DOCUMENT RESUME

ED 093 271 95 IR 000 752

AUTHOR Tennyson, Robert D.; Steve, Michael H.
TITLE Classroom Concepts: Learner Strategy Versus Instructional Strategy. Final Report.
INSTITUTION Florida State Univ., Tallahassee. Computer Applications Lab.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-2-D-027
PUB DATE Dec 73
GRANT OEG-4-72-0015
NOTE 38p.

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS College Students; Concept Formation; *Concept Teaching; Grade 7; Grade 8; Intermode Differences; Memorizing; *Organization: *Prompting; *Sequential Learning; Serial Learning

ABSTRACT

In the first of three studies, separately reported, the effects of prompting and sequencing on a science concept task were studied with college students. The data analysis showed that the prompting procedure was significantly different from a no-prompting condition; prompting seemed to negate the affect of the defined concept instructional sequence. The second study used college psychology students to examine concept learning under the variables of prompting and sequencing. The findings indicated that the prompted/sequenced treatment resulted in less time to complete the task and with fewer errors. The third study used seventh and eighth graders to study the effect of prior memorization of either examples or non-examples on concept formation. While there were no significant results, the three prior-memorization groups spent less time reaching criterion in the training program, but took significantly more total instructional time than did the no-prior-memorization group. (WH)

FINAL REPORT
Project No. 2-D-027
Grant No. OEG-4-72-0015

CLASSROOM CONCEPTS: LEARNER STRATEGY VERSUS INSTRUCTIONAL STRATEGY

Robert D. Tennyson
Michael H. Steve
Florida State University
Tallahassee, Fla. 32306

December 1973

U. S. DEPARTMENT OF
HEALTH, EDUCATION AND WELFARE

Office of Education
Bureau of Research

FINAL REPORT

Project No. 2-D-027
Grant No. OEG-4-72-0015

CLASSROOM CONCEPTS: LEARNER STRATEGY VERSUS INSTRUCTIONAL STRATEGY

Study 1 Strategy Prompting and Sequence Effect on Concept Acquisition
Study 2 Attribute Prompting and Task Sequence in Efficient Concept Acquisition
Study 3 The Effects of Prior Memorization of Definition Components on Concept Acquisition Using an Effective Training Program

Robert D. Tennyson
Michael H. Steve
Florida State University
Tallahassee, Fla. 32306

December 1973

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

Acknowledgements

This project was conducted as part of a continuing research program in the Computer Applications Laboratory at Florida State University on variables and conditions related to the instructional process. Investigations using computer technology involve a number of people to implement the study. One individual who contributed to the final design was Ms. Edna Reynolds, who programmed and coded the two experiments using computers. For the one study which was conducted at Bucknell University, the authors express thanks to Dr. Richard C. Boutwell for administering it. In addition, the authors acknowledge the contributions of Ms. Dorothy Carr and Ms. Linda Morse, who were responsible not only for the final report, but also for contributions during the project.

As a result of the sponsored research reported here, several other studies have been conducted as extensions. Certainly, the funds provided in this one research project will continue to have payoff in both future research activities and in the applied usage of the variables in the classroom environment.

CONTENTS

STRATEGY PROMPTING AND SEQUENCE EFFECT ON CONCEPT ACQUISITION

Robert D. Tennyson and Michael Steve
Florida State University

Abstract

Instructional task variables of prompting and sequencing were studied as an extension of a defined concept acquisition paradigm. The independent variable of prompting consisted of two components; a procedure for focusing the subject's attention on the critical attributes of the given concept, and a presentation of the strategy used to determine classification of the examples. Sequencing of instances involved an organized presentation based upon the relationship of the stimulus attributes. Subjects, college students, were presented a science concept task. The data analysis showed that the prompting procedure was significantly different from a no-prompting condition ($p > .05$). Concluded was the notion that prompting seemed to negate the affect of the defined concept instructional sequence.

STRATEGY PROMPTING AND SEQUENCE EFFECT ON CONCEPT ACQUISITION

Robert D. Tennyson[1] and Michael Steve
Florida State University

Facilitation of concept learning has been demonstrated in a series of studies (Tennyson, Woolley, & Merrill, 1972; Tennyson, 1973; Tennyson, Merrill, Young, & Low, 1974) by use of an instructional paradigm which sequences positive and negative instances by defined relationships of attributes and instance difficulty. The basic premise of the paradigm was that acquisition of a given concept can be optimized by the appropriate manipulation of task variables. These variables included: (a) the display characteristics of the instances, i.e., two instances are matched when their irrelevant attributes are as similar as possible and divergent when their irrelevant attributes are as different as possible; (b) relative difficulty of the instances; and (c) additional information given to facilitate attention to relevant aspects of an instance. The objective of the concept instructional paradigm was to insure correct classification behavior (all instances correctly identified) while preventing the errors of overgeneralization (negative instances similar to class members identified as positive), undergeneralization (positive instances identified as negative), and misconception (instances sharing a common irrelevant attribute(s) identified as class members) (Markle & Tiemann, 1969).

The purpose of this study was to extend the concept acquisition model by investigating the task variables of prompting and sequencing. Specifically, the prompting stimuli would be used to identify not only the critical attributes of a given positive instance, but also as a strategy for recognizing those attributes. The strategy procedure would: (a) focus the learner's attention to the critical attributes by demonstrating the divergent relationship between two examples; and (b) describe the method used to determine a given example classification. The hypothesis was that strategy prompting that identifies critical differences between instances would significantly increase the effectiveness of the concept acquisition paradigm.

