R842-09 ERIC REPORT RESUME

ED 010 199 1-31-67 56 (REV)
PROMPTING AS AN INSTRUCTIONAL VARIABLE IN CLASSROOM SIMULATION, FINAL REPORT.
TWELKER, PAUL A.
RQR60180 OREGON STATE SYS. OF HIGHER ED., MONMOUTH
NDEA-VIIA-1366
BR-5-0950
R-52
30-APR-66 OEG-7-47-9015-276
EDRS PRICE MF-$0.18 HC-$2.96 74P.

*PROMPTING, *STUDENT REACTION, *LEARNING THEORIES, *REINFORCEMENT, *TRANSFER OF TRAINING, STIMULUS DEVICES, MONMOUTH, OREGON

STUDENT PROMPTING WAS INVESTIGATED AS A MEANS OF ENHANCING LEARNING EFFICACY. IT WAS HYPOTHESIZED THAT BOTH (1) LEARNERS WHO ARE GIVEN GENERAL INFORMATION WHICH IDENTIFIED THE TYPE OF PROBLEM TO WHICH A RESPONSE IS REQUIRED (CLASSIFICATION PROMPTS) AND (2) LEARNERS WHO ARE GIVEN INFORMATION THAT HELPS IDENTIFY APPROPRIATE RESPONSES (STANDARD PROMPTS), WOULD TAKE FEWER TRAILS TO LEARN AND SCORE HIGHER ON A TRANSFER TEST THAN LEARNERS WHO ARE NOT GIVEN THESE PROMPTS. MATCHED GROUPS OF JUNIOR-LEVEL COLLEGE STUDENTS WERE USED TO STUDY THE TWO TYPES OF PROMPTS, AND A PRETEST WAS USED AS THE CONTROL VARIABLE. THE POST-TEST WAS ADMINISTERED AS A TEST OF TRANSFER FROM PREVIOUS INSTRUCTIONAL SEQUENCES, AND VARIOUS MEASURES OF LEARNING RATE, TRANSFER, AND AFFECTIVITY WERE OBTAINED. THE HYPOTHESES OF THE STUDY WAS NOT SUPPORTED BY THE RESULTS. THE PRESENTATION OF CLASSIFICATION PROMPTS HAD NO MEASURABLE EFFECT ON LEARNING OR TRANSFER. THE STANDARD PROMPTS MADE LEARNING MORE EFFICIENT IN NUMBER OF SESSIONS REQUIRED FOR LEARNING AND IN AIDING THE ADEQUACY OF THE SUBJECT@S FIRST RESPONSE IN TRAINING ON EACH PROBLEM ADMINISTERED. THE EFFECTS OF THESE LATTER PROMPTS, HOWEVER, WERE NOT CONSISTENT IN AFFECTING ABILITY TO RESPOND TO TRANSFER TASKS. (JH)

PROMPTING AS AN INSTRUCTIONAL VARIABLE IN CLASSROOM SIMULATION

Paul A. Twelker

Final Report

Title VII Project Number 5-0950
B-1366
National Defense Education Act of 1958

Grant Number 7-47-9015-276

teaching research

OREGON STATE SYSTEM OF HIGHER EDUCATIO

R-52

Prompting as an Instructional Variable in Classroom Simulation

Paul A. Twelker

Final Report
Title VII Project Number 5-0950
National Defense Education Act of 1958

April 30, 1966

Teaching Research Division
Oregon State System of Higher Education
Monmouth, Oregon

Acknowledgments

The author is indebted to many people and wishes to express his gratitude for their contributions:

Dr. Bert Y. Kersh, Teaching Research Division, for his encouragement and support throughout the project,

Dr. Jesse H. Garrison, Oregon College of Education, and Mr. Robert Myers, University of Oregon, who gave full cooperation as instructors of the courses from which experimental subjects were used,

Dr. Gary Martin and Dr. Deon Stevens, Teaching Research Division, who were the experimental instructors for the pilot study,

Mr. John Jones, Mr. Thomas Stipek, Mr. Peter Smith, and Mr. Gerald Girod, who were the instructors for the experiment proper,

Mr. Larry Janssen and Mr. David Neiss, Oregon State University Computer Services, who conducted the automated data analyses,

Mr. Fred C. Crowell and Mrs. Virginia Weigel, Teaching Research Division, who assisted in the statistical computations,

Mr. Timothy J. Pettibone, Teaching Research Division, who served as laboratory technician for the project,

The U. S. Office of Education, who gave financial support for the project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I OVERVIEW[1]

Classroom simulation, as developed by Kersh (1961; 1963a) is an instructional experience which allows student teachers to practice responding to classroom problems under supervision. Sound motion pictures, multiple projection techniques and printed materials are employed to simulate a single sixth grade class in a variety of problematic situations in a laboratory training facility.

As used currently, the simulation technique requires a student teacher (T) to respond to each problem sequence without explicit guidance. T is not prompted by the experimenter (E) with any principles of teaching or specific hints on how to behave, to guide his overt responding. In effect, a "learning by discovery" method is relied upon to help T develop modes of operation within certain behavioral limitations, and to teach T how to identify relevant cues that indicate problems in the classroom. Because the length of time required for instruction is long, it is difficult to provide individualized instruction for more than 30 student teachers per term.

A problem of instruction is to adequately provide forms of prompting which shortcut or do away with the need for trial and error learning, and hence increase learning efficiency without reducing transfer. Withholding principles in order that learners may "discover" for themselves, actually may reduce performance, increase the time required for learning, and decrease affectivity toward the learning experience (Wittrock, 1963; Twelker, 1964b). Further, attempts to introduce princi-

---

[1] Revision of a paper originally read at the American Educational Research Association, Chicago, February, 1966.

ples through feedback after a time of independent discovery by the learners, as compared with giving principles directly to learners, devoid of discovery activities, may result in decreased motivation to learn new subject matter (Twelker, 1964b).

In simulation training, two types of prompts may be appropriate: (1) problem classification prompts that help T identify the problem, and (2) principle or standard prompts that help T identify appropriate responses from a number of alternative responses. In other words, these prompts might cue T to the essential features of the situation which represent a problem to which he must respond, as well as guide T's overt response behavior. The main purpose of the project was to investigate empirically the effects of these two types of prompts on learning and transfer when college students are taught how to respond to problems of management and communication in a classroom simulation facility.

The prompts used were general rather than specific. Studies that have varied the specificity of prompts have generally found that presenting rules to learners enhances learning, retention, and transfer of learned principles to a greater extent than presenting specific answers to problems (Kittell, 1957; Forgus and Schwartz, 1957; Craig, 1956; Wittrock, 1963; Wittrock and Twelker, 1964a; 1964b; Twelker, 1964a; 1964b). Specific prompts might aid the student in making correct responses during learning, but might be inappropriate in a testing situation that requires a different answer.

Objectives

From the reasoning mentioned above, the following hypotheses were tested:

(1) Learners who are given general information that helps identify what type of problem is to be responded to (classification prompts), will take fewer trials to learn,

and will score higher on a transfer test than learners who are not given these prompts.

(2) Learners who are given information that helps identify appropriate responses (standard prompts), will take fewer trials to learn, and will score higher on a test of transfer than learners who are not given these prompts.

Procedure

Experimental Design. A two-factor design with matched groups (A x B x L design, Lindquist, 1953) was used to study the two types of prompts. A pretest was used as the control variable. Subjects were blocked into six levels, and randomly assigned to the various A-B combinations within each level. The four treatment groups were:

(1) Standard given; classification given

(2) Standard not given; classification given

(3) Standard given; classification not given

(4) Standard not given; classification not given.

A control group was given the pre and post test without simulation training to assess extraneous sources of influence that may have caused changes during simulation training, i.e., ongoing course of instruction.

Sample Plan. Eighty-nine junior level college students were given simulation training with existing materials developed by Kersh (1963b). The subjects were all under 25 years of age, and had no previous teaching experience.

Materials and Procedure. Simulation training involved four phases: (1) orientation; (2) pretesting; (3) instruction; and (4) post testing. The orientation phase introduced "Mr. Land's Sixth Grade" to Ts. Among other things, Ts were responsible for learning the names of the children and the important characteristics of each child in terms of class role, academic ability, and special problem areas.

The pretest was administered to Ts in small groups. Ts were requested to write their responses to each of 20 filmed problem sequences. Inter-rater reliability between four instructors was shown to be .94.

Simulation training employed 20 problem sequences. In brief, the experimenter explained the setting of the problem sequence to T. The filmed problem sequence was then presented and T was requested to enact his response to the problem. E compared T's response with pre-established standards and on this basis, selected and projected feedback sequences. Each sequence was repeated until T responded appropriately. After each presentation of a problem and feedback sequence, E and T discussed the experience. E withheld direct guidance and forced T to evaluate his performance on the basis of past knowledge, the feedback sequence, supporting records presented during the orientation phase of training, and prompts given, if any. At this time, E rated T's assessment of the problem. The simulated problem was repeated until T could verbalize all the cues that identified the problem. Reliability between the four instructors for the rating of the responses and the assessment of the problems was shown to be .97 and .95, respectively.

Prompts, as called for by the experimental treatment, were given to T after the description of the problem setting and before T enacted his response. Examples of a standard prompt and a classification prompt follow:

> Standard prompt - When learners appear disinterested or confused, stimulate a more active, interested response rather than make no effort to change the learner's response.
>
> Classification prompt - This is a management problem that involves a child's disregard of instructions from a teacher due to a fatigue reaction.

The post test presented 20 new filmed problem sequences and required T to respond to each problem and to identify the problem. The post test was sufficiently different from the instructional sequence that it constituted a test of transfer.

Various measures of learning rate, transfer, and affectivity were obtained.

Summary of Results

To assess whether the experimental treatment groups differed reliably on the post test measures from the control group which did not receive training, Dunnett's procedure was used (cf, Winer, 1962, pp. 89-92). This comparison indicated that the training procedure produced gains in post test performance that could not be attributed easily to chance. These findings indicate that further analysis of the data on the various measures listed below was appropriate and meaningful.

Number of Sessions in Training. The groups that were given the standards took about 10% less time to complete instruction than those groups that were not given the standards ($p < .05$). A statistically significant standard x classification interaction ($p < .05$) revealed that instruction was faster when standards were given without classifications than when standards were given with classifications. Clearly, the worst situation in terms of the number of sessions required for learning, was to give no prompts at all. The "no prompts" groups took about 20% longer than did the fastest group.

Total Repetitions in Training. This measure represents the maximum number of repetitions or trials required by T to meet criterion on the response or assessment measure. Findings indicate that the groups that were given standards required about 16% fewer repetitions than did those groups not receiving standards ($p < .001$). The presenting of classification prompts did not materially reduce the number of repetitions of the problem sequences required in learning.

Adequacy of First Response in Training. This measure represents the adequacy of T's first enacted response to each of the 20 problem sequences in training. The results indicate that when standards were given, Ts scored about 20% higher on thier first response in training than when standards were not given ($p < .001$). This finding indicates that the standard prompts given were meaningful and actually guided students to make appropriate responses when presented initially with the problem.

