array of unencountered instances used in training and testing. Just as the displays in the training program were designed to insure full generalization and proper discrimination, so the unencountered test instances were sampled from the domain of all possible attribute combinations.

Results

Subjects Failing to Meet Criterion

Multiple linear regression techniques, as outlined by Cronbach and Snow (1969) and by Bottenberg and Ward (1963), were used to analyze the 17 subjects who did not meet criterion in the training phase. California Test of Mental Maturity (CTMM) scores and age were coded as continuous vectors. A subject's group membership was represented by four dummy vectors of 1s and 0s. To test for possible Aptitude X Treatment interactions (ATI), interaction vectors were constructed between group and CTMM score vectors and between group and age vectors. The criterion variable was the dichotomous variable pass-or-fail from the training program. The stepwise procedure used here for the testing of main and interaction effects is a modificaiton of the Bottenberg and Ward approach (1963, p.95). The main effects were examined by creating a full model with Group, CTMM, and Age main effect vectors as predictors. The significance of each variable was tested by forming an appropriate reduced model and then testing for the reduction in the multiple correlation. The significance of interaction effects was tested by alternately including a CTMM X Group vector and a Age X Group vector with the other vectors included in the full model described above and then testing for the increase in the multiple correlation. The results of these analyses appear in Table 1. The only effect significant was the CTMM effect. Thus, a subject's

CTMM score could help predict whether that subject passed or failed one of the concept tasks in the training program.

- - - - - - - - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - - - - - - - -

Because the scores of those subjects who failed to meet criterion in the training phase were not included in the computation of posttest and latency statistics, a selection bias could have been operative in comparative group analyses on these variables. Because those subjects who were dropped had significantly lower scores than the group as a whole, results using the successful subjects are not readily generalizable to the experimental population as a whole. On the other hand, the fact that neither the Group nor the CTMM X Group or the Age X Group effects were significant supports the contention that group comparison tests using the curtailed data base are interpretable despite the possible selection bias.

Variables

An analysis of covariance statistical model was selected to test the implied null hypotheses. The two covariate variables were CTMM scores and age. There were no significant prior experimental differences among the group means for either variable.

Learning success. The 17 subjects who failed to meet criterion represented 19% of the total population. For the subjects reaching criterion, correct classification averages on the full 40 item posttest were 18.8 (94%) for the conjunctive concept and 15.3 (76%) for the disjunctive concept. Eighty percent of these subjects correctly classified 80% or more of the items on the full posttest.

On the posttest, mean correct classification scores and overgeneralization and undergeneralization mean error scores were analyzed for each concept separately and then in combination, resulting in nine separate $\underline{F}$-tests. Each of these tests resulted in $\underline{F}$s less than unity. These results suggest that there is no difference in the effectiveness of the four instructional sequences.

A within-subjects design was used to test the effects of the number of instances memorized. Only subjects from the EX and NEX groups were included in this ancillary analysis. For each of the nine dependent variables defined above, mean scores for the three-instance and one-instance cases were not significantly different.

Latencies. Three latency comparison tests were made: (a) training phase latency, (b) posttest phase latency, and (c) total program latency. Means are presented in Table 2. Analysis of training phase latency group means resulted in a significant $\underline{F}$ test ($\underline{F} = 3.37$; $df = 3/64$; $\underline{p} < .05$). A Newman-Keuls test was used to make pairwise group comparisons using the adjusted means. The only significant comparison showed that the NULL group took significantly longer in the training program than did the DEF group ($\underline{p} < .05$). On the average the more time each group took in the memorization phase, the less time they took to reach criterion in the training phase.

- - - - - - - - - - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - - - - - - - - -

An analysis of posttest phase latency mean scores revealed no significant differences. A comparision of mean total program latencies, defined as the sum of the memorization, training, and posttest phase latencies, resulted in an overall $\underline{F} = 7.44$ ($\underline{p} < .05$). A Newman-Keuls test on the adjusted group means showed that the NULL group took significantly less time to complete the total program than did either of the three proir memorization groups, $\underline{p} < .01$ for each of the three pairwise tests. There were no differences among the adjusted group means for the three prior memorization groups on the total program latency variable.

Discussion

No significant differences were found between any group means on any variable measuring the type of errors made or number of errors made either in the program or on the posttest. Furthermore, a regression analysis demonstrated that the instructional treatment variable was not a significant factor in determining which subjects would reach criterion in the training program. These results suggest two conclusions. First, there is no evidence to suggest that any treatment program determined an internal organization of the rule in the learner that was systematically more effective from that created by any other of the three treatment programs. Second, the results indicate that the four training programs are equally effective.

Because there were no differences among the major success indices for the different treatment programs, the instructional efficiency of programs needs to be assessed in order to decide the optimality of each program. Since the three prior memorization groups attained criterion earlier in the training program than the no prior memorization group, it appears that prior memorization of either examples, nonexamples, or subconcepts of a definition facilitates the acquisition of correct classification behaviors. However, the amount of time these three groups spent in the memorization phase was considerably more than the time they subsequently saved in the training phase. On the average, the three prior memorization groups spent 6.4 minutes in the memorization phase and 6.2 minutes (adjusted) to reach criterion in the training phase. The no prior memorization group spent no time in the memorization phase and took 7.0 minutes (adjusted) to reach criterion in the training phase.