The second task variable, not previously investigated in the concept acquisition research, but of concern in instruction design, is sequence of instances. An assumption of the concept model is that two examples should be simultaneously contrasted to focus on the divergency of their irrelevant attributes, and that the nonexamples matched to the examples should be likewise presented simultaneously. To study this sequence variable, a second hypothesis was that the organized presentation of instances would result in a significantly higher classification score than a random sequence of the same instances.

Method

Subjects and design. Students (total 155) enrolled in the core course on foundations of education at Bucknell University were used as subjects. Participation in the experiment was used to fulfill a course requirement. A posttest-only factorial design, with one main effect being the three conditions of prompting and the second, the two levels of sequencing, was used in this experiment (Campbell & Stanley, 1963).

Learning task. The instructional objective for the experimental task required the subjects to perform classification behavior by identifying previously unencountered instances of RX_2 crystals. This science concept was selected because it is similar to the type of classroom subjects taught in undergraduate chemistry curriculum, thus generalizability to other subject matters would facilitate applicability. Six self-instructional learning tasks were designed each using the same general format presentation, i.e., an introduction to the concept of crystals and the task requirements, presentation of the RX_2 crystal definition, additional information on crystal identification, the presentation of the instances, and the posttest on crystal identification. The definition (critical attributes) of RX_2 crystals focused the subject's attention to the basic, repeating, two-to-one ratio in crystal structure of the RX_2 crystal. Each page of the learning task consisted of two crystal pictures taken from Crystal Structures (Wyckoff, 1968). Reproductions of the pictures were made from photo copies that provided shaded crystals. Crystals were shaded so that depth perception would not confound identification.

Mode of presentation consisted of an inquisitory form which required subjects to identify an instance as either an example or nonexample. In each program, after the subject's response, the instances were displayed for a second time, but with the appropriate prompting treatment. Instances were grouped into quads according to the relationship of the stimulus attributes. The two examples per quad were selected by the divergency of their defined common irrelevant attribute(s), e.g., the dimension of symmetry, the dimension of ratio recognition (whether atoms were horizontal or vertical), the size and/or color of the atoms, the number of the atom structure, and the process of subdividing the crystal structure to obtain the underlying pattern. And, for each example in the quad, a nonexample was matched to it by the similarity of their respective irrelevant attributes. This followed the concept paradigm presented in the previous Tennyson studies (Tennyson et al., 1972; Tennyson, 1973). The two sequence variables, organized and randomized, used a range of instance difficulty obtained from a former empirical data analysis (Tennyson & Boutwell, 1973). In the organized sequence the crystal quads consisted of an example followed by a nonexample per page, with the quads progressing from easy to hard. While the randomized instances were presented in such a way as to have an example and a nonexample on each page, i.e., the quads, then, contained two examples and two nonexamples, but without the defined relationship. The learning task consisted of four quads for a total of eight examples and eight nonexamples.

The prompting variable consisted of additional information presented to the subjects following their responses to the unlabeled instances. The first consisted of a prompting technique that focused on the uniqueness of the critical attributes to that particular example. This was the same procedure followed in more traditional forms of prompting. However, the second prompting condition provided, in addition to the above prompting, a strategy on how to identify critical attributes by using information related to lower order attributes. These prerequisite attributes were the same as used by the experimenters to pair divergently the quad examples. Nonexamples were prompted with information explaining the absence of the critical attribute(s), and what type of procedures to use in recognizing that situation. A third condition, no-prompting, simply identified the instances as positive or negative.

The posttest was constructed to evaluate classification behavior, i.e., subjects responded to previously unencountered instances. Thirty instances, 15 examples, were selected from the same item pool as used to construct the programs. They met the criterion of selection based on their critical and irrelevant attributes, i.e., examples and nonexamples had a full complement of irrelevant attribute values, and nonexamples were lacking different critical attributes, this generally meant that the atoms were in different ratios because the ratio was the main critical attribute.

Procedures. Several experimental sessions were established to accommodate subject time schedules. The learning tasks were randomized prior to the sessions and assigned to subjects after the period began. Subjects, seated in alternate desks in a large classroom, were given the tasks, and read the directions silently while the experimenter read aloud. Once the subject began, no questions concerning the task were answered by the experimenter. Directions required the subjects to identify the four crystals per quad and mark their responses on the answer sheet. Following the responses per quad they proceeded to the next two pages to receive the given answers. Subjects continued through the self-instructional task until the final quad, at that point they were directed to either return for further study or to begin the posttest. The test was in a separate booklet with answer sheet and was given when requested by the subject.

Results

The dependent responses were analyzed according to errors on three scoring patterns: correct classification, overgeneralization, and undergeneralization. The first pattern, correct classification, represented the subjects' errors in identifying instances. Scoring patterns for the two classification errors were designed such that any response of a nonexample as an exemple was considered an overgeneralization error and failure to identify any example was an undergeneralization error. A separate two-way analysis of variance was used for each classification error. One main effect was the two levels of sequencing--organized and random. While the three forms of prompting, full (prompting strategy), partial, and no-prompting, was the second. Table 1 presents the means

TABLE 1

Mean Error Scores for Correct Classification and Overgeneralization

Sequence	Prompting Full	Prompting Partial	Prompting No
Organized	11.1[a]	11.2	12.3
	3.9[b]	4.2	5.1
	7.2[c]	7.0	7.2
Random	11.4	10.9	13.1
	3.8	4.0	5.9
	7.6	6.9	7.2

[a]The first rows are the correct classification error means.

[b]The second rows are the overgeneralization error means.

