DOCUMENT RESUME

ED 187 575 SE 031 002

AUTHOR Blosser, Patricia E., Ed.; Steiner, Robert L., Ed.
TITLE Investigations in Science Education, Vol. 6, No. 2. [Expanded Abstracts and Critical Analyses of Recent Research].
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 80
NOTE 66p.
AVAILABLE FROM Information Reference Center (ERIC/IRC), The Ohio State University, 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (subscription $6.00, $1.75 single copy).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Abstracts; Educational Research; Research Methodology; *Science Education; *Science Instruction; *Teacher Education
IDENTIFIERS *Science Education Research

ABSTRACT

This issue of "Investigations in Science Education" (ISE) provides analytical abstracts, prepared by science educators, of research reports in the areas of teacher education and instruction. Each abstract includes bibliographical data, research design and procedure, purpose, research rationale, and an abstractor's analysis of the research. (CS)

ED187575

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC® Clearinghouse for Science, Mathematics, and Environmental Education

Published Quarterly by

The Center for Science and Mathematics Education
College of Education
The Ohio State University
1945 North High Street
Columbus, OH 43210

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 50¢ for Canadian mailings and $1.00 for foreign mailings.

SE 031 002

Volume 6, No. 2, 1980

NOTES FROM THE EDITORS

This issue of I.S.E. contains articles clustered into two areas: teacher education and instruction. The four teacher education articles emphasize different concerns. Nelson examined teacher effectiveness after instruction in strategies for teaching SAPA lessons. Rice reported on an instructional model for improving teachers' questioning skills. Tamir et al. formulated an inventory for use in the assessment of preservice secondary science teacher education programs. Willson and Garibaldi looked at the effect of teacher participation in NSF institutes on pupil achievement.

Within the instruction cluster, Bowyer and Linn studied scientific literacy within the context of SCIS goals. Hall reported on a method used to introduce students to the idea of chemical change. Herron and his colleagues looked at the effects of instructional methods on concept formation. Mathis and Shrum examined kinetic structure in a college biology class using audio-tutorial instruction. Mintzes and his colleagues also looked at individualized instruction at various levels: high school, junior college, and university. Smith and Rosenshein investigated questioning skills as they worked with teaching associates in charge of lower division physics classes.

We are pleased to report that participation in producing I.S.E. is continuing at a relatively steady state. However, if we were to increase our number of abstractors, each of them would have to produce fewer abstract-analysis articles for I.S.E. We urge you to invite your colleagues to participate if they are not already involved.

Patricia E. Blosser
Editor

Robert L. Steiner
Associate Editor

TEACHER EDUCATION

Nelson, Bess J. "The Relationship of Fifth- and Sixth-Grade Students' Achievement to Preservice Science Teacher Preparation." Journal of Science Teaching, 15(2): 161-166, 1978.

Descriptors--*Achievement; Educational Research; Elementary Education; Elementary School Science; *Elementary School Students; *Grade Point Average; *Lesson Plans; Preservice Education; *Science Course Improvement Project; Science Education; *Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Leon Ukens, Towson State University.

Purpose

The purposes of this study were: (1) to determine what effect instruction presented to preservice teachers prior to the time they taught prepared science lessons had on their teaching effectiveness as measured by the achievement of fifth and sixth grade students and (2) to determine if there was a relationship between the preservice teacher's grade point average (GPA) in college science courses and the effectiveness of his/her instruction.

Rationale

A study by Richardson (1973) indicated increased learning by mathematics students when their teachers had been presented prior teaching strategies. This study attempted to transfer Richardson's results to science teaching as indicated by the first purpose stated above.

A theoretical framework using pupil achievement as a measure for effective teaching was developed through a historical perspective from studies dating from the 1950s through the 1970s. A rationale for the second purpose was based on previous research which had produced conflicting results. This study attempted to clarify some of these conflicting reports.

Research Design and Procedure

A Campbell and Stanley posttest-only control group design was utilized. Randomization of the subjects is a key feature of this technique and was carried out in this study.

Preservice teachers from two science methods classes were randomly divided into two groups; an experimental (N=17), and a control (N=16). Each of these groups was then subdivided into two additional groups based on the student's GPA in eight required credits of college science courses. The preservice teachers' mean GPA was 2.72 with a standard error of 0.15. The high level GPA treatment involved those teachers with a GPA more than one standard error above the mean while the low level GPA treatment involved those teachers with a GPA more than one standard error below the mean.

