DOCUMENT RESUME

ED 046 880 SP 004 621

AUTHOR	Orme, Michael E. J.
TITLE	Effects of Cueing and Modeling Variables in Teacher Training Systems.
INSTITUTION	Indiana Univ., Bloomington.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Educational Personnel Development.
PUB DATE	70
GRANT	OEG-0-9-247053-721
NOTE	20p.; Paper presented at annual meeting, AERA, Minneapolis, 1970
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Cues, Elementary School Teachers, Feedback, Microteaching, *Preservice Education, *Teaching Models, *Training Techniques

ABSTRACT

Theoretical considerations suggest that the differential effectiveness of teaching models and associated feedback procedures stems from their distinctive cueing properties. This led to the development of three treatment conditions which may be labeled "rating" (rehearsal of key discriminations), "observation" vicarious reinforcement), and "direct practice." Specific predictions were that increases in the response strength of probing (the dependent variable) would be greatest for the Direct Practice Group, next the Observer Group, and finally the Rater Group, and that within groups, Rater subjects would demonstrate the most significant within-repertoire shifts of sub-skills. Forty undergraduate teacher trainees were randomly assigned to three experimental and one control group. All were pre- and posttested in 5-minute video taped lessons which they taught for four fifth grade pupils. Each subject spent a total of 75 minutes in three treatment sessions over a 3-week period. Two trained raters coded relevant behaviors. Data treatment included analyses of variance and covariance and t tests. The first hypothesis was supported; the second was not although data was suggestive. Results provide further support for the theoretical assumptions. By increasing the distinctiveness of relevant model stimuli, observational learning is facilitated. (JS)

EFFECTS OF CUEING AND MODELING VARIABLES IN TEACHER TRAINING SYSTEMS

MICHAEL E.J. ORME
INSTITUTE FOR CHILD STUDY
INDIANA UNIVERSITY

Paper presented at AERA Convention, March, 1970

The research reported herein was performed in part pursuant to a contract (OEG-0-9-247053-721) with the U.S. Office of Education, Department of Health, Education, and Welfare, under the provisions of the EPDA Act. Additional support was provided by the Indiana University Faculty Research Fund.

While the current literature convincingly documents the facilitating effects of modeling in observational learning (e.g. Bandura and Walters, 1963), attempts to explain why modeling works are less successful.

Bandura (1965) makes a distinction between acquisition and performance in observational learning. He suggests that the acquisition of imitative responses can best be explained in terms of contiguity. On the other hand, the performance of such responses is thought to be due to reinforcements administered to the model (M). If this is so, then vicarious reinforcement or empathic learning (Mowrer, 1960) must be assumed to be operating in situations where S observes M performing given behaviors during the acquisition phase of learning, but does not practice the criterion performance.

The above argument is persuasive in the general case. However, its utility is quite limited when one begins to consider modeling as a presentation variable in an instructional system designed to impart teaching strategies.

To begin with, it cannot be assumed that the mere exposure of S to M consititutes a sufficient condition for imitative learning to occur. Exposing a person to a complex sequence of stimulation is no guarantee that he will attend to the entire range of cues; that he will select from the total stimulus complex the most relevant stimuli; or that he will even perceive accurately the cues to which his attention is directed.

Secondly, contiguity theory does not explain why one type of modeling should be more effective than another. This is an important consideration in teacher training, as multiple arrangements of modeling and associated feedback treatments are common. For example, trainees may be exposed to symbolic (written), symbolic-perceptual (filmed), or live models who demonstrate given teaching strategies. Or, the trainee may observe videotaped playbacks of his own performance, in which case his prior performance serves as a self-model. Finally, the frequency with which the model emits both desirable and undesirable behaviors can vary within and between each of these types of presentation. As the acquisition of a teaching skill has been shown to vary as a function of model presentation (Orme, 1967), factors other than contiguity must be operating in initial training.

One way of clarifying the nature of the stimulus event in associative learning is to propose that it is implicit responses to the individual features of stimulus objects (models) which function as stimulus events. These implicit responses then become associated with overt responses. Those stimulus features of a model which elicit such implicit responses may be termed cues. It follows that by increasing the distinctiveness of salient model cues, observational learning should be facilitated. Support for this view comes from research by Sheffield and Maccoby (1961) and Wulff and Kraeling (1961).