[c]The third rows are the undergeneralization error means.

for the correct classification, overgeneralization and undergeneralization error scores. The analysis of variance test for the undergeneralization dependent variable resulted in nonsignificance ($p > .05$) for both main effects and interaction, with statistical power (effect size of .30) at .65. In the other two classification behaviors, correct classification and overgeneralization, the main effect of sequencing and interaction were also nonsignificant at the .05 level.

The independent variable of prompting was investigated at three levels representing a no-prompting condition in which instances were just labeled as example or nonexample, a partial prompting treatment similar to previous methods used in the verbal information level of behavior in which mathemagenic material was used to help the learner remember the stimulus by acquiring information for cues during the criterion measure while the difference here was to focus the subjects' attention on the critical attributes in each example presented or the absence of such in the non-examples, and a third condition which did the above plus provided the subject with a strategy procedure for determining the critical attributes. Results of the correct classification analysis showed that there was a significant difference on this variable ($F = 4.12$, $df = 2/149$, $p < .025$). The difference between means showed the two prompting conditions being nonsignificant ($p > .05$), while the no-prompting error mean was significantly higher than the other two ($p < .05$). Likewise, on the overgeneralization analysis the F test was significant at .01 ($F = 7.92$), with the same mean relationship as the correct classification.

The independent variable of prompting was investigated at three levels representing a no-prompting condition in which instances were just labeled as example or nonexample, a partial prompting treatment similar to previous methods used in the verbal information level of behavior in which mathemagenic material was used to help the learner remember the stimulus by acquiring information for cues during the criterion measure while the difference here was to focus the subjects' attention on the critical attributes in each example presented or the absence of such in the non-examples, and a third condition which did the above plus provided the subject with a strategy procedure for determining the critical attributes. Results of the correct classification analysis showed that there was a significant difference on this variable ($F = 4.12$, $df = 2/149$, $p < .025$). The difference between means showed the two prompting conditions being nonsignificant ($p > .05$), while the no-prompting error mean was significantly higher than the other two ($p < .05$). Likewise, on the over-generalization analysis the F test was significant at .01 ($F = 7.92$), with the same mean relationship as the correct classification.

Discussion

In previously reported research on the concept acquisition paradigm the variables of prompting and sequencing were not investigated (Tennyson et al., 1972; Tennyson, 1973). Also, Clark's review (1971) showed that the prompting variable has not been studied in terms of isolating the critical attributes during the instructional period for help in focusing the learner's attention to the defined attributes. The purpose here was to present the instances in an inquisitory mode (previously only the expository method was used), and then explain why a given instance was either an example, because it had the critical attributes, or a nonexample, because it failed the conditions defined for class membership. An additional variable to reduce error in the acquisition of a concept, the prompting materials were supplemented with a verbal description of the strategy used to determine if that given instance was positive. However, this further information did not result in a significant decrease in errors over the former prompting condition. The directions introducing the three treatment conditions were the same, thus, the subjects in the strategy prompting program were not informed on the purpose or use of the additional information. Given directions on the purpose of the strategy could result in a useful tool for learning a skill in identifying new examples. The prompting conditions did show that a meaningful addition to the concept acquisition paradigm is the prompting of instances by specific attributes.

Sequencing of instances according to the organized method of two divergent examples matched to two nonexamples was less effective as a variable, and the test with a random condition seemed to be negated by the prompting variable. The purpose of the organized set of instances was to provide the subject with matching nonexamples to direct attention to the critical attributes of the examples. And, in the expository mode the subject would know the positive and negative, while in the inquisitory mode the subject had to select the identity before knowledge of results was given. Therefore, it seems that the prompting variable supplemented the need to present the organized sequence because the necessary condition of drawing attention to the critical attributes was given by the instruction.

Extensions of this study should investigate the effect of an interactive paradigm of prompting and sequence. That is, in a given set of instances the prompting would point out the matchedness or divergency by direct comparisons. From the results of this study and previous research, this would seem to have a positive effect on acquisition. Latency should also be included as a dependent variable to determine the time required to learn a given concept, and the time required to perform on the evaluation process. Such an interactive variable should result in a procedure for organization over a random presentation. However, with more research on the differing tasks it might be that sequence is a function of content, necessitating research on task classification variables, and their interaction with instructional design characteristics.

References

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

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

Markle, S. M., & Tiemann, P. W. Really understanding conceots. Champaign, Ill.: Stipes, 1969.

Tennyson, R. D., & Boutwell, R. C. Pre-task versus within-task anxiety measures in predicting performance on a concept acquisition task. Journal of Educational Psychology, 1973, 65, 88-92.

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

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

Wyckoff, R. W. Crystal structures (2nd ed.). New York: John Wiley & Sons, 1968.

Tennyson, R. D., Merrill, M. D., Young, J. I., & Low, W. Attribute matching, instance difficulty, and attribute isolation in concept acquisition. Submitted Journal of Educational Psychology, 1974.

Footnote

[1] Reprints may be requested from Robert D. Tennyson, Computer Applications Laboratory, Florida State University 32306.

ATTRIBUTE PROMPTING AND TASK SEQUENCE IN EFFICIENT CONCEPT ACQUISITION

Abstract

Investigated was the premise that efficient concept learning would result from a task designed to use the variables of prompting and sequence. The prompting variable consisted of isolating instance attributes as one treatment and no prompting as a second. The sequence variable tested an organized order of instances versus a random order. Undergraduate psychology students were measured according to performance responses and latencies for the task and tests. The findings indicated that the prompted/organized treatment resulted in less time to complete the task ($p < .01$), and fewer errors and less time on the posttest ($p < .01$) than the three other treatments.