The experimental group received 45 minutes of instruction on strategies they could use in teaching three lessons from Module 78 of SAPA II (Formulating Hypotheses, C) to a small group of fifth and sixth graders from a nearby school. The control group members received no such instruction but were instead given lesson plans and were told to prepare the lessons for presentation on their own. Students from the nearby school were randomly assigned to the 33 preservice teachers.

An evaluation instrument was administered, by the experimenter, to the fifth and sixth graders immediately following the last lesson. Scores on this instrument were used as the achievement measure. The evaluation instrument was one used with SAPA II, Module 78 and had been previously tested for reliability and validity.

Data were analyzed using a two-factor analysis of variance design. Experimental treatment (instruction and non-instruction of how to teach the lesson) and teacher science GPA (high and low) were the two factors used.

Findings

Analysis showed that the criterion variable means were higher for the experimental treatment at both GPA levels than for the control group. Further analysis indicated a significant difference between student achievement scores from the two treatment groups. The college science course GPA of the subjects, however, was not significantly related to the criterion variable. In other words, treatment affected the criterion variable, GPA did not.

Interpretations

The students of preservice teachers receiving instruction on how to teach a particular series of lessons have higher achievement on the lessons than students not receiving such instruction.

Science knowledge, as depicted by a science course GPA, indicated that preservice teachers can be effective in science teaching regardless of their science background knowledge.

Methods courses should devote more time to actually working through prepared science lessons with the focus of keeping the preservice teacher's mind set on completing the objectives of the lesson but not at the expense of flexibility.

Results may also indicate that preservice teachers could be valuable assets in local schools by coming into those schools, after receiving training, and acting as demonstration teachers.

ABSTRACTOR'S ANALYSIS

In order to establish how this study fits in with existing research, one must first look at the two questions studied and how each of them fits into a theoretical base. The two questions were: (1) How would specific preservice instruction on how to teach a series of science lessons affect how well a group of selected fifth and sixth graders would do on a test covering the subject matter, and (2) what effect would the preservice teacher's science GPA have on the performance of these subjects on the test?

While this study relates to these two important research areas, it contributes little new information. Although not strictly for replication, this research does try to fit in with the larger research area depicted in the rationale. While studies of this nature are important in developing a larger data base for educational decision making, this one seems to be particularly vulnerable to criticisms in two areas. One is using the preservice teacher's GPA as a measure of science knowledge in question number two, and the other is the use of such a small sample size in both questions.

While quoting the literature on the relationship between preservice training style, the first question was actually built out of one doctoral study by Richardson (1973) with mathematics preservice teachers. The researcher was attempting to transfer Richardson's strategies with math teachers to science teachers. While this is legitimate, the report did not indicate if the strategies used in this study were identical to Richardson's or were modified for this particular study. The exact relationship between the two studies was not developed.

A worrisome part of the entire study relates to the small number of people involved, especially the fifth and sixth grade students. The report indicates that each teacher taught more than one or two students but does not indicate the average class size. The data table indicates 30 fifth and sixth grade students were involved with the 33 preservice teachers. But if each teacher had more than one or two students, how could this be with 30 students and 33 teachers? One way would be to exclude those preservice teachers having a GPA falling within one standard error from the mean. The report does not indicate how many were in this group or what was done with them.

The second question was built out of conflicting research results from previous researchers. While research built on a conflicting base may add to which direction the conflict may go, this study made an assumption in studying this second question which may make the results invalid when viewed as a part of a larger picture. The researcher assumed the preservice teacher's GPA for eight semester hours of science content instruction taken prior to this study was related directly to that teacher's science knowledge. This certainly may not be a valid assumption especially if we examine some questions. If these preservice teachers had different instructors for the same course, can we not imagine a case where a C student knows more science than a B or even an A student merely because of a particular instructor? This seems to be especially significant when a GPA of .30 represents the difference between a teacher with a high science knowledge and one with low science knowledge. Also, when these courses were taken prior to this study might influence the teacher's knowledge at the time of the data collection. Perhaps another objection, although arguable, to this assumption involves the nature of science instruction in college classes, which is usually heavily content oriented, compared to what the preservice teachers are asked to teach: a process oriented series of lessons. Other ways of measuring science knowledge may be easier to accept.