For training research, the major implication of this view seems to be that the treatment applied should not only require the trainee to focus on the correct end response, but require him to respond to salient cues that occur during the course of learning. Lumsdaine (1961) reflects this concern in suggesting that programmed learning specialists have been preoccupied with reward schedules to the detriment of the manipulation of prompting cues. This is almost certainly the case in teaching supervision where undue attention has been paid to corrective feedback procedures, and initial response guidance has been neglected.

The present study reflects these concerns, and has a strong applied focus. However, it approaches problems of "supervision" in terms of critical presentation and feedback variables which underlie the training process. In this respect, it has a theoretical base as well.

In a recent experiment Orme (1967) tested the assumption that the rate and level of learning a given teaching strategy (probing) varies as a function of the mode of model presentation. Two types of modeling, symbolic and perceptual, were manipulated in combination with feedback variables including prompting, self-feedback, and confirmation. Predictions were based on theoretical considerations which suggest that the differential effectiveness of varying model and feedback procedures stems from their distinctive cueing properties. Training procedures were varied along a continuum of increasingly available cues on the criterion behavior.

While the data fit neatly with predicted differences along the treatment continuum, between-group differences were reflected primarily in terms of increases in response frequency. Rearrangements of particular types of questioning techniques within a given teacher's response hierarchy of probes were less pronounced.

Now, to return to the theme of the earlier discussion, (probing will be described later): training systems which seek to increase both the range and availability of sub-skills within a given teaching strategy must systematically highlight salient cues in the initial phases of training.

One way of doing this is to have the teacher (T) rehearse key discriminations in initial practice by having him classify and label sub-skills which make up the criterion strategy, rather than directly practicing them in a lesson. This constitutes a form of indirect or parallel practice. Once T had viewed a model who demonstrated relevant skills, he would then be required to rate or analyze the model tape for frequencies and types of sub-skills demonstrated.

The prediction is that this type of treatment would be more effective in reordering sub-skills within T's response hierarchy than would a direct practice treatment. In the latter condition, the rearrangement of responses within T's repetoire would be seconded to overall increases in response strength.

Support for these predictions can also be seen in research on problem solving. Davies (1969) provides some data which suggest that in learning to apply a principle (as in the case in probing), labeling and recalling key concepts or discriminations is more efficient than direct practice.

The predictions outlined above lead to two different treatments in which salient modeling cues are differentially highlighted in initial practice.

A third condition in which T is yoked to another who engages in direct practice of the criterion behavior, might be termed, vicarious practice. The rationale for this treatment is based on earlier discussion about the role of empathic learning in imitative response acquisition. It constitutes a vicarious reinforcement treatment where T observes a model (a direct practice trainee) who receives differential reinforcement and feedback for desirable behavior. Salient cues would be distinctively highlighted as the observer would be exposed to both positive and negative instances of the desired behavior. The absence of any kind of overt practice should be offset by the strong emphasis on cueing and prompting. These latter procedures should facilitate the association of initially implicit responses with eventual teaching performance.

In summary, theoretical considerations suggest that the differential effectiveness of teaching models and associated feedback procedures stems from their distinctive cueing properties. This led to the development of three treatment conditions which may be labeled Rating (rehersal of key discriminations), Observation (vicarious reinforcement), and Direct Practice.

Specific predictions were that increases in the response strength of probing would be greatest for the Direct Practice Group, next the Observer Group, and finally the Rater Group. It was also predicted that within groups, Rater subjects would demonstrate the most significant within-repertoire shifts of sub-skills.

METHOD

The Dependent Variable: The dependent variable that was employed in the study is termed Probing. This is a basic questioning technique in which the teacher requires students to go beyond first-answer responses. It is designed to be used in lessons where pupil participation is prerequisite to the goals of instruction, and is intended to upgrade the quality of such participation. Once the pupil has responded by means of a question, answer, or comment, the teacher may probe this response by means of one or more probing techniques. These sub-classes or categories of probing are termed: clarification, critical awareness, refocus, prompting, encouraging alternatives, and redirection. The labels in each case generally reflect the teacher's goal when using a given type of probing.

The teacher's objective in training is to probe each pupil response during the course of practice lessons. He may do this by employing one of the techniques outlined above. If he does not, then by definition, he has non-probed. The training problem thus becomes one of teaching the subject to respond to any given pupil response as an S^d which comes to elicit a teaching probing response.

General Procedures: Forty undergraduate teacher trainees were randomly assigned to one of three experimental groups and a control group. Subjects in each group were pre and post tested in five-minute video taped lessons which they taught to four fifth grade pupils. Three experimental sessions intervened between pretest and post-test.