Adequacy of First Assessment in Training. This measure represents the adequacy of T's first assessment of each of the 20 problem sequences in training. The results indicate that the presentation of neither prompt significantly affected this measure.

Post Test Response. This measure represents the adequacy of T's responses to each of the 20 novel problem sequences. A statistically significant standard x classification x level interaction was obtained ($p < .01$). Because of the complexity of the design, in terms of the number of levels, the interpretation of this interaction is difficult. However, it was evident that the three highest mean scores were obtained by the groups that were given standards, while the three lowest mean scores were obtained by groups not receiving standards. In every case, the high scoring groups represented the upper three (of the six) levels.

Post Test Assessment. This measure represents the adequacy of T's assessment of the problems in each of the 20 novel sequences. Results indicate that neither of the prompts produced statistically significant main effects or interaction effects.

Affectivity. This measure represented T's rating of the entire simulation experience, as measured by a Thurstone-type attitude scale. No statistically significant differences were found. However, attitudes generally were favorable.

Discussion

The results indicate that giving standard prompts that guided Ts' subsequent responses, made learning more efficient in terms of number of sessions required for learning, number of trials required to meet criterion, and adequacy of T's first response in training on each problem, as compared with not giving the same prompts. Also, the standard prompts apparently had some effect on the student's ability to respond to novel problems in the transfer task. However, this effect was not consistent across treatment combinations, and was limited to groups representing above-median performance on the pretest. The presentation of classification prompts that helped Ts identify the problems they should respond to in the simulated problems had no measurable effect on learning or transfer. Thus, the first hypothesis was not supported, and the second hypothesis was only partially supported.

Two implications arise from this study. First, evidence indicates that future classroom simulation techniques need not require students to respond to simulated problems as though they had no principles of behavior upon which to base their actions except that which they bring in with them to the simulation laboratory. The findings reveal that the presentation of standards of behavior had a positive effect on learning rate, and for some students, transfer. These findings support previous research on the efficacy of using meaningful prompts during instruction.

Although the evidence is far from conclusive, the possibility is raised that for some students, transfer performance may be reduced with the presentation of too much information. The evidence failed to support the superiority of one type of prompt over the other.

Second, simulation training may provide a powerful vehicle for teaching principles of instruction or principles of classroom management and control, because it provides common referents for students. That is, students are provided with realistic experiences that are examples of situations in which the use of these principles result in consequences on the part of the simulated class that are desirable. Students can consider the principles in light of these experiences, and weigh the consequences of their actions, shown through selected feedback sequences, against their own predictions. These findings confirm those of Vlcek (1966) and point the way to an eventual adoption of more efficient methods of teaching instructional principles than is available through current teaching techniques. It should be realized, however, that although the post test is a test of transfer, the effects of presenting prompts eventually should be assessed by direct observation in an actual classroom situation.

II RELATED RESEARCH

The Development of Classroom Simulation

In Oregon, simulation as an instructional medium in teacher education has been under development since 1961. In that year, Bert Kersh, armed with NDEA Title VII support, built a simulation facility and initiated the development of a variety of simulated classroom situations. These situations were simulated through the medium of sound motion pictures and printed materials.

The classroom simulation technique under consideration attempts to create for the student teacher all of the relevant features of a single classroom situation. Potentially, many different classrooms may be created, but presently the technique is limited to one group of sixth graders named, "Mr. Land's Sixth Grade." Mr. Land is the hypothetical supervising teacher with whom the student teachers work during their simulation training. A complete cumulative record file is available on each child, including standardized test data, achievement records, health records, a summary of the teacher's anecdotal records, and a snapshot. In addition, there are printed descriptions of the hypothetical school, "College Grove Elementary," and the community of "College Grove." The student teacher may be further oriented to Mr. Land's Sixth Grade through the use of motion picture sequences presenting the class under the direction of Mr. Land as they might appear during an observation session.

A technique of filming the youngsters in Mr. Land's Sixth Grade, so that they appear to be reacting directly to the student teacher who is viewing the sequences, is employed in 60 different problem sequences on film, all involving the same group of youngsters. The

60 problem sequences are divided into three sets of twenty sequences each. Each of the three programs correspond to one school day and are roughly parallel in terms of the types of problems included. The selection of problems was based on the work of Hughes (1959). One half of the filmed sequences pose problems in classroom management for the student teacher, and the remaining are classed as communication problems (e.g., inattention, interjection of new information by a student). In each case, the student teacher is expected to react to the film as if he were in an actual classroom.

To further enhance the simulated practicum from the instructional standpoint, feedback sequences are available for each of the 60 problem sequences. The feedback sequences show the student teacher how the youngsters might react to his handling of each of the problematic situations. Presently, there are at least two alternative feedback sequences available for each of the 60 problem sequences. By using three remotely controlled projectors simultaneously, the motion picture projection of the children may be changed from the problem sequence to the feedback sequence instantly.

The proposed technique is not intended to be rigid in its approach. As is true in the "real world," the learner does not learn precisely how to react to each classroom situation. Instead, the objective is to develop a learning set towards a specific class of teaching problems — a way of operating within certain behavioral boundaries.

The original theoretical basis for the technique was based on the operant conditioning model. It was assumed that the behavior that student teachers exhibit in practice teaching is controlled primarily by external stimuli in the classroom. On this basis, it was presumed that the classroom behavior of teachers is "shaped" by the different stimulus events that occur in the classroom and that these stimuli act as reinforcers. Hence, exposition of educational methods or principles could be expected to help the student teacher talk about teaching, but only classroom experience could train the beginning teacher to teach. It was concluded that a technique which would permit a supervising teacher to control the stimulus events in the classroom, via simulated experiences, could effectively shape the behavior of the student teacher in training by reinforcing successive approximations of the desired behavior.

Using this theoretical model, a single sixth grade classroom, "Mr. Land's Sixth Grade," was simulated, and a continuous program of research on different instructional variables has been undertaken. This research has stimulated further improvement in the instructional materials.

From this original theoretical basis, Kersh developed the concept of "controlled feedback," i.e., feedback which is controlled by the experimenter. Originally, the feedback was to be manipulated as "reinforcing stimuli." However, as developed, the feedback may or may not be reinforcing. The student teacher is presented with short sequences on film showing the most likely response of the simulated class to the teacher's behavior. The feedback, although "controlled," now is considered simply as information in the ongoing instructional process. Thus, Kersh shifted the basis for his instructional simulation experience to an "information-system approach," as formulated by Ryans (1963). Ryans treats the teacher as an information-processing system which receives information, evaluates, make decisions, encodes messages for the students, and transmits the messages using appropriate communication channels.

Research on Classroom Simulation

In the past four years, many instructional variables have been researched, using "Mr. Land's Sixth Grade." For example, two projects (Kersh, 1963b; 1965) have concerned themselves with realism variables: (a) size of image (small or large) in the orientation-pretest experience, and training; (b) mode of feedback (visual or verbal); (c) mode of response (enacted or verbalized); and (d) motion in image (moving or still). Findings have suggested that realism in simulation is not an important variable in enhancing transfer, in comparsion with factors such as instructor differences. The findings indicate that less realistic (small) projections result in very small, though statistically significant, differences in comparison with more realistic, life size displays. Kersh concluded that the studies support the use of small projections throughout orientation and instruction, thus making the simulation technique adaptable in much less elaborate and less expensive laboratory facilities. Other findings indicate no statistically significant differences in post test performance of students who enact responses to problems on film and those who simply describe how they would respond. These findings suggest that classroom simulation as an instructional medium may be adapted in individualized or group-paced instruction where students use smaller-than-life-size images and respond by describing rather than enacting what they would do. These studies have been described in greater detail elsewhere (Vlcek, 1965; Kersh, 1965).

Vlcek (1965; 1966) has also studied the effects of the classroom simulation experience, using "Mr. Land's Sixth Grade," at Michigan State University. Specifically, he was interested in testing the transfer value of simulation instruction in a real classroom setting. Vlcek's experimental group received approximately nine hours of orientation, testing and training while his control group received orientation and testing only. Both of the groups were tested in a simulation laboratory, and were rated by a procedure similar to that used by Kersh. The experimental group performed better than did the control group in enacting their responses to post-test problems. However, no statistically significant differences were found in the two groups' ability to describe accurately the problem after viewing the film sequence.

The follow-up of the student teachers occurred during the next semester after the experimental or control experience. Using an observation scale that detected the extent to which student teachers applied the teaching principles taught in the simulated classroom experience, Vlcek found that both groups applied approximately equal numbers of principles with effective results, but that the experimental group used a greater number of principles with ineffective results. Further, evidence recorded on a "confidence scale" and "attitude scale" provided support for Vlcek's hypothesis that classroom simulation increases students' confidence in ability to teach, and indicated that the simulation experience was considered valuable.

Although simulation training had positive effects noted above, evidence indicated that training did not insure that students would be more effective in directing the classroom behavior of their students. Students who did not have experience in the simulated classroom employed fewer "teaching principles" but also were judged to have greater success in their limited efforts. Vlcek interpreted these findings to mean that the particular principles learned in simulation training are not effective in classroom management and control. Kersh (1965) points out that, although the "experts" who established the standards may have been wrong, another interpretation is that Vlcek's findings indicate that teachers who try a greater number of different ways of handing problems in the classroom may fail in their attempts more frequently than teachers who experiment less frequently. The important thing is that the study showed that the experimental group learned to be flexible in their response.

In the studies on classroom simulation mentioned above, the student is rated on his ability to: (1) identify the problem, and (2) respond to the problem. Because these are essentially two different responses that are rated independently on the basis of pre-established criteria, it is possible that four types of behavior may be exhibited by the student teacher (T):

(1) T identifies the problem and responds to the problem correctly;

(2) T identifies the problem correctly, but responds to the problem incorrectly;

(3) T identifies the problem incorrectly, but responds to the problem correctly; and

(4) T identifies the problem and responds to the problem incorrectly.

To evaluate the types of behavior change produced by the classroom simulation experience, the present investigator analyzed the scores of 20 Ts who received simulation training (motion picture mode) in a recent experiment (Kersh, 1963b). Figure 1a shows the number of Ts' responses to pretest and post test problem sequences that fall into each of the four categories. Figure 1b shows the gain from pretest to post test for each category. It is interesting to note that correct identification of problems did not guarantee correct responses to those problems. Students often find it easy to recognize problems, but rather difficult to alleviate the problems. Just as interesting is the fact that, while some students could not identify problems correctly, they still responded appropriately. Apparently, there exist some response patterns that are appropriate for a number of specific problems.

It is clear that after simulation training, the frequency of appropriate responses increased as well as the frequency of correctly identified problems. It is also evident that there was room for improvement in T's post test scores. Eighty-three of the 400 problem sequences, or 21 per cent, still were identified incorrectly, while 67 of the 400 sequences, or 17 per cent still were responded to inappropriately after training.

		Problem Identified			
		Correctly		Incorrectly	
		Pretest	Post Test	Pretest	Post Test
Response Made	Correctly	170	274	111	59
	Incorrectly	43	43	76	24

Figure 1a. Frequency of Responses for Each of Four Types of Behavior Shown by T on the Pretest and Post Test.