ATTRIBUTE PROMPTING AND TASK SEQUENCE IN EFFICIENT CONCEPT ACQUISITION

This study investigated the premise that instructional design procedures result in increased learner effectiveness and efficiency. Recent studies (Tennyson, Woolley, & Merrill, 1972; Tennyson, 1973) investigated variables and conditions that have a direct application to the design of concept teaching. The instructional strategy (Tennyson, 1973) for concept acquisition consisted of presenting exemplars and nonexemplars to the learner in such a way that the critical attributes were clearly contrasted with the irrelevant attributes. The purpose of this study was to extend the Tennyson paradigm for concept instructional design by investigating the variables of attribute prompting (explanatory information indicating the critical attributes for each exemplar or the absence of critical attributes for each nonexemplar), and task sequence (display order of the exemplars and nonexemplars). The two main effects were crossed to form four treatments; (a) organized sequence with prompting, (b) organized with no prompting, (c) random with prompting, and (d) random with no prompting. Computer-aided instruction procedures were used to measure subject latency on the learning and performance tasks.

Method

Subjects and design. Subjects were undergraduate students from the general psychology subject pool at Florida State University, who were required to participate for course credit. Students were given the choice of selecting experiments from all available in the Department of Psychology. The experimental design was a 2 x 2 pretest/posttest design (Campbell & Stanley, 1963), in which the main effects were sequence and prompting. Dependent variables were error scores and latencies within the task and on the tests.

Learning task. A poetry task, modified from the Tennyson, Woolley, and Merrill (1972) study, was adapted for use on a computer teletype instruction system. The definition of trochaic meter (critical attributes) used in that previous study was given here for all four groups. The two organized treatment conditions received a series of four sets of instances composed of two divergent examples matched to two nonexamples. The prompting treatment included a statement which identified the initial attributes of the given concept and why they were relevant. For the nonexamples, the absence of the critical attributes was noted and explained. The two random tasks were developed by randomizing the instances from the organized sequence.

Apparatus. The learning tasks were presented by a Digital Equipment Corporation PDP/8 680 Communication System which is interfaced to an IBM 1500 Instructional System. This system supports 16 teletypes, of which a maximum of ten were used during any one session of this study. The terminals were located in an air-conditioned sound-deadened room. The computer-assisted instruction system administered the learning task and recorded the students' responses and latencies.

Procedure. An experimental session consisted of general directions read by the experimenter, a pretest taken on the teletype, a presentation of the definition in a printed booklet, followed by the treatment, and a posttest all taken on the teletype except for the definition. Subjects, ten at a time, were seated in the experimental room in front of a teletype. General directions were read by the experimenter, who then turned on the terminal and entered the subjects' number. Directions on the operation of the teletype and the program were given by the computer and in the booklet. After these brief directions, the subjects were given the pretest on the terminal. Subjects were required to identify examples of trochaic poetry by typing "Yes," and "No," if a nonexample. Following the pretest, subjects were asked if they had ever studied trochaic meter. Subjects were then instructed to read the definition of trochaic meter contained in the booklet. The booklet allowed the subjects to keep the definition for reference throughout the task. When the subjects had studied the definition, they proceeded to the task. Subjects in the prompted groups were given the poetry selections followed by the prompting. In the nonprompted conditions, the selections were labeled as examples or nonexamples.

At the conclusion of the task subjects raised their hand to indicate they had finished the program and were ready for the posttest. The program was nonspeeded so the subject could study at any point in the task. When the subject was ready for the test, the experimenter removed the task paper from the teletype, collected the definition, and entered the appropriate command to start the posttest. The posttest was designed using the same format as used in the previous study by Tennyson (1973), except there was no misconception error. At the conclusion of the test the subject's score was given by the computer and the subject was allowed to leave.

Results

This study used the same scoring procedure as the previous experiments (Tennyson et al., 1972; Tennyson, 1973) in obtaining the error mean dependent variables of correct classification, overgeneralization, and undergeneralization (Table 1). Each of the four treatments was hypothesized to result in a particular learning behavior. Given additional instructional information, subjects would show a tendency to be conservative in identifying instances as positive; therefore, the groups without prompting (Organized/No Prompts and Random/No Prompts) would overgeneralize. The sequence effect would affect the degree of classification behavior, in that the random groups (Random/Prompts and Random/No Prompts) would have a tendency to overgeneralize. Latency data, as a dependent variable, was collected on the pretest, the learning task, and the posttest. Because of the interdependence of the dependent variables, a multivariate analysis of covariance was used as the statistical design. The two covariates used in the analysis were sex and prior knowledge of trochaic meter.

TABLE 1

Experiment VII
Hypothesized Error Responses and Mean Error Scores

Behavioral Outcomes	Treatment Groups			
	OP	ONP	RP	RNP
Correct Classification	5.4[a]	7.1	6.8	9.3
	0[b]	7	6	7
Overgeneralization	8.3	7.0	8.4	5.3
	6	3	13	0
Undergeneralization	8.3	9.2	7.2	11.4
	6	13	0	13

Note.--The treatment groups are represented by capital letters: OP = Organized/Prompted; ONP = Organized/Not Prompted; RP = Random/Prompted; and RNP = Random/Not Prompted.

[a]First rows are the adjusted mean scores.

[b]Second rows are the predicted error scores.

The first series of multivariate hypothesis tests blocked on the two main effects, sequence and prompting. The sequence variable showed a significant difference between the organized and random groups ($U > .78$, $df = 9/1/81$, $p < .05$). The second independent variable, prompting, resulted in a significant difference ($U > .80$, $p < .01$) between the two conditions. To determine where the differences occurred, a series of univariate hypotheses on each of the dependent variables was performed. The data analyses are reported in two sections; the learning task error scores and the latency measures. Interactions appropriate to the design were tested, but none were significant ($p > .05$).