Before coming to the experimental sessions, all subjects were directed to prepare a five-minute lesson in which they would encourage discussion. The goal of the lesson was to employ questioning techniques in attempting to increase the quality of pupil responses.

To avoid undue attention being paid to the development of a new lesson for each teaching trial, teachers taught a different group of students in each lesson. This allowed them to retain the same basic subject matter in each lesson, while attending to improvements in probing techniques.

The experiment was run over a three week period. Time spent in each phase of treatment was controlled across groups. Each subject spent seventy-five minutes in treatment.

Subjects: Undergraduate teaching candidates enrolled in Educational Psychology classes constituted the initial pool from which random assignments to each of the four groups were made. The groups proved to be adequately matched on age, sex, grade point average, subject major, and class standing.

Typically, six of the ten subjects in each group were female, three-quarters of the group were elementary majors, and the mean age was between 20.6 and 22.7. Prior teaching experience, prior education courses, and prior psychology classes were also very similar across all groups. On the average, subjects within each group had taken three education courses, two psychology courses, and had taught less than one month prior to the experiment.

Fifth grade pupils were randomly selected from a pool of four such classes from a nearby school. All had prior experience with televised lessons. All training was conducted in a regular classroom at the school. Teams of four pupils were brought to the room when needed. No one team or pupil was exposed to more than four lessons. Each lesson was taught by a different teacher.

Treatments:

Group 1: Controls: Following initial instructions (common to S's in all groups) to come to the teaching sessions prepared to teach a five minute lesson, Group 1 teachers were videotaped in a five-minute pretest lesson.

They were then told that the sessions were designed to give them an opportunity to practice basic questioning techniques in discussion-type lessons, and given five minutes to plan the next lesson.

This cycle was repeated so that before his treatment was complete, each Control had been videotaped in four five-minute lessons with planning periods intervening between each teaching trial.

Group 2: Raters: Following the pretest, teachers in this group were allowed five minutes to read a handout which described probing, relevant subskills of the strategy, and examples of each type of probe. As with all other subjects, the purpose of the sessions was described as outlined above.

Raters then viewed Model tape A with E_I He cued subjects by pointing out when a probe was going to occur, identified each type of probe as it occurred, and indicated four occasions in which non-probing occurred.
1. Treatments in each group are summarized on the following page in Table 1.

---

1. Model A was a 33 year old male teacher who had been teaching at the fifth grade level for six years. He was pretrained in probing techniques, and was videotaped teaching a five minute lesson to four fifth graders.

He taught a general lesson which sought a general definition of the term architecture. Teacher-pupil discussion centered around an imagined contest in which the pupils were going to judge various buildings. Probing techniques were employed to develop criteria against which judgements could be made.

Two other model tapes were rated by subjects in Group 2. Model B taught an English lesson to fifth graders, Model C taught a lesson in critical thinking to an equivalent group.

Each of the three models were male, between 30 and 35, and had taught Grades 4, 5, or 6 for more than five years prior to training in probing.

Model A probed sixty percent of all pupil responses which occurred. Model B probed seventy-five percent of all possible responses, and Model C's probing rate was sixty-three percent.

All model lessons were taped under conditions like those for trainees. Lessons were five minutes long and taught to four fifth graders.

TABLE 1

SUMMARY OF STEPS IN TREATMENT
FOR ALL GROUPS
(N = 39)

STEPS IN TREATMENT	Experimental Groups (a)				
	I Controls (n = 10)	II Raters (n=9)	III Observers (n = 10)	IV Direct Practice (n = 10)	Time in Treatment
1. Teach: Pretest:	X	X	X	X	5 min.
2. Read Directions:		X	X	X	5 min.
3. View Model A:		X	Observe	X	15 min.
4. Plan: Trial Two:	X	Introduce Rating	Observe	X	5 min.
5. Teach: Trial Two:	X	Rate Model A	Observe	X	5 min.
6. Trial Two Feedback:		Feedback on Model A	Observe	X	10 min.
7. Plan Trial Three:	X	Rate Model B	Observe	X	5 min.
8. Teach: Trial Three	X	Feedback on Model B	Observe	X	5 min.
9. Trial Three Feedback:		Rate (b) Model C	Observe	X	10 min.
10. Plan Postest:	X	X	X	X	5 min.
11. Teach Postest:	X	X	X	X	5 min.
Total Time in Treatment	35 min.	75 min.	75 min	75 min.	75 min.

(a) Explanation of Symbols: (X) indicates that subjects in this group were exposed to the step in treatment indicated in Column 1.