		Problem Identified	
		Correctly	Incorrectly
Response Made	Correctly	104	-52
	Incorrectly	0	-52

Figure 1b. Gain From Pretest to Post Test for Each of the Four Types of Behavior Shown by T.

As was mentioned in Chapter I, the simulation technique as used in the above mentioned studies requires a student teacher to respond to each problem sequence without explicit guidance. The student is not prompted by the experimenter with any principles or standards of teaching, or specific hints on how to behave, to guide his overt responding. In effect, a "learning by discovery" method is relied upon in training.

Research on Prompting

Studies that have compared prompted techniques with unprompted, "discovery"-type techniques have generally shown prompting to be of value in increasing learning or transfer. As far back as forty years ago, it was shown that prompts which suggested the response term facilitated paired-associate learning (Pan, 1926). Maier (1930) found that providing prompts helped subjects to solve his double pendulum problem. Irwin, et al. (1934) found that learning was enhanced when a rule or principle was presented, and its application by the learner was practiced.

Many studies on prompting appear under the heading, "learning by discovery." For the purposes of the present review, the more recent studies on learning by discovery will be examined.[2]

Craig (1956), using college students, required two experimental groups to solve a series of verbal problem tasks. The "directed" group was given a short general statement of the rules invovled, while the "independent" group was not given this information. Knowledge of results was provided by means of a punchboard. Craig showed that the group which was given the rules learned and retained more than the group which was required to derive the rules. The results failed to show a difference on the test of transfer to new rules. Craig concluded that large amounts of external direction help to provide the learner with an adequate background of knowledge to direct his future discovery.

Kittell (1957) placed sixth grade students into three treatment groups. The "minimum" group received only the test items and the instructions that there was an underlying principle for each item. The "intermediate" group was given the same list of test items, as well as a verbal statement of the principle involved. The "maximum" group received the same stimuli as the other groups, but in addition, was given the correct answers to the test items before the learners responded. All three groups received knowledge of results through the use of punchboards. On the learning test, both the "intermediate" and "maximum" groups scored dependably higher than the "minimum" group. On tests of retention, transfer to new examples, and transfer to new principles, the "intermediate"

---

[2] A comprehensive review of learning by discovery has been reported by Ausbel (1963, pp. 139-175). Some recent studies have been analyzed by Kersh and Wittrock (1962).

group was reliably higher than both the other groups. Kittell concluded that presenting rules to learners enhances retention and transfer to learned principles and provides the background which promotes discovery of new principles.

Forgus and Schwartz (1957) taught college students a new alphabet by three methods. The "observer" group received the new alphabet and its English equivalent, with instructions explaining the rule which was the basis of the alphabet construction. The "participant" group was given the same list, but were told only that there was an underlying rule which they were to derive in writing. The "memorization" group was given the same symbols as the above groups, but in a rearranged order so subjects could not easily derive the rule. It was shown that the groups which either were given the rule or were required to derive the rule were superior on tests of retention, transfer to new examples, and transfer to new rules to the group which only had to memorize the alphabet. Dependable differences were not found between the group that was given the rule and the group that was required to derive the rule.

Haslerud and Meyers (1958) presented to college students 20 coding problems designed to give the students two types of learning experiences. One half of the problems included rules which informed the subjects how to encode cryptograms. The other half of the problems required the subjects to derive the rules for themselves. The given and derived problems were alternated. On tests of learning, the subjects scored higher on the problems which included the rule than the problems which required the subjects to derive the rule. On a delayed test of transfer to new examples, no reliable difference in performance on the two types of problems was found. However, the scores were dependably increased for the problems which had no rule given, while the scores were dependably decreased for the problems which had the rule given. The authors took this finding as evidence that independently derived principles are more transferable than those given. Wittrock (1963) points out that if practice at "discovering" were generalized to related items, the experimental design may have contaminated the results since all the subjects were given equal practice on each of the two experimental treatments.

Wittrock (1963) taught college students to decipher transpositional cryptograms by one of four treatments. The "rule and answer given" group was presented with rules as well as answers to cryptograms, and were required to copy the answers. The "rule given, answer not given" group received rules, but were required to solve the cryptograms without benefit of having the answers. The "rule not given, answer given" group was required to derive rules from answered examples, while the "rule not given, answer not given" group was required to derive the rule without benefit of the answers. The treatment which presented rules and answers produced the greatest learning, but the treatment which presented rules and required the subjects to apply the rule to unanswered examples produced the greatest retention and transfer to new examples. These results were in close agreement with Craig (1956) and Kittell (1957) and were later supported by Wittrock and Twelker (1964a; 1964b).

In a recent study reported by Twelker (1964a), secondary school students were taught one rule, which when correctly applied to appropriate examples, made it possible for the subjects to decipher one type of transpositional cryptogram. Two variables were involved: (a) rule, given or not given, and (b) answer, given or not given. It was shown that giving rules enhanced transfer to new examples, but that not giving answers "standards" or rules of procedure, which apply to the problem. A list of each type of problem represented in simulation training appears in the appendix.

The instructional standards are so stated as to make the behavioral alternative clear. In other words, what is considered desirable behavior is contrasted with what would be considered undesirable. For example, one standard covers classroom problems involving rules of procedure when T is not informed of the rules. The standard states that in problems involving rules of procedure, T should defer to a person in authority, rather than establish his own rules. This is considered important since it is presumed that T is a student teacher on his first day's assignment. Each problem classification represents a unique combination of standards. For example, all problems that involve disorderly behavior are represented by the seventh and eighth standards. A summary list of standards appear in the appendix.

While revising the original standards and instructional procedures for purposes of the present research, it became evident that most problem sequences involved more than one standard. In cases where four or five standards might be considered relevant, only the two most important standards were chosen to rate T's responses. In a few cases, only one standard was selected. In cases where two standards were involved, the rating of three (3) was assigned when Ts behavior was considered effective by both standards; a rating of two (2) was assigned when one of the two standards was met but not the other; and a rating of one (1) was assigned when Ts behavior was considered ineffective by both standards. A zero (0) was assigned when T failed to respond at all to the problem. Variations in this rating procedure were employed when problems involved only one standard.

"hypotheses," "predictions," "expectancies," "sets," or "strategies") may be conceived of as mediators that act as determinants of overt responses. As these mediators are developed, a learner is equipped with a mechanism by which he can test the relevance or irrelevance of various stimulus cues. This conception of a mediational process is significant because it suggests that information that makes an overt response highly probable may not be especially suited to developing a number of systematic hypotheses or search models that may be used in an unprompted (transfer) situation by the learner to find an appropriate answer.

In simulation training, students are required to attend selectively to cues in the simulated environment so that they can make an appropriate, overt response that will solve the problem. A prompt, such as, "Look for the girl in the red dress who is sleepy," might identify the problem correctly for the learner, but give him little opportunity for developing search models to aid him in perceiving similar cues, and deciding upon their relevance or irrelevance in other problems. On the other hand, a prompt such as, "This is a problem that involves inattention," might cause the learner to test alternatively various hypotheses. He may systematically look for group inattention, as shown by general restlessness or drowsiness. If the group is perceived as attentive, he may then test the hypotheses that an individual is inattentive, and look for deviant behavior of a particular individual. Since the prompt does not tell the learner what specific type of inattention is involved, or how many are inattentive, the learner engages in the development of several strategies of observing a class, all of which may be used at a later time in an unprompted situation. It can be seen that this type of practice might be minimized by the presentation of a specific prompt.

From the reasoning above, the following hypothesis emerges:

> Hypothesis I: Learners who are given general information that helps identify what type of problem is to be responded to (classification prompts), will take fewer trials to learn how to recognize problems, and will assess problems more adequately on a transfer test than learners who are not given this information.

In the same manner, it may be reasoned that prompts that specify precisely the correct solution to problems may be of value as eliciting stimuli for the immediate occasion, but that they may not be suited to the learner's development, on his own, of various strategies in handling the problem. Thus, a prompt such as, "Stimulate an active response" may be of greater value in an unprompted, transfer situation than, "Have the pupil work at the chalkboard." The former prompt will give the learner opportunities for developing a number of alternative responses that might include, among others, chalkboard work. Further, the general prompt might act as a mediator that applies to a wide variety of different stimulus situations, rather than as a specific choice response that may apply to only a small variety of problems.

The second hypothesis involves the effect of the presentation of information of a general nature about the appropriate responses:

> Hypothesis II: Learners who are given information that helps the identification of appropriate responses (standard prompts), will take fewer trials to learn how to respond appropriately, and will respond more adequately on a test of transfer than learners who are not given this information.

In brief, the experimentation attempted to find if there was any value in prompting students with information relev[illegible] to how to respond, and to what to respond to, during a simulation experience, as compared with letting students "discover" for themselves the relevant standards involved and the situational cues to which they should respond. No hypothesis was made concerning the interactive effect of the two types of prompts. That is, the effects of each prompt was assumed to be additive.

III PROCEDURES AND MATERIALS

The procedures described below were used for both the pilot study and the experiment proper. The procedures are similar to those used in recently conducted research, wherein student teachers are instructed individually in a specially designed laboratory facility. This facility is diagramed in Figure 2.

Orientation

Student teachers were oriented to classroom simulation in four phases. First, Ts were told of the technique by the principal investigator during a regular class session. This orientation, which lasted approximately 15 minutes, covered the history and development of the technique, the instructors, the location of the laboratory, and other pertinent information. Students were told that everyone would receive a meaningful training experience, although the training may differ from student to student because an experiment was being conducted.

Immediately after this talk, Ts were oriented to the simulated classroom, "Mr. Land's Sixth Grade," with a slide-tape presentation. This second phase of orientation lasted approximately 20 minutes, and gave Ts an opportunity to see the simulated class in session, to learn the names and faces of the children, and to learn pertinent facts about each child. The script used in the presentation appears in the appendix. This phase of orientation was terminated by a drill, at which time various Ts were asked to review information previously presented.

As a follow-up to the above-mentioned activity, students were given cumulative record folders, which contained a picture of each child, achievement and health cards, anecdotal summaries of teachers' comments about each child, and a description of the school and community. Ts were instructed to study these materials in preparation for training.