Learning task. The first univariate test on the pretest, consisting of 16 items given to all subjects, showed no significant difference between the four groups ($p > .05$). The pretest error means indicated minimal prior knowledge of the trochaic meter concept used in the task. Using the correct classification scoring scheme, the four groups did perform significantly different ($F = 2.78$, $df = 3/81$, $p < .05$). A Duncan's new multiple range test was used to determine differences among the groups. The organized/prompted (OP) group had an error mean score significantly different from the Organized/Not Prompted (ONP) group

($p < .05$), and the Random/Not Prompted (RNP) group ($p < .01$); there was no difference with the Random/Prompted (RP) group ($p > .05$) (Table 1). The other statistically significant comparison was between the two random sequence groups, with RP having the lower error mean score ($p < .05$). There was a difference between the two not prompted groups at the .07 level.

The univariate analysis on the overgeneralization dependent variable resulted in a significant F test ($F = 3.09$, $p < .05$). Duncan's test showed the RNP group as having the lowest significant error score ($p < .05$), except for the no difference with the ONP group ($p > .05$) (Table 1). There were no other significant differences between the groups ($p > .05$). On the undergeneralization univariate test ($F = 2.92$, $p < .05$), the RP group using Duncan's test, had a significantly lower score than the RNP group ($p < .05$). There were no other differences ($p > .05$).

Latencies. Three latency times were collected to determine instructional efficiency of the four treatments. The pretest latency univariate test was nonsignificant ($p > .05$), that is, all four groups took approximately four minutes to finish the pretest (Table 2). Task latency refers to the total time spent in the learning program. The F test for task latency resulted in a significant difference ($F = 2.94$, $p < .05$) between the four groups. Duncan's test showed that the OP group spent less time on the task than groups RNP and ONP. The RP

TABLE 2

Experiment VII
Adjusted Mean Latencies

Latencies	Treatment Groups			
	OP	ONP	RP	RNP
Pretest	4.1	4.2	4.2	5.1
Task	8.3	10.3	9.7	10.8
Posttest	9.1	13.1	11.1	13.1

group differed from the OP group at the .03 level. However, on the posttest latency, the OP group was significantly different ($p < .05$) from the RP group ($p < .05$), and the RNP and ONP groups. The RP group's time was significantly lower than the RNP and ONP group's ($p < .05$).

Discussion

The systems approach to instruction proposes that learner acquisition of knowledge is improved, as well as requiring less time, than traditional forms of teaching. The purpose of this study was to investigate this premise while testing the variables of prompting and task sequencing. Prompting of the examples and nonexamples was done to focus the learners' attention on the presence or absence of the critical attributes of the given concept. Such prompting increased the amount of reading material given the subject during the instructional portion of the task. That is, the subjects in the no prompting conditions received only the definition of trochaic meter and the instances. Thus, the required reading length was more than doubled for the prompted groups.

The results demonstrated that the prompting condition, according to the given operational definition, reduced subject time spent on the learning task. The sequence effect was not, however, a factor in subject latency on task. Subjects in the two prompting conditions seemed to read the given material per instance and continue through the program at a steady pace, while subjects in the no prompting condition spent more time per instance. Without the prompts the subjects were forced to apply the rule of trochaic meter to determine why a given instance was labeled positive or negative.

The posttest measures of performance and latency showed that the optimal treatment (organized/prompted) resulted in increased effectiveness and efficiency. Subjects in the prompted groups not only had fewer errors on the correct classification score, but finished the test in significantly less time than the no prompted. This would indicate that the level of acquisition was also better because performance required significantly less time. The combined treatment of organized/prompted demonstrated this assumption when the subjects' latencies were less than the random/prompted, even though their performance scores were the same. The effectiveness of the optimal treatment is shown in the overgeneralization and undergeneralization scores which demonstrated that the subjects were not making these errors as the other subjects in the groups.

References

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

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 non-exemplar variables which produce correct concept classification behavior and specified classification errors. Journal of Educational Psychology, 1972, 63, 144-152.

THE EFFECTS OF PRIOR MEMORIZATION OF DEFINITION COMPONENTS ON CONCEPT ACQUISITION USING AN EFFECTIVE TRAINING PARADIGM

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

Abstract

The effect on concept acquisition of requiring memorization of either examples or nonexamples prior to going through a theoretically effective training program was compared to the performance of groups who either memorized nothing or memorized key words in the concept definitions. Correct classification scores, undergeneralization and overgeneralization 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 spent less time to reach criterion in the training program, but took significantly more total instructional time than did the no-prior-memorization group.

THE EFFECTS OF PRIOR MEMORIZATION OF DEFINITION COMPONENTS ON CONCEPT ACQUISITION USING AN EFFECTIVE TRAINING PARADIGM

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

Learning a defined concept ultimately entails the capability to correctly classify instances as examples or nonexamples according to a definition. Gagné (1970) classified learning from defined concepts as a specific type of rule using. In his theory, the definition serves as a rule statement which is used by the learner while acquiring classifications skills. When a student is able to correctly classify unfamiliar instances according to the definition, his behavior is said to have become rule-governed. The capability to classify instances is usually preferred to the capability of recalling either subsets of examples and nonexamples or the definition itself. The objective of this study was to assess the effects of memorization of examples or nonexamples or subconcepts in the definition on classification behavior when the memorization occurs prior to a training program whose goal was to teach correct classification skills.