It appeared from the report that the purpose of this study was to provide more information in areas where there were conflicting results from previous studies. Because of some of the objections raised earlier in this abstract, this study does not seem to add significantly to the decision-making process. Generalizing the results to a larger sample would be extremely arguable as this study used science knowledge based on courses at a particular college and used a very unrealistic teacher-student ratio in actually carrying out the research.

REFERENCE

Richardson, L. I. "Student Achievement in Solving Verbal Problems as Related to Teacher Preparation." Unpublished Doctoral dissertation. Institute on School Learning and Individual Differences, Nashville, Tennessee, 1973.

Rice, Dale R. "The Effect of Question-Asking Instruction On Preservice Elementary Science Teachers." Journal of Research in Science Teaching, 14(4): 353-359, 1977.

Descriptors--Educational Research; Elementary Education; *Elementary School Science; *Inquiry Training; Instruction; *Preservice Education; *Questioning Techniques; Science Education; Teacher Education; *Teaching Techniques

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Robert W. Johnson, Towson State University.

Purpose

The purpose of the study was to document the efficacy of an instructional model designed to attain a singular objective--improve the questioning behaviors in preservice teachers undergoing basic professional training. The hypothesis tested by the study was clear and uncomplicated. It stated that preservice elementary teachers receiving instruction on specific question-asking strategies would show significant improvement over similar persons not receiving such instructions.

Rationale

The conduct of this study was influenced by a growing awareness on the part of teacher educators of problems related to teaching competencies. Asking questions is one of the most basic skills practiced by teachers. The research climate in teacher education over the past two decades has nourished numerous and, on occasion, notable studies articulating varying degrees of incompetency in this important teaching skill. A complete review of the research would uncover a preponderance of evidence suggesting just how inept classroom teachers have been at developing critical thinking skills in their students. A recurring theme allows that teachers, irrespective of grade or discipline taught, lacked the formal training in the development and application of effective questioning strategies.

A basic assumption of this study was that skill in questioning could be taught, practiced, and evaluated. Further, it was assumed that improvement in teacher questioning behavior could be measured quantitatively.

Research Design and Procedure

The study focused on a small sample of 10 randomly selected subjects. All were training to become elementary school teachers. They came from a pool of undergraduate majors enrolled in an elementary science methods course. The subjects were divided equally into two groups--Instructional and Non-Instructional.

Three performance criteria constituted the dependent variables of the study. The instructional group was targeted for instruction on (1) wait-time after each question, (2) number of questions asked per minute, and (3) cognitive level of questions asked.

The special training incorporated into the research design for the instructional group utilized and included (1) film presentations, (2) reading/discussion periods of selected research studies on teacher questioning behavior, and (3) instructor/student post-lesson conferencing sessions. The latter assisted the subjects in analyzing lesson presentations with regard to each of the three performance criteria. The noninstructional group received no such special training during the conduct of the investigation.

All ten subjects prepared and taught a series of six science lessons to elementary school children over a period of two weeks. They used materials from the SCIS Program.

Quantitative observations on the study's three dependent variables were made according to specified criteria by a panel of three rater/judges. Each subject's performance was assessed from the written transcripts obtained from audiotapes of lessons one and six. These observations came to represent pretest and posttest performance scores.

For the statistical treatment of the data, the author obtained and reported interrater reliability coefficients with respect to each of the three dependent variables. Analysis of variance was used to test the null hypothesis of the study. With such a small subject sample, the ANOVA was employed to show significance in the F test between the instructional and noninstructional group. The Newman-Keuls *post hoc* test assisted the author to determine difference between individual means.

Various controls were imposed to insure maximum reliability and validity of results.

Findings

The results of the study unequivocally supported the hypothesis. The instructional group significantly increased the mean wait-time between asking a question and calling for a student response. The five subjects were asked 56 percent fewer questions per lesson segment at the conclusion of the investigation. Lastly, this group scored a measurable increase in the cognitive placement of questions asked over the period of the study.

The noninstruction group failed to show any significant change with respect to the three performance variables.

Interpretations

The study's investigator concluded that improved questioning skills can be taught to preservice teachers by methods instructors who carefully target their instruction toward attaining that objective. Important strategies to focus on, it was suggested, were (1) wait-time,

(2) number of questions asked, and (3) the asking of higher cognitive level questions.

Evaluating the study, the author stated that the children taught by the study's subject teachers became more involved in the manipulation and exploration of science equipment. Curiosity was stimulated rather than stifled by more effective questioning. He attributed this corollary effect to a growing awareness and subsequent improvement in the questioning performance of the subject teachers.