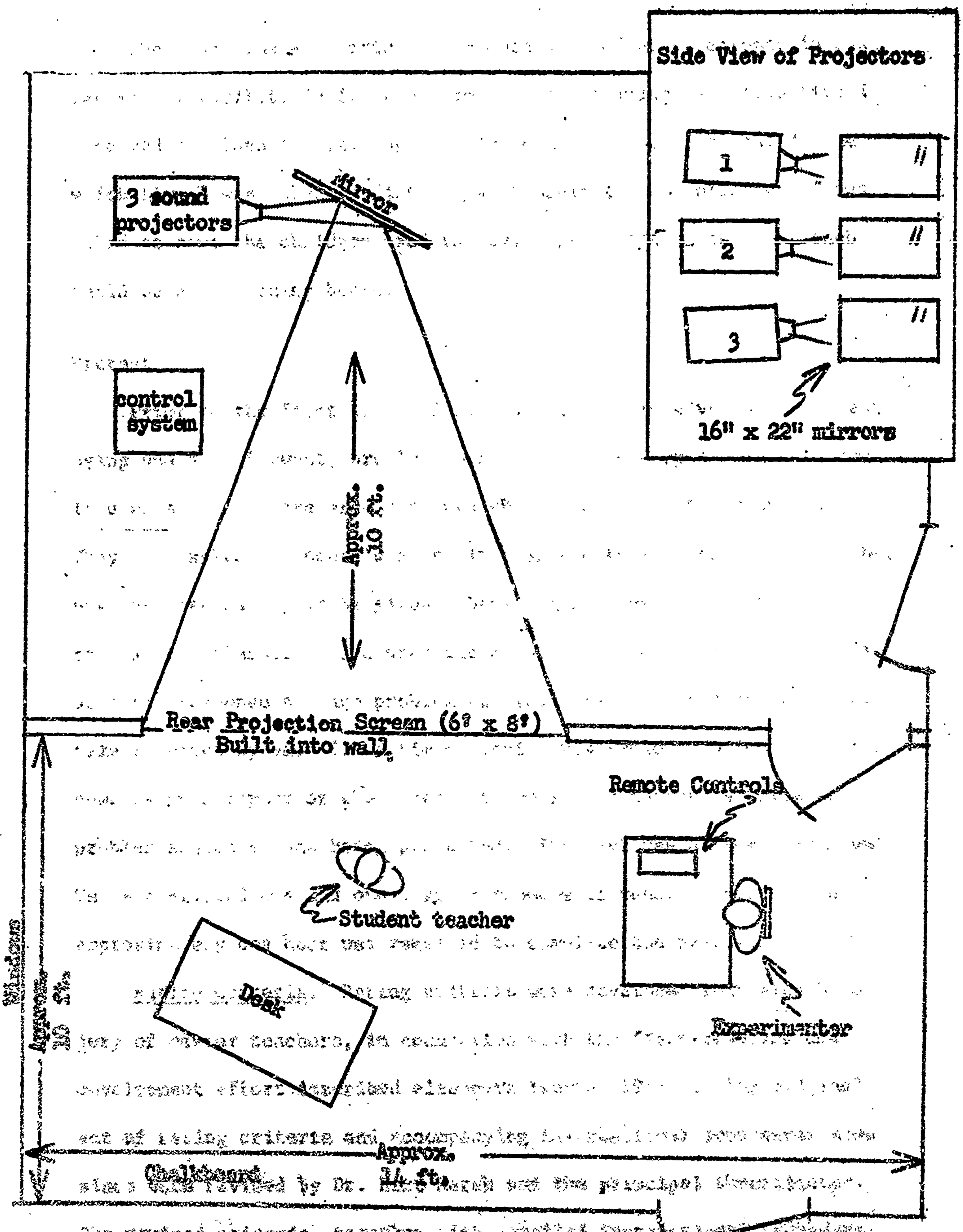


Figure 2. Simulation Facility Located at Oregon College of Education, 1965 Configuration.

The final phase of orientation consisted of an experience in the laboratory facility during the first day of training, at which time T observed Mr. Land interacting with the children (on film) and during which time T was asked to "introduce" himself to the children. T was asked to name the children, and to review pertinent facts about each child before training began.

Pretest

Prior to the first day of instruction, Ts were given a group test using one set of twenty problem episodes that make up a simulated day. Ts used specially prepared response sheets to record their answers. They were asked to record what their response to each problem would be, how the response would be given, where the response would be given from in the classroom, and when the response would be given. The timing of their response to each problem was indicated by a record of elapsed time recorded by each T. A timing device was employed which projected numbers on a corner of the screen at two second intervals while each problem sequence was being projected. The test was a timed test, and Ts were allowed one and one quarter minutes to record their responses. Approximately one hour was required to complete the pretest.

Rating Criteria. Rating criteria were developed initially by a jury of master teachers, in connection with the first research and development effort described elsewhere (Kersh, 1963b). The original set of rating criteria and accompanying instructional procedures have since been revised by Dr. Bert Kersh and the principal investigator. The revised criteria, together with detailed instruct[illegible]nal procedures, are included in the appendix. Each of these "scripts" include, among other things, the classification of each problem, and the instructional

[illegible] involved only [illegible] of [illegible].

"standards" or rules of procedure, which apply to the problem. A list of each type of problem represented in simulation training appears in the appendix.

The instructional standards are so stated as to make the behavioral alternative clear. In other words, what is considered desirable behavior is contrasted with what would be considered undesirable. For example, one standard covers classroom problems involving rules of procedure when T is not informed of the rules. The standard states that in problems involving rules of procedure, T should defer to a person in authority, rather than establish his own rules. This is considered important since it is presumed that T is a student teacher on his first day's assignment. Each problem classification represents a unique combination of standards. For example, all problems that involve disorderly behavior are represented by the seventh and eighth standards. A summary list of standards appear in the appendix.

While revising the original standards and instructional procedures for purposes of the present research, it became evident that most problem sequences involved more than one standard. In cases where four or five standards might be considered relevant, only the two most important standards were chosen to rate T's responses. In a few cases, only one standard was selected. In cases where two standards were involved, the rating of three (3) was assigned when Ts behavior was considered effective by both standards; a rating of two (2) was assigned when one of the two standards was met but not the other; and a rating of one (1) was assigned when Ts behavior was considered ineffective by both standards. A zero (0) was assigned when T failed to respond at all to the problem. Variations in this rating procedure were employed when problems involved only one standard.

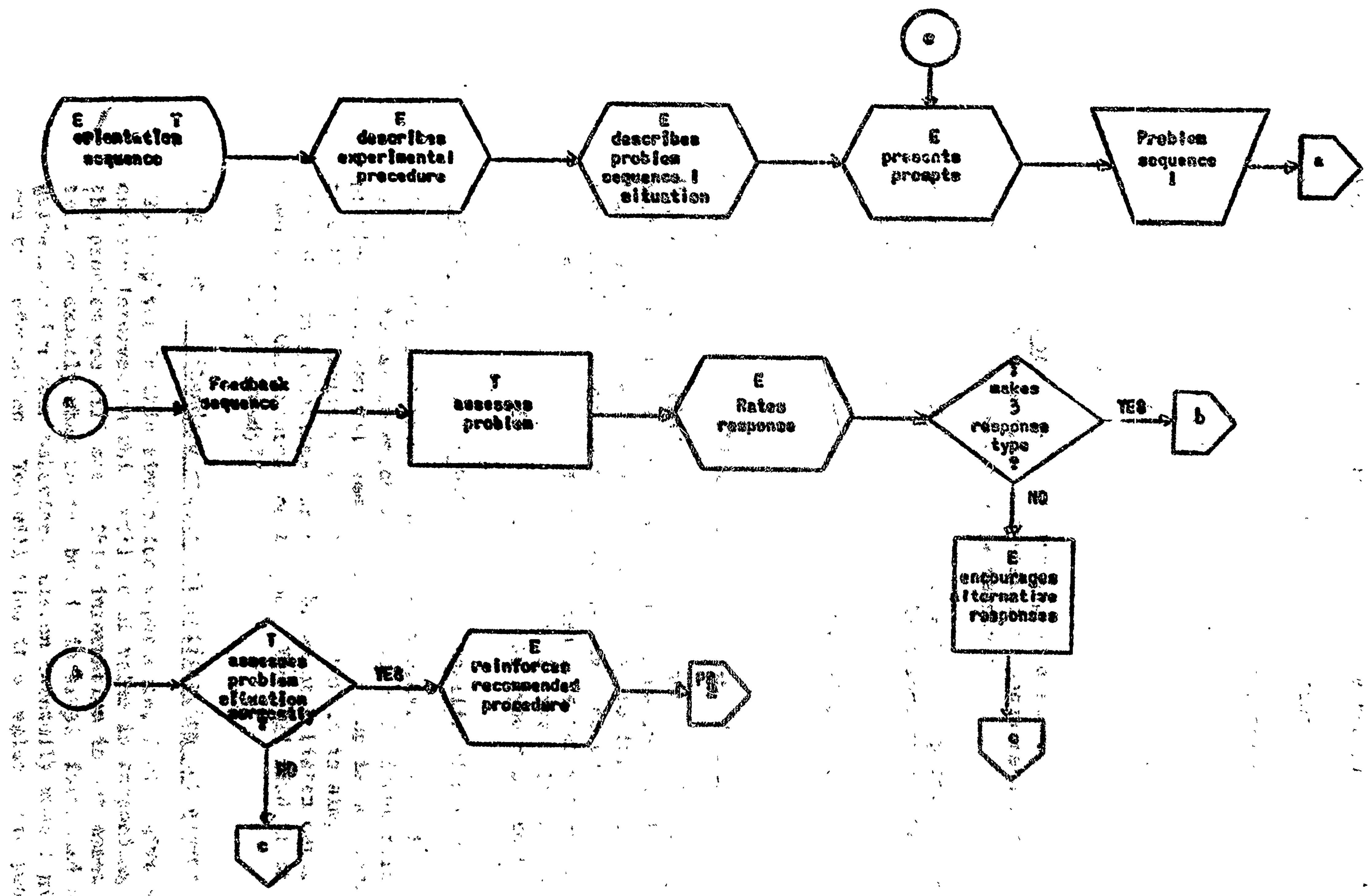
The scripts also contained information relevant to Ts assessment of each problem. Ts assessment was rated by noting his description of the stimulus situation and tallying the number of items of information which corresponded to each of those listed for each problem episode. The selection of relevant items of information was made using the standards for each problem as criteria. In addition to these items, some problems involved information which was included in the cumulative files or which was transmitted previously in the particular (simulated) day. This information was considered important, without which a particular pupil's behavior may have been misinterpreted. A composite rating for each T was made by summing the numerical ratings assigned T.

Instruction

At the termination of the orientation sequence, T was told that the next phase of simulation training involved the showing of twenty problems. The verbal instructions appear below.

> "Now, we are going through twenty problems, one at a time. First, I will describe the setting for each problem. I will indicate what the children are doing, and tell you where you are in the room. If I say that you are standing near the right side of the room, stand near the right side of the screen. I will then start the motion picture sequence which shows the problem as it develops. You are to look at the screen until you think of something that should be done about the situation. Enact your response at your first impulse. I will stop the picture on a frame--you just go right ahead and act out your response. If you feel that you should get closer to the situation, take a step or two in the direction you would move. Do you have any questions?"

Each problem sequence was repeated until Ts performance reached criterion for the most effective response and most adequate problem assessment. Numerical ratings were assigned for each T according to the rating procedure outlined above. The instructional procedure is shown by flowchart in Figure 3.



* According to Experimental Treatment

Figure 3. Flowchart Showing Instructional Procedure.

Prompts were presented immediately before the problem episode was shown on the screen. Ts that were given standard prompts received the relevant standard(s) for a given problem episode. Ts that were given classification prompts were presented with a narrative statement of the classification. Examples of a standard prompt and a classification prompt follow:

> <u>Standard prompt</u> - When learners appear disinterested or confused, stimulate a more active, interested response rather than make no effort to change the learner's response.