(b) Rater subject's feedback session on Model B was 10 minutes; the "rate Model C" session lasted 5 minutes.

At the conclusion of the tape, E informed Rater subjects they would learn more about probing before teaching again by analyzing the model tape. In the fourth step in treatment, a simplified coding or rating sheet was introduced and explained. It is illustrated below in Table 2.

Note that Raters were required first to discriminate Pupil Responses. These act of course as cues which signal the occurence of a teacher probe or non-probe. If a probe did in fact follow a pupil response, then the Rater was required to further classify it as to type of probe.

The illustration in Table 2 was repeated three times on the single sheet given to each Rater subject. This allowed him to record his analysis for each of the three model tapes shown in treatment. This format and the use of the form were described in a five minute period.

TABLE 2

ILLUSTRATION OF THE ANALYSIS FORM USED BY
RATER SUBJECTS ON EACH OF THREE MODEL TAPES.

RATER TRAINING IN PROBING

NAME ______________

CLASS ______________

RESPONSES TO BE RATED:		FREQUENCY OF RESPONSES	TOTALS
PUPIL RESPONSES:	QUESTIONS:		
	ANSWERS:		
	COMMENT:		
	NON-RESPONSE:		
		ALL PUPIL RESPONSES TOTAL=	
TEACHER NON-PROBE (X)		N-P =	
TEACHER PROBE (/)		P =	
TYPES OF PROBES:	CLARIFICATION:		
	CRITICAL AWARE:		
	REFOCUS:		
	PROMPTING:		
	REDIRECT:		

The Model A tape was then replayed and analyzed by the subjects. E provided immediate feedback by replaying this tape, and identifying each of the relevant behaviors as they occurred.

This rate-then-feedback cycle was repeated again, this time showing a different teacher (Model B) employing probing techniques with a different group of fifth graders. Finally, a third lesson (Model C) was shown. However, instead of receiving feedback on this third tape, S's were given a five-minute planning session, and then taught their post-test lesson.

Group 3: Observers: Like S's in the Rater group, Observers were pretested, read the initial directions on probing. They were then yoked to Direct Practice Ss in Group 4. One observer was yoked to each Group 4 subject, and observed him through all phases of treatment from pretest, through viewing Model A with E, planning, and receiving feedback from E on direct practice trails. The observer then had a planning session in which his pretest lesson could be revised to incorporate more probing. Finally, he was posttested.

Group 4: Direct Practice: This treatment differed from all others in that Ss viewed playbacks of their Trial 2 and 3 lessons with E. E provided feedback based on S's prior performance by pointing out when probing had occurred (reinforcement) and identifying the type of probe employed. When non-probes occurred, E cued S by suggesting alternative teacher probes, and provided a specific example.

By the time each Direct Practice S had completed treatment, he had taught four times, viewed Model A, and received feedback based on two of his own lessons.

Measurement Procedures: During the study each trainee's lessons were videotaped for later analysis. Two raters were trained to code relevant behaviors. They worked from a Rater's Manual which defined each behavior in detail, and provided copious examples.

In addition to recording response frequencies of the behaviors outlined in Table 2, the raters also coded amount of Teacher Talk, Teacher Questions, Teacher Repetition, and Teacher Reinforcements.

Initial training was conducted on non-experimental tapes similar to those to be rated, and continued until the raters had reached 85% or better agreement on all sub-categories of probing and pupil responses. Percentage of agreement was computed on a minute-to-minute basis.

Once criterion was reached, independent, blind rating (i.e., raters coded each tape alone, and did not know which group or trial a given tape represented) of experimental tapes began.

Six five-minute tapes were lost due to recording problems and subject illness. Of the remaining 114 tapes, 22 were selected for reliability checks as analysis proceeded. For approximately each five tapes rated, one was selected for double rating. Raters were not told when they were coding reliability tapes. Tape selection ensured equal representation from each Group and each trial.

Inter-rater agreement on all behaviors rated varied from 67 percent on Teacher Repeats to 98 percent on Teacher Probes. Agreement on probing sub-categories varied from 79 percent on Critical Awareness to 93 percent on Clarification and Prompting. When a given behavior occurred infrequently percentage of agreement was severely curtailed. This is artifactual, and suggests the reported percentages are conservative.

RESULTS

Predictions about between-group differences in the response strength of teacher probes were ordered in terms of theoretical expectations about the type of initial practice. It was predicted that the Direct Practice treatment (Group 4) would lead to the greatest increases in probes, followed by the observer or vicarious practice group (Group 3), and finally by the rater or parallel practice group (Group 2).