Teachers at all levels of professional development, the study implied, need to become cognizant of the importance of developing questioning strategies that encourage both critical as well as creative thinking skills in their pupils.

ABSTRACTOR'S ANALYSIS

The author of this study selected three important questioning strategies around which he tailored and executed the study. All three strategies had been identified and thoroughly researched by previous investigators in numerous studies. The results of these investigations, as well as their implications on teacher training programs, have been duly recorded and reported in the literature. Mindful of these original contributions to the research dimensions of education, it must be said that this study contributed little to advance novel constructs to the growing body of teaching theory.

Perhaps the greater contribution of the study resided in its methodology. It demonstrated and confirmed the importance of targeting instruction to develop competency in a skill-related area.

Comparison data obtained from pretest and posttest indices validated the effectiveness of the study's instructional model. The study attained its objectives with regard to the subject group receiving that training. Nonetheless, in assessing the overall research design, little can be said about the study's effect on the non-instructional subject group. No change in their questioning performance was observed. It follows that no instruction produces no learning. Therefore, no subsequent change in their behavior could be expected. One questions why such treatment groups, even as controls, are built into pilot studies testing models of instruction. In no way does their performance support or negate the efficacy of the model under investigation. It merely adds weight but no substance to the conduct of the inquiry.

As one control measure, the section of each audiotape analyzed for performance data was taken from the first 10 minutes of the 20-minute lessons taught by all subjects. The author reported the subjects used the SCIS program and materials to plan and execute their teaching lessons. Therefore, let us assume the lessons taught were of an inquiry nature. The author never fully spelled out the instructional format used in the lessons being taught. Where did the lessons fit into the context of a unit presentation? How much inquiry by children can take place in a 20-minute lesson? Does the first 10-minute segment of a lesson, so often routine, best lend itself to the asking of higher cognitive questions or the lengthening of the wait-time period? In teaching an elementary school science lesson, irrespective of format, the second 10-minute segment might have yielded even more significant performance data regarding the three variables under investigation.

Was the instructional format of lesson 6 controlled? Again, that aspect was not suitably described for the reader. Lesson format would certainly affect a teacher's questioning strategies depending upon whether the lesson was introductory, demonstrative, investigative, or summative in nature.

The previous questions notwithstanding, the report was clearly written. It is becoming more widely recognized by teacher educators that effective teacher questioning behavior doesn't happen by chance. The author makes the cogent point that teacher training programs need to focus on the teaching of strategies that will raise the level of questioning competency in both preservice and inservice teachers. He suggests that the science methods course is a good place to begin.

A plethora of research documenting the need and validating various instructional models to accomplish that singular objective are presently available. What is needed is not more research in the area, but rather a greater awareness of its substance on the part of teacher educators everywhere. Ultimately, the implementation of training programs to upgrade teacher competency in asking questions reside with them.

Tamir, Pinchas; Vincent N. Lunetta; and Robert E. Yager. "Science Teacher Education: An Assessment Inventory." Science Education, 62(1): 85-94, 1978.

Descriptors--*Educational Assessment; *Educational Research; *Evaluation; Instrumentation; *Measurement Instruments; Science Education; *Science Teachers; Secondary School Science; *Teacher Education.

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Walter S. Smith, University of Kansas.

Purpose

This article presents a lengthy eight-page inventory, the "Science Teacher Education Inventory" (STEI), which teacher educators can use for self assessment of their preservice secondary science teacher education program. Results can be used for describing and revising programs but not for grading specific programs nor for justifying any particular approach to preparing science teachers.

Rationale

The STEI was developed in order to more fully describe the following nine principles of an ideal secondary science teacher education program identified by Yager, Lunetta and Tamir (in press).

(1) Experiences in teacher education are planned for a span of several years and are integrated with the total academic program.

(2) The program consists of a broad curriculum that goes beyond the separate science disciplines.

(3) The nature of science in a historical, philosophical, and social perspective is a central component.

(4) The program is based upon stated objectives, generally expressed in performance terms that delineate a variety of instructional skills and competencies.

(5) Experiences for improving communication and interpersonal relationships are included.

(6) Preservice teachers are actively involved in a variety of teaching experiences; a significant number of these occur with students in the public schools.