> <u>Classification prompt</u> - This is a management problem that involves a child's disregard for instructions from a teacher due to a fatigue reaction.

The following verbal instructions were given, accordingly to the experimental treatment.

<u>Treatment 1: Standard given: classification given:</u>

> "To help you in this simulation experience, I will give you certain information which will be of value in identifying the problem you will see, and responding to it. This information is not meant to tell you exactly what to do or what to look for, but it is simply furnished as guidelines for your behavior. Listen carefully when I give this information to you. You will find it of value. Are you ready to begin?"

<u>Treatment 2: Standard not given: classification given:</u>

> "To help you in this simulation experience, I will give you certain information which will be of value in identifying problems that you will see. This information is not meant to tell you exactly what to look for, but it is simply furnished as guidelines for your behavior. Listen carefully when I give the information to you. You will find it of value. Are you ready?"

<u>Treatment 3: Standard given: classification not given:</u>

> "To help you in this simulation experience, I will give you certain information which will be of value in responding to the problem you will see. This information is not meant to tell you exactly what to do, but it is simply furnished as guidelines for your behavior. Listen carefully when I give you the information. You will find it of value. Are you ready?"

Treatment 4: Standard not given; classification not given. No special instructions were given to T. E simply asked, "Are you ready to begin?"

After the appropriate prompts were given, the problem episode was shown on the screen. When T began to enact his response, E stopped the projector on a single frame. E then compared Ts response with the standards, and presented the appropriate feedback (either filmed or verbal). T was then asked to describe the problem situation to E. "Neutral" questions were asked, such as, "What happened in this situation?" or "What things contributed to the problem?", in contrast to "leading" questions such as, "Did you see anything happen before Mr. Land announced the fight?" E was permitted to ask, "Do you remember anything else that might have something to do with the problem?" Ts assessment was compared with the criteria as outlined under Stimulus Situations (see instructional procedures in appendix).

If T did not make an acceptable, enacted response to a problem, alternative responses were encouraged. Information which might have suggested alternative responses directly was not given until it was clear that T had no other alternative to suggest, or began repeating responses. T was encouraged with such statements as: "Can you think of another way to handle this situation?" or "Now that you have seen the problem and thought about it, can you think of a better way to handle it?" If T made a correct response, and assessed the problem situation correctly, he was reinforced in an indirect manner. Explanations were avoided unless T asked further questions about the episode.

The scripts for each problem sequence were prepared so that E could refer to them quickly and accurately during instruction or testing. In order to reduce the possibility of giving prompts other than those appropriate for a given experimental treatment, sets of scripts were used that showed only the relevant prompts to be presented for each treatment.

The scripts contain a description of the situation, typed exactly as it was to be communicated to T. Next, a description of the problem scene was typed as a reminder for E. The problem scene was not communicated to T before being projected. As part of the problem description, a "hold cue" was indicated which specified where the film was to be stopped in the event that T did not respond while the problem episode was being projected. Examples of the detailed instructional procedures for each problem episode are included in the appendix. All instructional procedures are written for use with a specially designed control system.

Post Test

One week after the termination of instruction, an individual post test was given in the simulation facility. The post test presented twenty new filmed problem sequences and required T to respond to each problem and to identify the problem. The post test was sufficiently different from the instructional sequence that it constituted a test of transfer. The post test took approximately one hour to administer.

Reliability between the four instructors for the rating of the responses, and the assessment of the problems, was shown to be .97 and .95 respectively, computed according to a procedure outlined by Winer (1962, pp. 124-132.)

At the termination of the post test, an affectivity measure was obtained from each S by means of a Thurstone-type attitude scale. Ss were also given a chance to make any comments they had about simulation training, and to suggest any improvements.

[illegible] 2 x 2 factorial [illegible] was [illegible] to study [illegible] given or not given, [illegible] Classification [illegible] given [illegible]. The four experimental groups [illegible] these two [illegible]:

1. [illegible] classification given [illegible]
2. [illegible] classification given [illegible]
3. [illegible] classification not given [illegible]
4. [illegible] classification not given (C_{Ng})

[illegible] given the pre and post test without simulation [illegible] of influence that may have caused changes during simulation [illegible]

[illegible] The subjects were [illegible] in the [illegible] at the University of Oregon. [illegible] subjects [illegible] participated in the experiment. [illegible] failed to finish the [illegible] phase of training [illegible] the post test. Data from the remaining [illegible] were analyzed. Participation in the experiment [illegible] for each [illegible] individually [illegible] to each of the experimental [illegible]

[illegible]

IV PILOT STUDY

The preliminary work consisted of a pilot study which tested the feasibility of using forty problems during instruction, as well as testing the suitability of the prompts.

Method

Experimental Design. A 2 x 2 factorial design was used to study two variables: (a) Standard: given or not given, (b) Classification: given or not given. The four experimental groups based on these two variables were:

1. Standard given; classification given (S_gC_g)
2. Standards not given; classification given ($S_{ng}C_g$)
3. Standard given; classification not given (S_gC_{ng})
4. Standard not given; classification not given ($S_{ng}C_{ng}$)

A control group was given the pre- and post test without simulation training to assess extraneous sources of influence that may have caused changes during simulation training, i.e. ongoing course of instruction.

Subjects. The subjects were undergraduate students enrolled in the School of Education at the University of Oregon. There were twenty-eight subjects who originally participated in the experiment. Of these, twelve failed to finish the instructional phase of training, or take the post test. Data from the remaining sixteen students were analyzed. Participation in the experiment was mandatory for all students. Subjects were individually assigned at random to each of the experimental conditions.

Materials and procedure. The procedure followed was outlined in Chapter III. However, forty problem episodes were used in training, rather than twenty as has been the case in previous research. The first twenty of these problem episodes were novel to X, while the last twenty were the same problems as those used for the pretest.

Results

Much difficulty was experienced because of the long training program. The instructional phase of training took so long that post tests were given during or after the final examination period of the term. It was impossible to administer the post test to any of the control group Ts, and to many of the experimental group Ts. As a result, data from only sixteen subjects were obtained. Because of these administrative problems, it was decided to run the experiment proper using only twenty problem sequences.

To assess the main and interactional effects of the two independent variables, a 2 x 2 factorial experiment was used (cf., Winer, 1962, pp. 241-247). Treatment group means for each dependent variable are shown in Table 1.

Adequacy of First Response in Training. This measure represents the adequacy of T's first enacted response to each of the 20 problem sequences in training. As shown by Table 2, the groups that were given information relevant to making an overt response performed more adequately on the first trial for each problem episode than groups that did not receive this information ($F = 9.31$, $p < .05$). Further, groups that were given information relevant to the classification of the problem episodes also scored higher than did the groups that were not given this information ($F = 5.23$, $p < .05$).

Total Repetitions in Training. On another measure of learning rate, the total number of repetitions of each problem episode required during learning, the standard variable also produced statistically significant differences ($F = 35.54$, $p < .01$). As shown by Table 2 the groups that were given standard prompts took less repetitions to learn than did groups who were not given these prompts.

Table 1

Treatment Group Means for the Measures of

Learning and Transfer

	Standard given; Classification given (n = 3)	Standard not given; Classification given (n = 5)	Standard given; Classification not given (n = 5)	Standard not given Classification not given (n = 3)
Training				
Adequacy of First Response	104.00	93.00	88.00	69.00
Total Repetitions	53.00	61.00	63.00	69.67
Transfer				
Post Test Response	53.00	45.40	47.20	43.70
Post Test Assessment	19.66	18.20	18.60	18.33

Table 2

Summary of the Two-way Analyses of Variance for the Learning Measures

Source	Adequacy of First Response			Total Repetitions		
	df	MS	F	df	MS	F
Standard (A)	1	1508	9.31*	1	1844.06	35.54**
Classification (B)	1	848	5.23*	1	202.03	3.91
A x B	1	60	- -	1	75.48	1.45
Error	12	162		12	51.89	

* $p < .05$
** $p < .01$

Table 3

Summary of the Two-way Analyses of Variance for the Transfer Measures

Source	Post Test Response			Post Test Assessment		
	df	MS	F	df	MS	F
Standard (A)	1	85.57	2.29	1	.83	- -
Classification (B)	1	8.00		1	2.94	1.65
A x B	1	60.67	1.62	1	1.21	
Error	12	37.39		12	1.78	- -

Post Test. Examination of Table 3 shows that no statistically significant differences were obtained ($p < .05$) on the post test response and assessment score. The evidence above, taken as a whole, only partially supported Hypotheses 1 and 2. The prompts did reduce the amount of repetitions required for learning, as well as increase the adequacy of T's first response to each problem episode. However, the results failed to indicate any statistically significant differences on the transfer test. With the extremely small *n*, the fact that statistically significant differences were obtained at all was surprising and most gratifying.

In summary, the evidence was taken to indicate that the prompts were meaningful to the students, since the information did produce changes in learning. A most important finding of the pilot study was that the training period was too long and resulted in administrative difficulties which could not be overcome under the present conditions of the experiment.

V THE EXPERIMENT

The materials tested in the pilot study were used to investigate the effects of the two independent variables, standard prompts and classification prompts. The main and interactional effects of these two prompts on learning and transfer were measured.

Method

Experimental Design. A two-factor design with matched groups (AxBxL design, Lindquist, 1953) was used to study the two types of prompts. The group pretest was used as the control variable. Subjects were blocked into six levels, and randomly assigned to the various A-B combinations within each level. The four treatment groups were:

(1) Standard given; classification given (S_gC_g)

(2) Standard not given; classification given ($S_{ng}C_g$)

(3) Standard given; classification not given (S_gC_{ng})

(4) Standard not given; classification not given ($S_{ng}C_{ng}$)

A control group was given the pre- and post test with simulation training to assess extraneous sources of influence that may have caused changes during simulation training, i.e., ongoing course of instruction.

Subjects. Eighty-nine junior-level college students were given simulation training with existing materials developed by Kersh (1963b). Subjects were drawn from two schools: the University of Oregon, and the Oregon College of Education. Subjects' scores on the pretest ranged from 18 to 49. A summary of the pretest data appears in Table 4.

Materials and Procedures. The procedure followed was similar to that used in the pilot study, and was described in Chapter III. Twenty problems were used in training.

Table 4a

Treatment Group Means for the Pretest Response Score

	Classification		
Standard	**Given**	**Not Given**	**Combined**
Given	38.44	38.94	38.69
Not Given	39.22	38.94	39.08
Combined	38.83	38.94	

Table 4b

Means for Each Level for the Pretest Response Score

(Summed over all Experimental Groups)

Level	Mean
1	29.58
2	35.92
3	37.67
4	40.58
5	43.00
6	46.58

Each treatment group contained 18 subjects, while the control group contained 17 subjects. In one of the experimental treatments, a subject dropped school part way through the experiment. Data missing on this subject were replaced with treatment group means.

To assess whether the experimental treatment groups differed significantly on the performance measures from the control group which did not receive training, Dunnett's procedure was used (cf., Winer, 1962, pp. 89-92). In this manner, the effectiveness of simulation training was compared with chance performance.