One-way analyses of variance on pretest scores between all groups for all variables were performed to the outset to determine whether or not there were significant differences in entry behavior. There were no significant differences. For teacher probes, the F ratio (df=3,36) = 0.94. (Given these degrees of freedom an F ratio of 2.88 is required for significance at the .05 level).

F ratios for sub-categories of probing were well below significance as well. They varied from. 0.50 for Critical Awareness to 1.27 for Clarification. Pupil responses were non-significant as well (F 3,37 = 1.68).

Following the initial analysis on pretest scores, multiple analyses of variance for between-group differences on the posttest were conducted. In addition, covariance analyses using pretest scores as covariates to adjust postest scores were run as a further precaution.

Treatment differences on the dependent variable were highly significant and in the predicted direction. F ratios for major response categories on posttest between-group scores are summarized in Table 3. The differences between covariance and unadjusted means were so small that significance levels for any given F ratio were not affected. For this reason the results of analyses of variance are presented.

Note that on Total Teacher Probes, F was a highly significant 12.99. This is borne out by the results for sub-categories of probing. With the exception of Refocus, all of them were significant.

Between-group differences on teacher probes are illustrated in Figure 2, where treatment means for each group across relevant trials are plotted. The dotted lines for the Rater and Observation groups indicate that Trial Two and Trial Three means are not plotted. These groups taught only at pretest and posttest.

Treatment means in Figure 2 were brought to a common point by dividing treatment means within each group by the Trial One mean. This was done to facilitate comparisons between the learning slopes.

TABLE 3

F RATIOS AND SIGNIFICANCE LEVELS FOR MULTIPLE ANALYSES OF VARIANCE OF POST-TEST MEANS BETWEEN GROUPS ON THE MAJOR RESPONSE CATEGORIES

RESPONSE CATEGORY	F RATIO	DEGREES of FREEDOM	LEVEL of SIGNIFICANCE(a)
Total Pupil Responses	5.31	3,36	<.005
Total Teacher no-Probes	6.66	3,36	<.005
Total Teacher Probes	12.99	3,36	<.001
Sub-Skills of Probing:			
(1) Clarification	16.55	3,36	<.001
(2) Critical Awareness	3.11	3,36	<.05
(3) Prompting	3.88	3,36	<.05
(4) Redirection	5.40	3,36	<.005
Teacher Probes/Pupil Responses	9.51	3,36	<.001

(a) Required F ratios and their associated significance levels for df =3/36 are as follows:

F = 2.88, .05 level F= 4.43, .01 level

F = 3.54, .025 level F = 5.24, .005 level

F = 7.00, .001 level

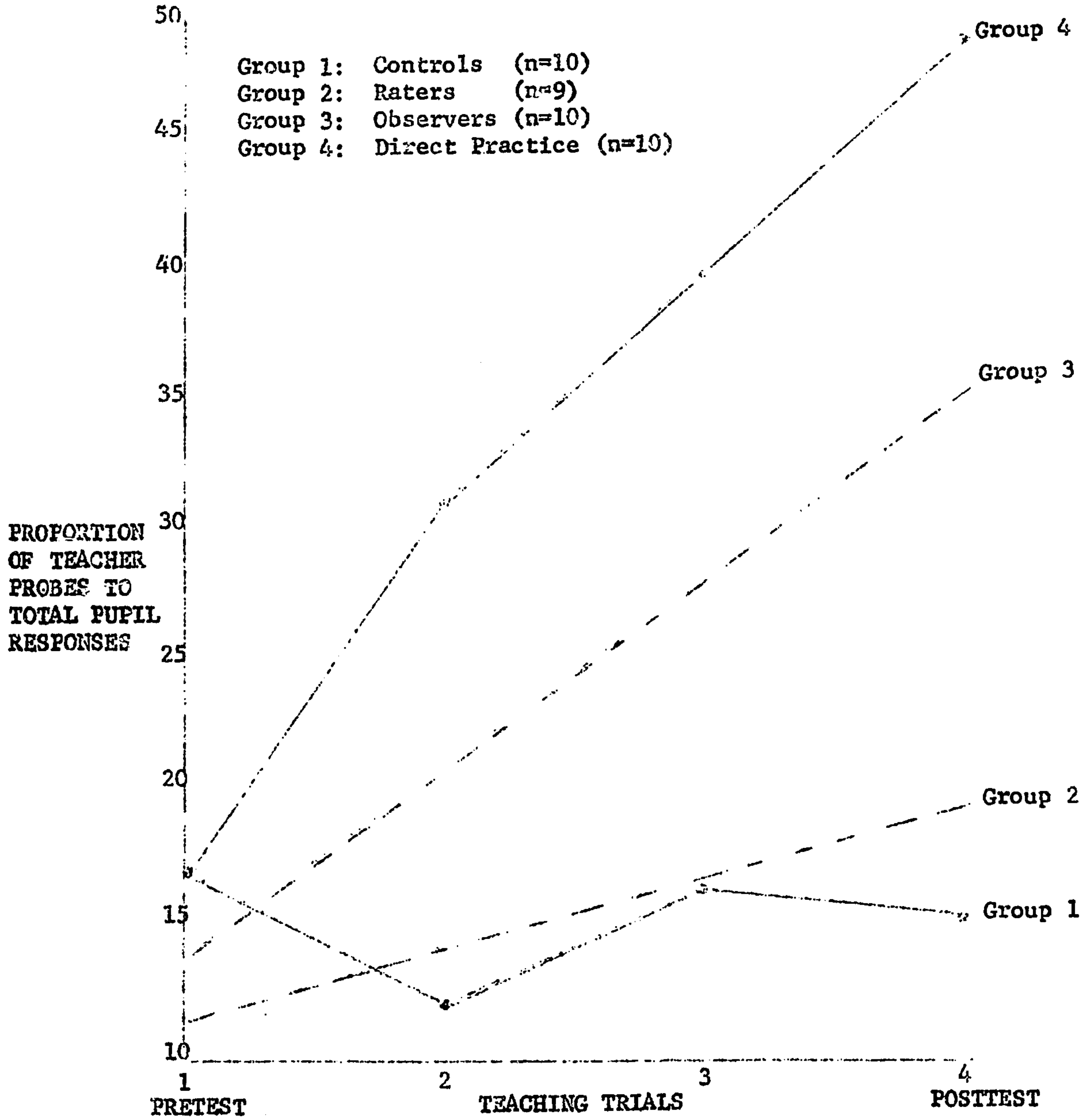

FIGURE 1. Treatment means for each group on relevant trails for proportions of teacher probes to total pupil responses.