(7) Experiences are provided in evaluation and in the application of research to learning and teaching.

(8) The preservice program is but a first step in a continuous cycle of professional growth and inservice education.

(9) The program is based upon a continuing evaluation of needs and program effectiveness; it includes the continuous assessment of the skills of individual preservice teachers.

Research Design and Procedure

Since the procedure for developing the STEI is not described in the article, it can only be assumed that a rational procedure of literature review and consultation with experts was employed to identify inventory items.

Findings

The STEI is provided in the article, but there is no reporting of its use in any setting. For example, no measures of reliability and validity are included.

ABSTRACTOR'S ANALYSIS

This article functions to publish an inventory which secondary science teacher educators can use to describe their own programs within the framework of the nine principles identified by Yager, Lunetta and Tamir (in press). Thus, the article is similar to program standards published by NCATE (1977) or AAAS (1971), in that it provides a checklist of criteria, but no means except "professional judgment" of measuring the extent to which the criteria are met. The inventory can be used only to describe a program and to stimulate thinking about program revision. As the authors rightfully point out, the inventory cannot be used in any way for empirically evaluating any specific program either against other programs or a model program, for neither reliability, validity, nor normative data are presented for the instrument.

In the STEI the article provides a starting point--the instrument, but the instrument's usefulness is severely limited by lack of information about the procedure for the instrument's development (e.g., how were these nine areas and the accompanying items selected?) and its use (e.g., is it reliable?).

REFERENCES

American Association for the Advancement of Science. Guidelines and Standards for the Education of Secondary School Teachers of Science and Mathematics. Washington, D.C., 1971.

National Council for Accreditation of Teacher Education. Standards for Accreditation of Teacher Education. Washington, D.C., 1977.

Yager, R. E.; V. N. Lunetta; and P. Tamir. "Trends in Science Education During the Past Decade." School Science and Mathematics (in press).

Willson, Victor L. and Antoine M. Garibaldi. "The Association Between Teacher Participation in NSF Institutes and Student Achievement." Journal of Research in Science Teaching, 13(5): 431-439, 1976.
Descriptors--*Achievement; Educational Research; Inservice Education; *Inservice Teacher Education; Science Education; *Science Institutes; *Secondary School Science; *Teacher Education; Teacher Improvement

Expanded Abstract and Analysis Prepared Especially for I.S.E. by George G. Mallinson, Western Michigan University.

Purpose

The problem of the study reported here was to seek answers to the question, "Is there any evidence that precollege student cognitive achievement has been increased because of teacher participation in NSF-sponsored institutes?" The problem statement should have used the term "NSF-supported" rather than "NSF-sponsored" since the NSF, at least verbally, disclaims sponsorship, only indicating support.

Rationale

The authors indicated that from 1958 until the time of the study the National Science Foundation had spent nearly $750 million for teacher training and upgrading. It was claimed, and there is no reason to question the claim, that most of the money has been used to support various training institutes, such as the Academic Year Institutes (AYI), In-Service Institutes (ISI), Sequential Summer Institutes (SSI), and Unitary Summer Institutes (USI). One goal of these NSF precollege programs was to improve the scientific literacy of American school children through the improvement of teachers' knowledge. The improvement involved two facets: (1) increased scientific knowledge and (2) the improved capability to impart that knowledge to students.

Although not stated directly in terms of specific hypotheses, one may infer from a statement by the investigators that some increments in improvement were evident. The statement is, "Some evidence has been provided to show that the ultimate consumers, school children, have benefited" from teacher participation in institutes. A second statement, "This report provides some additional evidence of benefit," may be interpreted as a hypothesis that replication, or extension, of previous studies would support the evidence in the earlier statement alleging such benefits. Consequently, one may assume that the rationale for this study was that (1) evidence of benefit was available but inconclusive, but (2) further study would provide additional supporting evidence. So the effort was oriented toward investigating any relationships that might exist between teacher participation in NSF institutes and student achievement.

Research Design and Procedure

Obviously, an experimental design involving the random assignment of teachers for participation or nonparticipation in NSF-supported institutes and the random assignment of students to teachers followed by a comparison of the achievements of students of the two groups of teachers would have been optimal. However, the manner in which institutes have been supported and the way teachers have been selected made such a design impossible. Thus, it was necessary to make such comparisons on data available after the teachers had participated or not participated and the students were tested. Obviously, this type of post hoc design involved a number of variables that could not be controlled and which may have affected any comparisons. The source of the data was not categorically stated in the report, but a telephone conference with the first investigator supported the inference that it was generated in a study by Gullickson and Welch (1972). The study was part of an independent evaluation of five NSF-funded Comprehensive Teacher Training Projects in which schools were systematically sampled on the basis of factors such as urban-rural, geographic region, junior and senior high school. The geographic regions were centered in Wyoming, South Dakota and Mississippi for science, and in California and Indiana for mathematics.