Dunnett's procedure is appropriate regardless of the over-all *F* ratio between groups. The harmonic mean was used in place of *n* to make allowance for the unequal sample size of the control group, in comparison with the experimental group. The .05 level test (two-tailed) was made on all comparisons. Table 5 summarizes the findings for the two post test measures.

On the post test response score, the three groups that were given prompts scored significantly higher than the control group ($p < .05$). The group that was not given prompts failed to score significantly better on this measure. However, this same group that did not receive prompts was the only group that scored significantly better than the control on the post test assessment score.

This analysis suggests that simulation training did produce scores above that shown by chance performance. It is also clear that the prompts had a differential effect on post test performance, depending on what prompts were given, and what criterion was being used.

Number of Sessions in Training. Examination of Tables 6 and 7 shows that the standard variable produced a significant effect on the number of training sessions required ($F = 6.26$, $p < .05$). The groups that were given

Table 5

Summary of the Comparison of all Experimental Treatment Group Means with the Control Group Mean for the Post Test Response and Assessment Scores

	Response			Assessment		
Group	Mean	D	t	Mean	D	t
1	45.66	5.72	2.95*	28.39	1.80	1.28
2	44.44	4.50	2.52*	28.11	1.52	1.09
3	44.72	4.78	2.46*	28.72	2.13	1.51
4	41.88	1.94	1.00	31.00	4.41	3.13*
Control	39.94			26.59		

Note 1: The symbol D refers to the difference between each experimental group mean and the control group mean.

Note 2: An asterisk indicates that the control group mean differed significantly from the particular experimental group mean, $p < .05$, two-tailed, as determined by Dunnett's procedure.

Table 6

Treatment Group Means for the Number of Sessions in Training

	Classification		
Standard	Given	Not Given	Combined
Given	3.00	2.72	2.86
Not Given	3.00	3.39	3.19
Combined	3.00	3.06	

Table 7

Summary of the Analysis of Variance for

the Number of Sessions in Training

Source	df	MS	F
Standard (S)	1	2.00	6.26*
Classification (C)	1	.06	.17
Level (L)	3	.56	1.74
S X C	1	2.00	6.26*
S X L	3	.10	.31
C X L	3	.29	.90
S X C X L	3	.37	1.15
Error	48	.32	

* $p < .05$

the standard took about 10% less time to complete instruction that those groups that were not given the standard. A statistically significant standard x classification interaction ($F = 6.26$, $p < .05$) revealed that instruction was faster when standards were given without classifications than when standards were given with classifications (see Figure 4). Clearly, the worst situation in terms of the number of sessions required for learning, was to present no prompts at all. The $S_{ng}C_{ng}$ group took about 20% longer than did the fastest group.

This evidence only partially supports the two hypotheses. Standard prompts reduced the time required for learning. However, the standard x classification interaction seems to indicate that too much information may be detrimental, in terms of learning speed.

Response Repetitions in Training. This measure represents the number of repetitions or trials required by T to meet the response criterion on each of the 20 training problems. Examination of Tables 8 and 9 shows that giving standards reduced the number of repetitions required in comparison with not giving standards ($F = 32.27$, $p < .001$). The groups that were given standards required about 16% less repetitions to meet response criterion than did the groups that were not given standards. Hypothesis 2 stated that learners who are given standard prompts will take fewer trials to learn. This hypothesis was supported. This finding suggests that for quickest learning, subjects should be prompted with information that helps them identify appropriate responses.

Assessment Repetitions in Training. This measure represents the number of repetitions required for T to meet the assessment c[illegible]rion on each of the 20 training problems. Tables 10 and 11 show the results of this analysis. An analysis of a statistically significant classification x

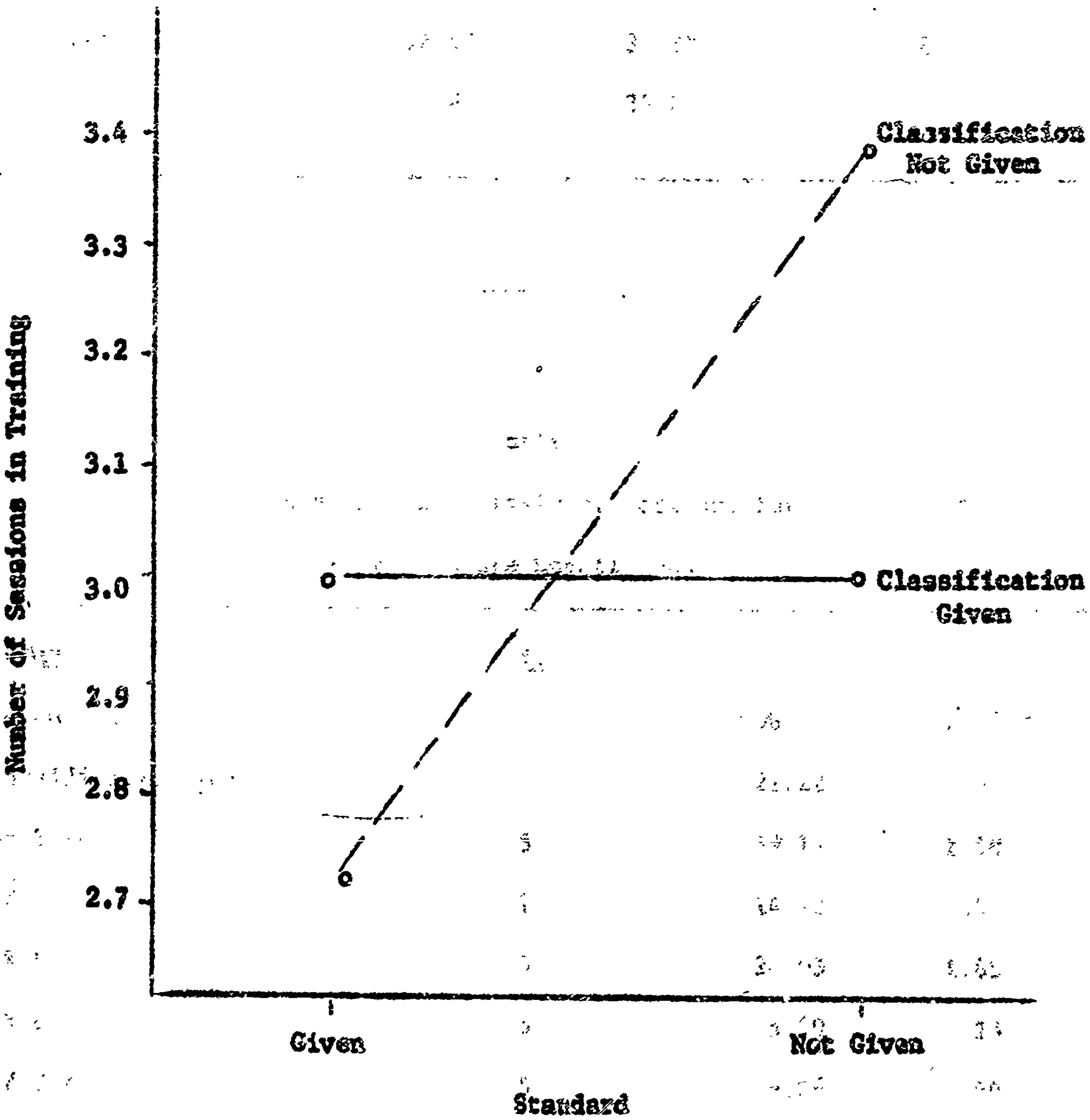


Figure 4

Profiles of Means Showing the Standard x Classification Interaction. The Dependent Variable is the Number of Sessions in Training.

Table 8

Treatment Group Means for the Number of Response Repetitions in Training

	Classification		
Standard	Given	Not Given	Combined
Given	30.11	32.11	31.11
Not Given	37.67	37.89	37.78
Combined	33.89	35.00	

Table 9

Summary of the Analysis of Variance for

the Number of Response Repetitions in Training

Source	df	MS	F
Standard (S)	1	800.00	32.27***
Classification (C)	1	22.22	.90
Level (L)	5	39.19	1.58
S X C	1	14.22	.57
S X L	5	26.03	1.05
C X L	5	5.59	.23
S X C X L	5	9.86	.40
Error	48	24.79	

*** $p < .001$

Table 10

Treatment Group Means for the Number of Assessment Repetitions in Training

	Classification		
Standard	Given	Not Given	Combined
Given	21.72	22.11	21.92
Not Given	21.78	23.00	22.39
Combined	21.75	22.56	

Table 11

Summary of the Analysis of Variance for

the Number of Assessment Repetitions in Training

Source	df	MS	F
Standard (S)	1	4.01	1.28
Classification (C)	1	11.66	3.74
Level (L)	5	2.05	.66
S X C	1	3.12	1.00
S X L	5	3.71	1.19
C X L	5	9.71	3.11*
S X C X L	5	4.63	1.48
Error	48	3.13	

* $p < .05$

level interaction ($F = 3.11$, $p < .05$) did not reveal any easily interpreted findings (see Figure 5). The data suggests that presenting classification prompts reduces the number of repetitions required to meet criterion. Of the four groups that required the least number of repetitions, three of these groups were given classification prompts. Further, of the four groups that required the most repetitions in training, three of these groups were ones that did not receive classification prompts.

These data provide only tentative support for Hypothesis 1. The main effect for the classification variable was not significant ($p > .05$), and the significant interaction revealed that the effect of presenting classification information depended on the pretest level of subjects.

Total Repetitions in Training. This measure represents the maximum number of repetitions or trials required by T to meet the response and assessement measure for each of the 20 problem episodes. Examination of Tables 12 and13 indicate that the groups that were given standards required about 16% fewer repetitions than did those groups not receiving standards ($F = 30.93$, $p < .001$). The presenting of classification prompts did not materially reduce the number of repetitions of the problem sequences required during learning. These data indicate that Hypothesis 2 was supported, and that Hypothesis 1 was not supported. Evidently, the presentation of standard prompts had a much more significant impact on the student's performance during learning than did the presentation of classification prompts. It is quite possible that the classification prompts were not entirely meaningful to Ts because they had not had prior training involving these concepts.