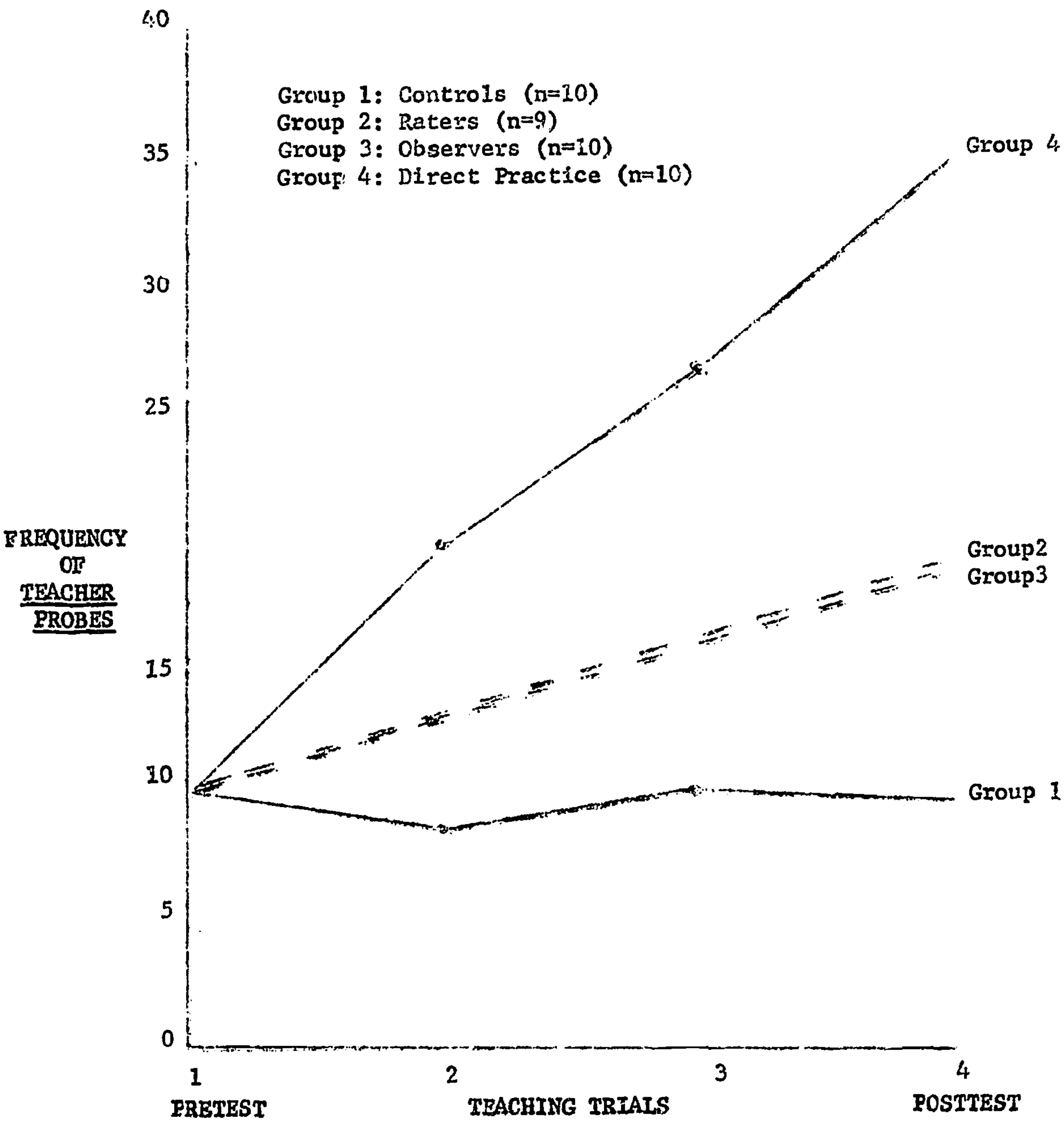

FIGURE 2. Treatment means for each group on relevant trials for teacher probes, pretest means brought to a common point for slope comparisons.

The functional utility of <u>Teacher Probes</u> is limited by, and clearly related to <u>Pupil Responses</u>. Since the number of probes is to some degree a function of the pupil responses that occur in a given lesson, there is a ceiling limit. Thus a proportionate analysis of <u>Probes/Total Pupil Responses</u> is highly desirable. However, it is somewhat problematic, as Rao, (1966) has predicted that a Cauchy distribution with an infinite mean and variance might result. This is a theoretical issue which is not evident in the practical situation. (If it were, the computer would still be running). In any event, there is some question as to the appropriateness of analyses of variance in such situations.

Descriptive data are certainly useful in this context, and are presented in Figure 1. In addition, analyses of variance and covariance (with pretest means as covariates) were performed as well. F ratios in posttest <u>Probe/Pupil Responses</u> between groups were significant beyond the .001 level for both types of analysis ($F_{3,34} = 9.51$ for ANOVA, and 10.636 on CANOVA). The results are entirely consistant with the results of other analyses.

<u>T</u> tests for between-group differences on <u>Probes/Pupil Responses</u> showed that <u>Direct Practice</u> subjects were significantly different from all other groups (Group 4 vs Group 3: $t = 1.74$, $p < .05$; Group 4 vs Group 2; $t = 4.23$, $p < .001$; Group 4 vs Group 1; $t = 5.34$, $p < .001$).

<u>Observers</u> (Group 3) were significantly different from <u>Raters</u> ($t = 2.03$ $p < .025$) and Controls ($t = 2.70$, $p < .01$). <u>Rater</u> performance was not significantly different from <u>Controls.</u>

T tests for differences between groups following a significant F ratio on posttest probing means, mirrored the differences reported above for <u>Probes/Pupil Responses</u>.

The second hypothesis stated that the Rater treatment would lead to greater shifts within subjects' probing response hierarchies than would the other treatment conditions.

Initial test have been performed, but further analysis is required. An analysis of covariance for between-group differences was performed on Trial Four Clarify/Total Probes proportions, with Trail One proportions as covariates.

The rationale for this is that on pretest performance, Clarification types of probes account for between 40 to 70 percent of Total Probes. Most likely they are the most available probing response, because they are easy to use, and at a superficial level require only a "What else"? or "Explain further" response from T. This holds true for groups of experienced and inexperienced teachers who have received training in California, Massachusetts, Florida, and Indiana. Thus, if the proportionate contribution of Claraficationa to Probes can be reduced and a concomitant increase in other types of probes brought about, an effective reordering of probing responses will have been achieved.

The results of the Clairfy/Probes covariance analysis on posttest means fell short of significance. The F ratio for df 3/33 was 1.97 (F = 2.88 is required for significance at the .05 level).