The learners' capability to classify unencountered instances may be incremented by memorizing a greater number of examples or nonexamples or by memorizing more of the concept definition than was required in this experiment. However, any increment in posttest performance would probably not justify the amount of time learners would have to spend memorizing concept task components. Thus, if the goal of instruction involves only the correct classification capabilities of learners, the results of this investigation suggest that prior memorization of concept ponents is not an advisable instructional technique.

Clark's (1971) review concluded that disjunctive concepts in concept attainment tasks were more difficult to learn than conjunctive concepts. The poorer performance on the disjunctive concept items in this experiment would appear to extend Clark's conclusion to tasks involving a deductive teaching mode (Glaser, 1968). The correlation between correct classification scores on the posttest for the disjunctive and conjunctive concepts was $\underline{r}$ = .09. This low correlation suggests that the learning requirements for classification tasks involving conjunctive concepts are functionally different than those of disjunctive concepts.

While many subjects did well, 19% of the total subject pool could not correctly classify 4 instances in a row after seeing the definition on at least five occassions and after 40 different instances were correctly classified for them. The subjects who had problems may have benefited from a completely different teaching strategy. On the other hand, intrinsic learning problems could have been corrected with a remedial sequence of instruction. Verbal interaction with some of the subjects who failed to reach criterion revealed that two major problems existed. First, the disjunctive rule in concept definitions appeared to be unfamiliar to students and difficult to use. Second, working with the first two critical attributes in the disjunctive concept was difficult because they were stated in the negative. Both of these difficulties help explain why the disjunctive concept task was so difficult.

References

Bottenburg, R.A., & Ward, J.H., Jr. Applied multiple linear regression. (Technical Documentary Report PRL-TDR-63-6) Lackland Air Force Base, Texas: Aerospace Medical Division, November, 1966 reprint (Available from National Technical Information Service, AD No. 413128, Springfield, Va. 22151).

Clark,.D.C. Teaching concepts in the classroom: A set of teaching prescriptions derived from experimental research. Journal of Educational Psychology Monographs, 1971, 62, 253-278.

Copi, I.M. Introduction to logic. (4th ed.) New York. MacMillan Company,1972.

Cronbach, L.J., & Snow, R.E. Individual differences in learning ability as a function of instructional variables. (Final Report, United States Office of Education Contract No. OEC 4-6-061269-1217). Stanford, Calif.: Stanford University, 1969.

Feldman, K.V., & Klausmeier, H.J. The effects of adding emphasis of relevant attributes and a concept definition to a rational set of concept examples and nonexamples. Paper presented at the meeting of the American Educational Research Association, New Orleans, March, 1973.

Gagné, R.M. The conditions of learning. (2nd ed.) New York: Holt, Rinehart, and Winston, Inc., 1970.

Glaser, R. Concept learning and concept teaching. In R. Gagné and W.J. Gephart (Eds.), Learning research and school subjects. Itasca, Ill.: F.E. Peacock Publishers, 1968.

Markle, S.M., & Tiemann, P.W. Really understanding concepts. Champaign, Illinois: Stipes, 1969.

Markle, S.M., & Tiemann, P.W. Some principles of instructional design at higher cognitive levels. Presentation to the Lehrmaschinen in Verbingung mit einer Fachausstellung, Berline, Germany, April 5, 1972.

Merrill, M.D. & Boutwell, R. Instructional development: Methodology and research. In F.N. Kerlinger (Ed.), Review of research in education. Vol. 1. Itasca, Ill.: F.E. Peacock Publishers, 1973.

Merrill, M.D., & Tennyson, R.D. The effects of types of positive and negative examples on learning concepts in the classroom. (Final Report, United States Office of Education Contract No. 0-H-014). Provo, Utah: Brigham Young University, 1971.

Tennyson, R.D. Effect of negative instances in concept acquisition using a verbal learning task. Journal of Educational Psychology, 1973, 64, 247-260.

Tennyson, R.D., Woolley, F.R., & Merrill, M.D. Exemplar and nonexemplar variables which produce correct concept classification behavior and specified classification errors. Journal of Educational Psychology, 1972, 63, 144-152.

Footnotes

[1]This experiment was supported by the United States Office of Education, No. 2-D-027. The authors would like to express their appreciation to Edna Reynolds for her part in programming the instructional materials on the computer.

[2]Requests for reprints should be sent to Michael H. Steve, Computer Applications Laboratory, Florida State University, 32306.

TABLE 1

Summary of Multiple Linear Regression

Analyses with Pass-Fail Criterion

Effect	df	% of Variance	F	P
Group	3,85	04	1.39	
CTMM	1,85	20	23.75	< .001
Age	1,85	00	< 1	
CTMM X Group	1,82	03	1.14	
Age X Group	1,82	04	1.64	

TABLE 2

Latency Means for Training and Testing

Phases, and Total Program

	Latencies (in min.)			
Groups	Memorization Phase	Training Phase	Posttest Phase	Total Program
EX	(5.6)	(8.6)9.1[a]	(5.2)5.2	(19.4)20.1
NEX	(6.3)	(8.9)8.2	(5.4)5.4	(20.5)19.5
DEF	(7.5)	(7.6)7.3	(4.8)4.8	(19.9)19.7
NULL	(0)	(9.6)9.9	(5.3)5.3	(14.9)15.4

[a]Times enclosed in parentheses are unadjusted means; those not enclosed are adjusted means.