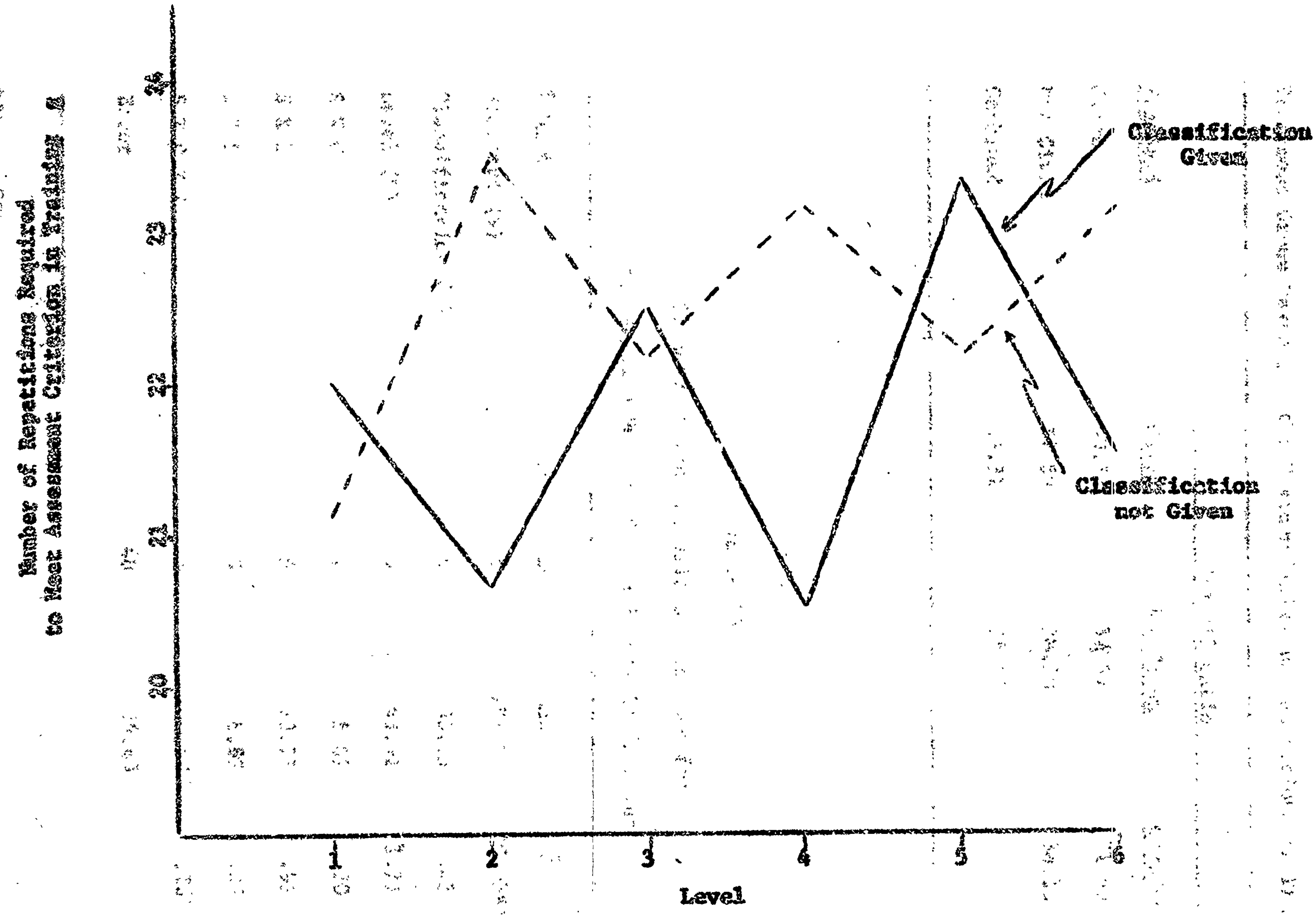


Figure 5

Profiles of Means Showing the Classification x Levels Interaction. The Dependent Variable is the Number of Repetitions Required to Meet Assessment Criterion in Training

Table 12

Treatment Group Means for the Maximum Number of Repetitions in Training

	Classification		
Standard	Given	Not Given	Combined
Given	30.72	32.67	31.69
Not Given	37.89	38.78	38.33
Combined	34.31	35.72	

Table 13

Summary of the Analysis of Variance for the Maximum Number of Repetitions in Training

Source	df	MS	F
Standard (S)	1	793.35	30.93***
Classification (C)	1	36.13	1.41
Level (L)	5	43.91	1.71
S X C	1	5.01	.20
S X L	5	23.71	.92
C X L	5	6.82	.27
S X C X L	5	6.58	.26
Error	48	25.65	

***$p < .001$

Adequacy of First Response in Training. This measure represents the adequacy of T's first enacted response to each of the 20 problem sequences in training. The examination of Tables 14 and 15 indicate that when standards were given, Ts scored about 20% higher on their first response in training than when standards were not given (F = 28.82, p < .001). This finding indicates that the standard prompts given were meaningful and actually guided students to make appropriate responses when presented initially with the problem. This evidence is in support of Hypothesis 2.

Adequacy of First Assessment in Training. This measure represents the adequacy of T's first assessment of each of the 20 problem episodes in training. The results indicate that neither prompt significantly affected this measure (see Table 16 and 17).

Post Test Response. This measure represents the adequacy of T's responses to each of the novel problem episodes. Examination of Tables 18 and 19 reveal a statistically significant standard x classification x level interaction (F = 3.45, p < .01). Because of the complexity of the design, in terms of the number of levels, the interpretation of this interaction is difficult (see Figure 6). To aid interpretation, the first-order interaction effects for treatment combinations at each level were assessed by t-tests (cf, Lindquist, 1953, p. 22). The standard x classification interaction for the lowest pretest level (mean = 29.6) was not statistically significant (p > .05). Futher, none of the groups produced what could be termed "good" performance. For the second pretest level (mean = 35.9), the results showed a significant first-order interaction (t = 2.33, p < .05). The best performance was shown by the group that was given classification prompts, but not given standard prompts. The worst performance was shown by the

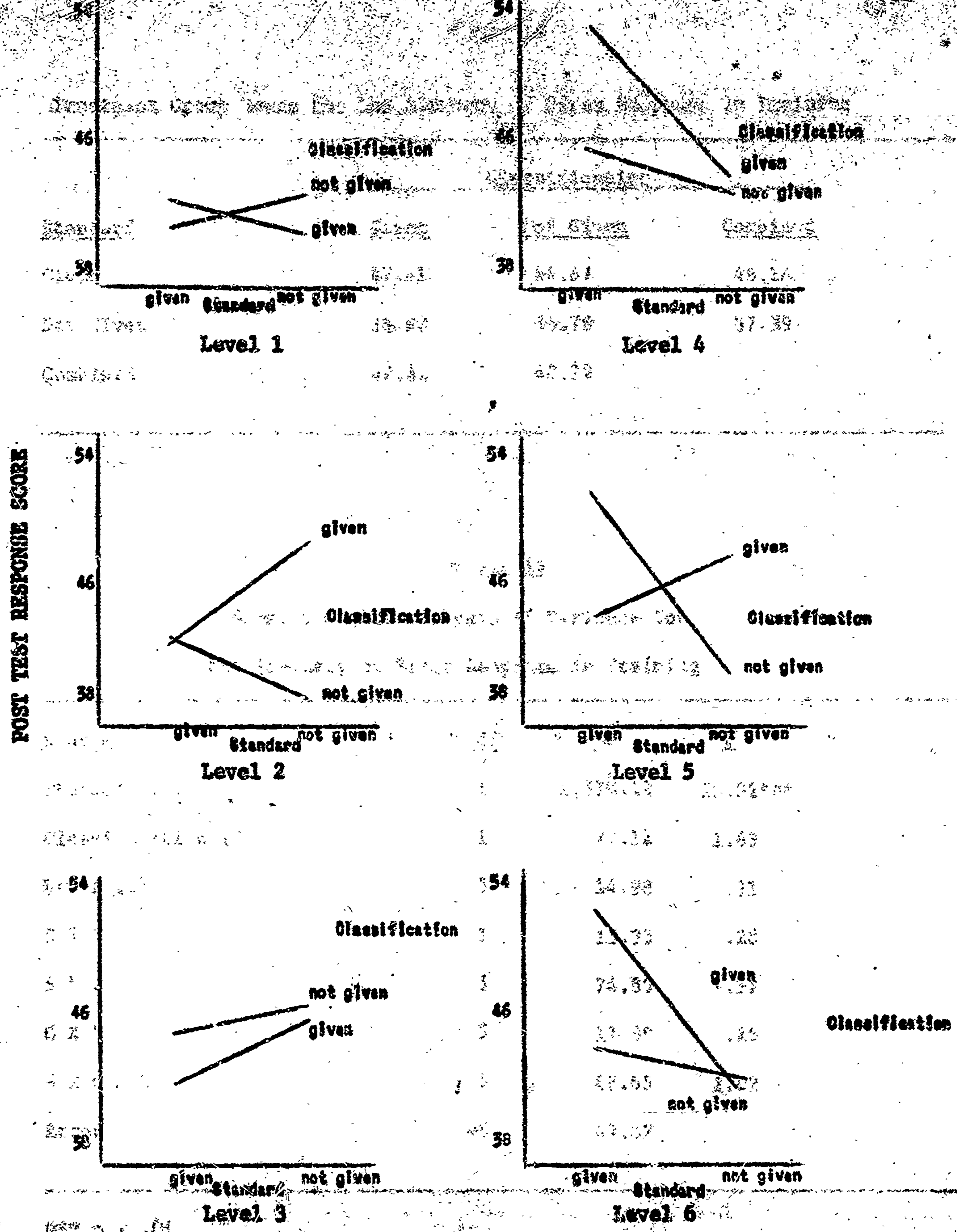


Figure 6
Profiles of Means Showing the Standard x Classification x Levels Interaction. The Dependent Variable is the Post Test Response Score

Table 1[illegible]

Treatment Group Means for the Adequacy of First Response in Training

	Classification		
Standard	Given	Not Given	Combined
Given	47.61	44.67	46.14
Not Given	38.00	36.78	37.39
Combined	42.81	40.72	

Table 15

Summary of the Analysis of Variance for

the Adequacy of First Response in Training

Source	df	MS	F
Standard (S)	1	1,378.12	28.82***
Classification (C)	1	78.12	1.63
Level (L)	5	14.98	.31
S X C	1	13.35	.28
S X L	5	74.89	1.57
C X L	5	11.89	.25
S X C X L	5	48.65	1.02
Error	48	47.82	

*** $p < .01$

Table 16

Treatment Group Means for the Adequacy of First Assessment in Training

Standard	Classification Given	Not Given	Combined
Given	29.00	29.50	29.25
Not Given	28.22	31.56	29.89
Combined	28.61	30.53	

Table 17

Summary of the Analysis of Variance for the Adequacy of First Assessment in Training

Source	df	MS	F
Standard (S)	1	7.35	.41
Classification (C)	1	66.12	3.70
Level (L)	5	50.55	2.83*
S X C	1	36.13	2.02
S X L	5	23.41	1.31
C X L	5	10.79	.60
S X C X L	5	36.33	2.04
Error	48	17.85	

* $p < .05$

Table 18

Treatment Group Means for the Post Test Response Criterion

	Classification		
Standard	Given	Not Given	Combined
Given	45.67	44.72	45.19
Not Given	44.44	41.89	43.17
Combined	45.06	43.31	

Table 19

Summary of the Analysis of Variance for the Post Test Response Criterion

Source	df	MS	F
Standard (S)	1	74.02	3.05
Classification (C)	1	55.13	2.27
Level (L)	5	48.05	1.98
S X C	1	11.67	.48
S X L	5	53.11	2.19
C X L	5	28.76	1.18
S X C X L	5	83.52	3.44**
Error	48	24.31	

** $p < .01$

group that did not receive prompts at all. For the third level (mean = 37.7), no significant first-order interaction was obtained. However, the fourth level (mean = 40.6) produced a significant interaction ($t = 2.51$, $p < .05$). The group that received both prompts was the only high scoring treatment combination. The fifth level (mean = 43.0) also showed a significant interaction ($t = 2.71$, $p < .05$), and revealed that groups receiving only one of the two prompts scored higher than groups receiving either both or neither of the prompts. The sixth level (mean =46.6) showed no significant interaction ($p > .05$).