Within-group differences moved in the predicted direction. The Rater group showed a 28% drop in the proportion of Clarification to Probes, while the Observer group showed a 1% increase in Clarification. Direct practice differences between pre and posttest showed a 4% drop on this variable.

In short, the data while suggestive, do not provide adequate support for the hypothesis.

Finally, other findings of note should be briefly mentioned. Between-group shifts on non-probes consistantly dropped as probing increased. Significance levels were comparably high, and in the predicted direction.

Amount of Teacher Talk decreased from pretest to postttest in both the Direct Practice and Observer conditions. While between-group differences on the posttest were nonsignificant, Observer subjects showed a 20% pre to post test drop, and Direct Practice subjects showed a 10% drop. Controls on the other hand showed a 10% increase.

DISCUSSION

The results provide further support for the assumption that the diff-ential effectiveness of different teaching models stems from their distinctive cueing properties. By increasing the distinctiveness of relevant model stimuli, observational learning is facilitated.

While the Direct Practice treatment proved to be significantly more powerful than the Observer condition, the latter is highly attractive when one considers training large numbers of teachers. The number of probes in relation to Total Pupil Responses increased by two and one-half times from pre to post test. On this basis, it would appear that, in comparison with data from earlier experiements (Orme 1967), the Observer treatment is much more powerful than self-feedback procedures, whether or not models are used in initial presentation. In addition, the acquisition rate appears to be greater for the Observer condition than a treatment where trainees view the model alone, but receive feedback from E (supervision) based on their own prior performance.

If the number of trainees who are yoked to Direct Practice subjects can be substantially increased without sacrificing power, this type of training would be highly useful. However, it would likely be most efficient with trainees who engaged in considerable probing at pretest, for if initial response rates are low, doubling them may not be enough.

The results of Rater subjects while suggestive provide only limited support for the hypothesis that rehersing and classifiying key discriminations will lead to significant shifts within questioning hierarchies. Further, the treatment did not produce enough increases in overall response strength to make it practical. However, it should not be too difficult to increase the power of this treatment while retaining its basic character. E may have to shape the discriminations more systematically than in this experiment. Increases in time-in-training may lead to significant gains. And of course, if Rater training was incorporated into the Direct Practice treatment, significant increases in range of techniques and availability of selected techniques might be realized.

Finally, it should be mentioned that while observational learning from models is described as imitative response acquisition in the general literature something more appears to be involved in teacher training.

It would be difficult to defend the view that increases in probing were primarily due to imitation. Trainees go well beyond imitation of the model. In probing, the teacher must select a given probe both in terms of the goals of the lesson and a particular pupil response. Thus direct imitation is neither possible nor desirable. Considerable transfer of generalizations appears to be involved. This is an area which requires further research.

REFERENCES

Bandura, A., Influence of model's reinforcement contingencies on the acquisition of imitative responses, Journal of Personality and Social Psychology, 1965, 1, 589-595.

Bandura, A., and Walters, R.H., Social Learning and Personality Development. New York: Holt, Rinehart, and Winston, 1963.

Davies. I.K., "Structure and Strategy: A model for instructional decision making". Unpublished manuscript (mimeo.), Indiana University, Institute on Instructional Development, November, 1969.

Lumsdaine, A.A., Some theorecical and practical problems in programmed instruction, in Coulson, J.S. (Ed.). Programmed Learning and Computer - Based Instruction. New York: Wiley, 1961, 134-154.

Margolius, G.J., and Sheffield, F.D., Optimum methods of combining practice with filmed demonstrations in teaching complex response sequences: Serial learning of a mechanical assembly task, in Lumsdaine, A.A., and Glaser, R., (Eds.). Teaching Machines and Programmed Learning: A Source Book. Washington, D.C.: NEA, 1961, Chapter 3.

Mower, O.H., Learning Theory and Symbolic Processes. New York: Wiley, 1960.

Orme, Michael, E.J., The Effects of Modeling and Feedback Variables on the Acquisition of a Complex Teaching Strategy. Stanford University, Unpublished Ph.D., Dissertation, 1967.

Rao, C.R. Linear Statistical Inference and its Application. New York: Wiley, 1966.

Wulff, J.J., and Kraeling, D., Familiarization procedures used as adjuncts to assumbly-task training with a demonstration film, in Lumsdaine, A.S., (Ed.). Student Response in Programmed Instruction. Washington, D.C.: NRC, 1961, 141-154.