Much of a student's day is devoted to learning defined concepts (Gagné, 1970; Carroll, 1964). The applicability of most concept identification and formation research to this important type of concept learning has yet to be demonstrated. Instructional science research, however, 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) and Tennyson (1973) 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 effectiveness was not tested. Instead, further instructional design modifications were introduced and evaluated.

Correct rule using behavior is not guaranteed by the memorization of rule statements alone (Gagné, 1970). Likewise, correct concept classification behavior is not guaranteed by the memorization of the concept definition or of subsets of examples and nonexamples. However, this does not assume that prior memorization of different definition components could not influence the way in which learners classify unfamiliar instances (rule-governed behavior) or in the abstraction of the definition (the rule) that the learner has internalized. Mandler and Pearlstone (1966) contrasted the performances of two groups whose task was to impose a conceptual rule on a standard stimulus array. A free group was asked to derive a categorization rule as it met instances; a constrained group was given an experimenter-established rule before meeting any instances. The findings were:

> "Given identical conceptual categories, the free Ss attain criterion much faster, and in nearly half the cases they do so on the first trial; i.e., these subjects are imposing a conceptual rule on the stimulus array before they have inspected all the instances of the array." (p. 130)

Mandler and Pearlstone hypothesized that the constrained subjects were also formulating individual hypotheses and that the subjects used these as they met successive instances. However, subjects had to discard most of these hypothesized rules as more instances became available.

Results of the Mandler and Pearlstone (1966) study suggest that the type of information which is presented and processed affects the formation of the conceptual rule and subsequent classification behaviors. Thus, the prior memorization of different definition components could be expected to influence the internal organization of the rule, and this difference in organization could be reflected in both in-task and posttask indices of classification behavior. Such a hypothesis was tested in this study. The independent variables were the type and amount of information required of students to memorize prior to classification training and testing.

Method

Subjects

A total of 93 subjects participated in this study. Data from two subjects were discarded because one subject had to leave before finishing and a second subject responded indiscriminately on the posttest. Of the 91 remaining, data from 17 subjects was not analyzed because these subjects failed to reach criterion on one of the concepts. This left 74 subjects from whom complete data was collected. Of these 74 subjects, 33 were seventh graders, 41 were eighth graders. Twenty-eight subjects were males, 46 were females.

Learning Task

The experiment consisted of three main phases: (a) memorization of definition 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 students. The instructional objective of the training phase was: Given an unfamiliar instance, the subject will identify it as 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 generalizability of results found. Four, a standard dictionary format was followed for both concepts. Both definitions describe first the general class to which the concept belongs, and then go on to describe how the defined instances differed from other members in the general class. This is known as 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. Five, 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, 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 (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 attributes, 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.

The following is a sample of the six examples and nonexamples used for each of the concepts:

Skeethand (conjunctive):

Examples:	#1	2	3	4	5	9
	#2	2	4	5	6	9
	#3	2	5	7	8	9
Nonexamples:	#4	2	3	3	5	9
	#5	2	3	5	9	10
	#6	2	3	4	5	6

Derf (disjunctive)

Examples:	t	aeiou	quq
Nonexamples:	at	maeiou	qum

It was also explained to the subjects why each of these six instances was classified as an example or nonexample. An attempt was made to reference these explanations as much as possible to the critical attributes in the concept definition. The following explanations are referenced to the Skeathand teaching display above and to the Skeathand definition:

Hand #1 meets all the requirements.

Hand #4 does not meet requirement #1.
(Note the two 3s.)

Hand #2 meets all the requirements.

Hand #5 does not meet requirement #2.
(Note the 10.)

Hand #3 meets all the requirements.

Hand #6 does not meet requirement #3.
(Note the lack of a 9.)

The next sample of explanations are referenced to the Derf teaching display above and to the Derf definition:

The t is an example because it contains no vowels.
The at is a nonexample because it does not meet any requirement.
The aeiou is an example because it contains no consonants.
The maeiou is a nonexample because it does not meet any requirement.
The quq is an example because it contains two q's.
The qum is a nonexample because it does not meet any requirement.

It is important to note that the definition and the six instances always appeared together on the cathode ray terminal (CRT) screen and that the explanations were presented separately on the screen while the definition and instances were visible. 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. Two groups memorized either one or three examples (EX) or one or three nonexamples (NEX). Two control groups were used. Members in one control group were required to memorize selected subconcepts of the definition (DEF), and those in the other memorized nothing and were passed directly to the training phase (NULL). 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 analyses, 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 again shown the example(s) or nonexample(s) and asked to type them again from memory.

The initial randomization procedure determined for which concept subjects memorized three instances and for which concept students memorized only one instance. Consequently, subjects were not only randomized to the EX or the NEX groups, but were also randomized to one of two subgroups within these groups (A or B). For example, the EX-A group memorized one example of the Skeethand concept and three examples of the Derf concept. 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 (major subconcepts) in each definition. Words memorized in the Skeethand concept were: five, no, once, all, 10, 25, and 9. For the Derf concept, the words memorized were: letters, either, no, more, and twice. The task for this group was given an incomplete definition to type in the missing key words correctly. Each time the subject did not respond correctly, he was again shown the complete concept definition and asked again to type the missing words from memory into an incomplete definition.

Procedure

A random number table was used to randomly assign subjects to one of the four experimental conditions and to a CRT booth. They were instructed that the program was individualized and that they should proceed at their own pace until completed. Instructions on how to operate the terminals were given on the CRT. Samples of familiar definitions were displayed, and it was explained how these definitions could be used to divide instances into examples and nonexample groups. Subjects were then told that 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 concept.

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 definition components with 100% accuracy.