In summary, it seems that the two lower scoring groups, in terms of pretest performance, benefited most from the presentation of classification prompts without the standard prompts. Among the students scoring somewhat above the mean performance on the pretest, the significant first order interactions suggested that only the presentation of both prompts was effective in enhancing transfer performance. On the other hand, in the next higher level (level 5), it was shown that presenting both prompts, or no prompts at all, was less effective than presenting only one of the two.

Looking at the data in a different way, it was evident that the three highest mean scores were obtained by groups that were given standards, while the three lowest mean scores were obtained by groups not receiving standards. In every case, the high scoring groups represented the upper three (of the six) levels.

Although the evidence is far from conclusive, the possibility is raised that for some students, transfer performance may be reduced with the presentation of too much information. The evidence failed to support the superiority of one type of prompt over the other.

Post Test Assessment. This measure represents the adequacy of T's assessment of the problems in each of the 20 novel sequences. Examination of Tables 20 and 21 reveal no statistically significant treatment effects. The levels factor did produce significant differences ($F = 3.59$, $p < .05$). Means for the assessment score, summed over all experimental groups, for levels 1 through 6, respectively, are: 30.58, 31.67, 26.83, 27.42, 30.50, and 27.33. The Newman-Keuls procedure was used to make comparison between the means (cf, Winer, 1962, pp. 238-239). From the outcome of the tests, it was concluded that level 2 differed from levels 3, 4, and 6, and that level 1 differed from level 3. With the exception of the sixth level, there appears to be a general trend for the middle ability level groups to perform poorer than either the high or low levels. This evidence would point up the importance of examining the factor of ability level in experimentation on laboratory simulation.

Affectivity. This measure represented T's rating of the entire simulation experience, as measured by a Thurstone-type attitude scale. Examination of Tables 22 and 23 reveal no statistically significant differences. The evidence, does reveal that students had a generally favorable attitude toward simulation. Statements on the attitude scale that were representative of the students' reactions were "Classroom Simulation does not waste my time." "Classroom Simulation has the reputation of being valuable." "Classroom Simulation is concerned with practical, down-to-earth matters."

Table 20

Treatment Group Means for the Post Test Assessment Criterion

	Classification		
Standard	Given	Not Given	Combined
Given	28.31	28.72	28.56
Not Given	28.11	31.00	29.56
Combined	28.25	29.86	

Table 21

Summary of the Analysis of Variance for

the Post Test Assessment Criterion

Source	df	MS	F
Standard (S)	1	18.00	1.23
Classification (C)	1	46.72	3.20
Level (L)	5	52.39	3.59*
S X C	1	29.39	2.02
S X L	5	22.30	1.53
C X L	5	2.49	.17
S X C X L	5	17.96	1.23
Error	48	14.58	

* $p < .05$

Table 22

Treatment Group Means for the Affectivity Score

	Classification		
Standard	Given	Not Given	Combined
Given	4.13	4.10	4.11
Not Given	4.27	3.89	4.08
Combined	4.20	4.00	

Table 23

Summary of the Analysis of Variance for the Affectivity Score

Source	df	MS	F
Standard (S)	1	.02	.02
Classification (C)	1	.73	.75
Level (L)	5	.71	.72
S X C	1	.57	.58
S X L	5	.49	.50
C X L	5	1.67	1.71
S X C X L	5	.98	1.00
Error	48	.98	

VI DISCUSSION

The results indicate that giving prompts that guided students' subsequent responses, made learning more efficient in terms of the number of sessions required for learning, the number of trials required to meet criterion, and adequacy of T's first response in training on each problem, as compared with not giving the same prompts. Also, the standard prompts apparently had some effect on the students' ability to respond to novel problems in a transfer task. However, this effect was not consistent across treatment combinations, and was limited to groups representing above-median performance on the pretest.

The presentation of prompts that helped students identify what stimulus features they should respond to in the simulated problems had little measurable effect on learning or transfer. Evidence did indicate that instruction was faster when standards were given without classification prompts. On the other hand, data from the post test response criterion may be interpreted to indicate that for the initially low-scoring groups, the presentation of classification prompts without the standard prompts was more beneficial than the presentation of both prompts. This evidence suggests that for some students, learning and transfer may be reduced with the presentation of too much information. The extent to which performance is reduced seems to depend on both the ability level of the student, and the type of information presented. Previous studies have shown that presenting prompts that enhance learning may retard transfer (cf. Twelker, 1964b).

It is clear that simulation training may provide a powerful vehicle for teaching principles of instruction or principles of classroom management and control. This is evidenced by the large increments in performance on the various learning measures. One cannot argue that the standard prompts were not meaningful to a large majority of students in light of these data. However, the lack of evidence that shows clearcut superiority on the post test from the presentation of standard prompts might indicate that the standards presented were not entirely appropriate for enhancing transfer performance. That is, they may not have been suited to the learner's development, on his own, of various strategies in handling problems similar to those taught.

It could be reasoned that exposure to standards immediately before the presentation of the problem may not be an instructional technique powerful enough to affect transfer performance. Perhaps, standards should be taught by means of a technique that makes the instructional principle more meaningful in terms of how and when it should be applied. One or two examples of situations in which it would be appropriate to employ the standard might be presented, possibly by film, before the student is required to use the standard, and respond in an appropriate manner with behavior that reflects the standard.

Simulation would be an ideal way of conducting such training, as it provides common referents for students. That is, students are provided with realistic experiences that are examples of situations in which the use of certai rinciples result in consequences on the part of the simulated class that are desirable. Students can consider the principles in light of these experiences and weigh the consequences of their action, shown through selected feedback sequences, against their own predictions. Thus, simulation training provides a powerful vehicle for teaching principles of instruction or principles of classroom management and control. Observational data from

Vlcek's study show that laboratory simulation is indeed a way of teaching principles that subsequently are employed in practice teaching soon after completion of the simulation training (Vlcek, 1965).

Evidence also indicates that future classroom simulation techniques need not require students to respond to simulated problems as though they had no principles of behavior upon which to base their actions except that which they bring in with them to the simulation laboratory. The findings reveal that the presentation of standards of behavior had a positive effect on learning rate, and for some students, transfer. These findings support previous research on the efficacy of using meaningful prompts during instruction.

The results have bearing on a current curriculum development effort at Teaching Research Division, which will produce "low-cost" classroom simulation materials for use either with individuals in a self-instructional manner, or with groups. To the extent that these findings are generalizable, they suggest that information pertaining to the educational principles, or "norms," to be taught, may be communicated directly to the student, rather than "discovered" through a time-consuming process of deduction.

An interesting finding of this present effort was the failure of prompting to increase student's performance on the post test above that gained by students in previous projects. The highest scoring group on the post test response score performed only 15 per cent higher than the control group which received no training. This increment, in view of a theoretical 39 per cent possibility for improvement, is small indeed.

Many questions remain to be answered. The evidence is not at all clear-cut concerning the specificity of prompts. It is possible that those presented were too specific, but this judgment awaits further investigation that establishes whether or not the prompts can be more meaningfully presented to students so as to increase transfer. Further, it is not known what effect the sequence of presenting the problems had on transfer. Different types of problems were presented in a more or less random fashion, as compared with presenting sequences that pertain to a particular problem classification all at one [illegible]. It is possible that this resulted in a failure for some or all of the standards to be established as effective mediators. Also, it remains to be seen how prompting interacts with the instructional method (group or self-instructional), and the length of the training period, in terms of the number of simulated problems shown. Further research is needed to uncover the importance of these and other variables.

REFERENCES

CRAIG, R. C. Directed versus independent discovery of established relations. J. educ. Psychol., 1956, 47, 223-234.

FORGUS, R. H., & SCHWARTZ, R. J. Efficient retention and transfer as affected by learning method. J. Psychol., 1957, 43, 135-139.

HUGHES, M. et al. A research report: Assessment of the quality of teaching in elementary schools. Salt Lake City: University of Utah Press, 1959.

IRWIN, F. S., KAUFMAN, K., PRIOR, G., WEAVER, H. B. On 'learning without awareness of what is being learned.' J. exp. Psychol, 1934, 17, 823-827.

KERSH, B. Y. The classroom simulator: An audiovisual environment for practice teaching. Audiovisual Instruction, 1961, Nov., 447-448.

KERSH, B. Y. Simulation in teacher education. Programed Instruction. 1963(a), 2, 3.

KERSH, B. Y. Classroom simulation: A new dimension in teacher education. Title VII Project No. 886, NDEA, Final Report, 1963(b).

KERSH, B. Y. Classroom Simulation: Further studies on dimensions of realism. Title VII Project No. 5-0848, NDEA, Final Report, 1965.

KIMBLE, G. A., & WULFF, J. J. "Response guidance" as a factor in the value of student participation in film instruction. In A. A. Lumsdaine (Ed.) Student response in programmed instruction. Washington, D. C. National Academy of Sciences--National Research Council, 1961.

KITTELL, J. E. An experimental study of the effect of external direction during learning on transfer and retention of principles. J. educ. Psychol., 1957, 48, 391-405.

LINDQUIST, E. F. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin, 1953.

MAIER, N. R. F. Reasoning in humans: I. On Direction. J. comp. Psychol., 1930, 10, 115-143.

PAN, S. The influence of content upon learning and recall. J. exp. Psychol., 1926, 9, 468-491.

RYANS, D. G. An information-system approach to theory of instruction with special reference to the teacher. SP-1079, System Development Corporation, 1963.

TWELKER, P. A. The effect of prompts on transfer of training. Paper read at the Western Psychological Association, Portland, April, 1964(a).

TWELKER, P. A. Rules, answers and feedback in learning, retention, and transfer of concepts. Doctoral dissertation, University of California Los Angeles, 1964(b).

WINER, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

VLCEK, C. W. Assessing the effect and transfer value of a classroom simulator technique. Michigan State University: Title VII, NDEA, Grant No. 7-32-0410-264 (Ed.D. dissertation project), 1965.

VLCEK, C. W. Classroom simulation in teacher education. Audiovisual Instr., 1966, 11, 86-90.

WITTROCK, M. C. Verbal stimuli in concept formation: Learning by discovery. J. educ. Psychol., 1963, 54, 183-190.

WITTROCK, M. C., & TWELKER, P. A. Prompting and feedback in the learning, retention, and transfer of concepts. Brit. J. educ. Psychol., 1964(a), 34, 10-18.

WITTROCK, M. C., & TWELKER, P. A. Verbal cues and variety of classes of problems in transfer of training. Psychol. Rep., 1964(b), 14, 827-830.

The appendix to this report is bound separately.

See the following document:

APPENDICES TO FINAL REPORTS

Classroom Simulation: Further Studies on Dimensions of Realism
Title VII Project Number 5-0848
Bert Y. Kersh, Principal Investigator

Prompting as an Instructional Variable in Classroom Simulation
Title VII Project Number 5-0950
Paul A. Twelker, Principal Investigator