The same teaching display presented in the memorization phase was presented as the first teaching display in the training phase. After studying the display and the six explanations, the screen was cleared and subjects were tested on four unfamiliar 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 their fourth attempt through the training phase for either concept, they were dropped from the study and their data were not analyzed.

After successfully passing through the memorization and training phases for both concepts, subjects in the EX, NEX, and DEF groups were shown the definition 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 unfamiliar 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 unfamiliar test instances were developed to assess these skills.

In-task items. The selection of each set of four test instances in the training phase was based on the algorithm that if an instance is an example, then include only the minimal number of critical attributes necessary; if it is a nonexample, include one less than the minimal number of critical attributes necessary. In each set of four, an attempt was made to include the fullest range of irrelevant attributes possible.

Posttest items. Twenty Skeethand and 20 Derf unencountered instances made up the posttest. As in the training phase, the task was to correctly classify instances as examples or nonexamples. For the conjunctive concept (Skeethand) there were 8 examples and 12 nonexamples, and for the disjunctive concept (Derf) the 20 instances were made up of 12 examples and 8 nonexamples. Again, a full range of irrelevant attributes was present.

Results

Subjects Failing to Meet Criterion

In an attempt to find systematic reasons why the particular seventeen subjects did not meet criterion in the training phase, multiple linear regression techniques were used as outlined by Cronbach and Snow (1969) and by Bottenburg and Ward (1963). 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 for the testing of main and interaction effects employed here is a modification of the Bottenburg and Ward approach (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 a reduced model by dropping its vector from the full model and then testing for the reduction in the multiple correlation. The significance of interaction effects was tested by alternately adding the CTMM X Group vector and the Age X Group vector to the full model described above and then testing for the increase in the multiple correlation. The results of these analyses appear in Table 1.

TABLE 1

Summary of Multiple Linear Regression Analyses with Pass-Fail Criterion

Effect	df	% of Variance	F	P
Program	3,85	04	1.39	
CTMM	1,85	20	23.75	<.000
Age	1,85	00	<1	
CTMM X Group	1,82	03	1.14	
Age X Group	1,82	04	1.64	

Because the scores of those students 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 students are not readily generalizable to the experimental population as a whole. On the other hand, the fact that neither the Group nor the Group X CTMM or the Group X Age predictor variables accounted for significant portions of the variance supports the contention that group comparison tests using the curtailed data base are interpretable despite the selection bias.

Variables

In this investigation, the critical dependent variables were: (a) correct classification scores, (b) undergeneralization error scores, and (c) overgeneralization error scores. Overgeneralization scores refer here to the number of nonexamples erroneously classed as examples, while undergeneralization scores refer to the number of examples classed as nonexamples. In addition, in-task and posttest latencies were collected to assess instructional efficiency. An analysis of covariance statistical model was chosen to test the implied null hypothesis. No significant prior experimental differences were found for the covariates of CT/M scores or age.

Learning success. On the posttest, mean correct classification scores and the two types of mean error scores were analyzed for each concept separately and then in combination, resulting in nine separate *F* tests. Each of these tests resulted in *F*s less than unity (see Table 2).

TABLE 2

Posttest Classification Score Means

	Concepts								
	Conjunctive			Disjunctive			Combined		
Group	C[a]	O	U	C	O	U	C	O	U
	20	12	8	20	8	12	40	20	20
EX	18.6	.7	.7	15.5	3.2	1.2	34.2	3.9	1.9
NEX	18.6	.8	.6	15.1	3.4	1.5	33.7	4.2	2.2
DEF	19.2	.4	.5	14.9	3.3	1.8	34.1	3.6	2.3
NULL	18.9	.1	.9	15.5	3.1	1.4	34.5	3.2	2.3

[a]These capital letters represent the following behaviors: C = correct classification, O = overgeneralization, and U = undergeneralization. Numbers beneath the letters refer to the total score possible for each variable.

Overall correct classification averages on the posttest were 18.8 (94%) for the conjunctive concept and 15.3 (76%) for the disjunctive concept.

Two other indices of learner success within the training program were analyzed. The mean percentage of subjects failing at least one test item in the conjunctive concept training program was 35% and 45% in the disjunctive concept training program. Of those who did fail at least one item in the training program, the mean test item number where the last classification error was made was 4.4 for the conjunctive concept and 6.8 for the disjunctive concept. No significant between-group differences existed for either variable.

For the combined posttest scores, the CTIM covariate resulted in Fs of $F = 21.13$ ($p < .001$), $F = 25.67$ ($p < .001$), and $F = 2.00$ ($p > .05$) for correct classification, overgeneralization, and undergeneralization scores, respectively. CTIM scores correlated positively with correct classification scores and negatively with overgeneralization error scores. The age covariate did not account for significant portions of the variance in any of these analyses.

Latencies. Memorization phase latency was collected, but analyses on group means were not made. Three latency comparison tests were made: (a) training phase latency, (b) posttest phase latency, and (c) total program latency. Analysis of training phase latency group means resulted in a significant F test ($F = 3.37$; $df = 3/64$; $p < .05$).

TABLE 3

Latency Means for Training and Testing Phases, and Total Program

	Latencies			
Group	Memorization Phase	Training Phase	Posttest Phase	Total Program
EX	(5.6)[a]	(8.6)9.1[b]	(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]Latency times are in minutes.

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

A Newman-Keuls test was used to make pairwise group comparisons using the adjusted group means. The only significant comparision found revealed that the NULL group took significantly longer in the training program on the average than the DEF group ($p < .05$). It is interesting that the more time each group took in the memorization phase, the less time they took to reach criterion in the training program phase.

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 $F = 7.44$ ($p < .05$). A Newman-Keuls test on the adjusted group means revealed that the NULL group took significantly less time to complete the total program than did either of the three prior memorization groups, $p < .01$ for each of the three pairwise tests. There were no differences among the three prior memorization group means on the total program latency variable.

For the CTMM covariate, Fs for inclusion within the full model for training phase, posttest phase, and total program latency scores were $F = 28.39$ ($p < .001$), $F < 1$, and $F = 24.93$ ($p < .001$), respectively. CTMM scores correlated negatively with training phase and total program latency scores. For all analysis of covariance tests on latency means, the age covariate did not account for significant portions of the variance.

Correlations. The intercorrelations between CTMM scores, age, Posttest classification scores, and latency scores were calculated (see Table 4).

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. 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 different from that created by the other three treatment programs. Second, these results support the null hypothesis that no difference exists in the effectiveness of the four training programs. This latter conclusion can be further substantiated when another result is considered. That is, when a prediction equation was formed predicting whether a student would or would not reach criterion in the training phase, dropping the group vector from the equation resulted in no significant loss in prediction.

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. That is, on the average, the three prior memorization groups spent 6.4 minutes in the memorization

TABLE 4

Intercorrelation Matrix

	CTMM	Age	Posttest[a] Disjunctive Concept Items	Posttest[a] Conjunctive Concept Items	Posttest[a] Total Items	Latency Training Phase	Latency Posttest Phase	Latency Total Program
CTMM		.41	.48	.34	.56	-.47	-.05	-.40
Age			.21	.24	.28	-.03	.03	-.19
Disjunctive Concept Items				.09	.89	-.44	-.10	-.32
Conjunctive Concept Items					.53	-.23	.17	-.13
Total Items						-.48	.01	-.32
Training Phase							.17	.18
Posttest Phase								.12
Total Program								

[a] Refers to correct classification scores.

phase and 6.2 minutes (adjusted) to reach criterion in the training phase, while 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.

Requiring learners to memorize a larger number of examples or nonexamples or to memorize more of the definition than was required in this experiment may increment the learners' capability to classify unfamiliar instances. However, any increment in posttest performance would probably not justify in most educational situations the large amount of time learners would have to spend memorizing definition 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 definition components is not an advisable instructional technique.

The consistently high positive correlations between CTMM scores and the posttest indices and the high negative correlations with training program and total program latency measures could have been expected. The task in this experiment was largely one of internalizing a rule from verbal statements and instances and then demonstrating correct rule-governed behaviors. As already stated, the learning task was constructed so as to be unfamiliar to the subjects while allowing easy identification of the critical attributes within the instances. The capabilities needed for the subjects to succeed in this experiment were similar to capabilities needed to do well in aptitude tests like the CTMM.

To the extent that the learning requirements in an experimental task environment model the learning requirements in actual task environments, generalization of experimental results to the actual task environments is insured. In this study special care was taken to insure that the simulated experimental environment placed the same types of requirements on learners as did concept acquisition tasks germane to the classroom. Subjects' success as measured by in-task and posttask indices suggest that an effective concept teaching paradigm is available for implementation. In cases where the critical and/or irrelevant attributes are less familiar to the learner or less easily identifiable in instances, the paradigm introduced here is as theoretically powerful. Research in testing the worthiness of the paradigm in these situations with both concrete and defined concepts would appear to be potentially rewarding.

The correlation between the correct classification concept was surprisingly low ($r = .09$), especially considering how much variance these scores share with CTMM scores. This low correlation suggests that the learning requirements for classification tasks with conjunctive concepts are functionally different than those with disjunctive concepts.

Eighty percent of all students who reached criterion in the training phase classified 80% or more of the 40 unfamiliar instances on the posttest correctly. However, a few scores were at the chance level on the posttest, and 19% of all subjects failed to meet criterion in the training phase. This latter group could not correctly classify 4 instances after seeing the definition on 5 separate occasions and after seeing a total of 40 instances correctly classified for them. The correlation between success in the training phase and success on the posttest indices suggest that the wide differences in student achievement levels is not simply an experimental artifact.

Possibly some of the students who had problems could have benefited most from a completely different teaching strategy. On the other hand, possibly learning problems could have been identified within the program and addressed most profitably in a remedial sequence of instruction. Verbal interaction with some of the students 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. Student difficulties with both these things help explain why the disjunctive concept task was so difficult. It would be expected that both difficulties could have been addressed most efficiently in remedial instructional sequences. Decisions about the structure and function of such a remedial scheme and about whether it should be learner-controlled or program-controlled would have to be made.

In conclusion, it is questionable whether all students of junior high school age are able to correctly classify nonfamiliar instances in accordance with standard dictionary definitions without training aimed at this objective. Including the memorization of definition components prior to a training program has questionable merits for instructional efficiency reasons. The objective of the training paradigm used in the training phase was to produce correct classification behaviors. While being effective for the majority of students, some type of adaptation to learner characteristics or to specific learning problems would probably be worthwhile.

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).

Carroll, J. B. Words, meanings, and concepts. Harvard Educational Review, 1964, 34, 178-202.

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.

Mandler, G., & Pearlstone, Z. Free and constrained concept learning and subsequent recall. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 126-131.

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 Lehrsysteme '72 Symposion uber Programmierte Instruktion und Lehrmaschinen in Verbingung mit einer Fachausstellung, Berline, Germany, April 5, 1972.

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 non-exemplar variables which produce correct concept classification behavior and specified classification errors. Journal of Educational Psychology, 1972, 63, 144-152.