Within each school in the sample the principal selected one teacher at random from the science or mathematics faculty and each teacher selected one of his/her classes at random. The teachers and students were then administered a series of questionnaires including background questions for the teachers and the National Teacher Examinations in Physics-Chemistry-Science or Mathematics (1970). The science students were then administered the Test of Achievement in Science (TAS) (1972) consisting of 40 items selected from the National Assessment of Educational Progress in science. The mathematics students were administered the Mathematics Achievement Test (MAT) consisting of 40 items from the National Longitudinal Study of Mathematical Abilities (NLSMA) pool in mathematics (1972). Two forms of each science and mathematics test were developed, one for junior-high-school students (8th grade) and one for senior-high-school students (11th grade).

A total of 346 teachers and classes participated in science and 211 teachers and classes in mathematics. It should be noted that not all students took the achievement tests. The teachers were instructed to randomly divide the students into groups, some of whom took the achievement tests whereas others took attitude and process tests. The numbers in the latter two groups were not indicated. Consequently, the class means used in the analyses were estimates.

Data concerning the reliabilities of the tests appear in Table 1.

Table 1

Reliabilities of Tests

Test	N	Reliability
NTE (Mathematics)	Not available	.94
NTE (Science)	Not available	.90
TAS		
Senior High	1921	.87 (KR-20)
Junior High	981	.87 (KR-20)
TAM		
Senior High	1261	.86 (KR-20)
Junior High	1424	.92 (KR-20)

Analyses and Findings

Three analyses were made:

1. The first analysis dealt with the possible differential assignment of institute attendees to high-ability classes. This analysis tested the independence of participation in NSF institutes from the ability group of the class to which the achievement tests were administered. The ability groups were based on assessments of the teachers in four categories: "high ability," "average ability," "low ability," and "mixed ability." Participation was defined as cumulative participation in AYI, ISI, SSI, USI programs. Cumulative participation was placed in three categories, "No"-43 percent of teachers; "Low"-one or two institutes-29 percent of teachers; and "high"-three to fourteen institutes-28 percent of teachers. The chi-square statistic was used for the analyses with each test and none were significant at the 5 percent level. Thus, it was concluded that assignment of teachers to classes of different levels of ability was not related to institute participation.

2. A second analysis using the chi-square statistic was made to test the independence of the type of class in senior high science (biology, chemistry, and physics) to which the teacher was assigned from extent of institute participation. As with level of class ability, the chi-square statistic involving the type of class taught was not found to be significant at the 5 percent level. Thus, assignments to teach the three science course in question was deemed to be independent from the NSF institute participation.

3. The third analysis, namely the relationship between institute participation and student achievement to institute participation, was most germane to the main problem of the study. The first step was to partition possible factors in student achievement by means of analysis of variance. This statistical treatment indicated that one fixed factor was regional effect. The second fixed factor

considered was cumulative participation in NSF institutes. In order to equate participants and nonparticipants in NSF institutes statistically, teacher achievement as measured by the NTE, was included in the analysis as a covariate. This was based on the assumption that achievement was a criterion for the selection of teachers for attending NSF institutes. Consequently, the analyses of covariance were computed with two fixed crossed factors, (1) geographical region with Mississippi, South Dakota and Wyoming for science and California and Indiana for mathematics, and (2) level of NSF institute participation classified as "No," "Low," and "High" for four sets of data. These four sets were the scores obtained on the tests for science and mathematics at both the junior- and senior-high-school levels. However, disproportionate cell sizes were encountered for the science data and so the design was made proportional to aid interpretation. This was accomplished by randomly eliminating subjects in oversize cells and, in the case of the Mississippi junior-high-school science participating group, including cell means as scores for three dummy subjects. Disproportionality was not so great for mathematics so no adjustments were made.

According to the investigators, the findings indicate that the marginal means of student achievement for NSF participation showed a consistent trend toward better student performance with increased teacher participation in NSF institutes for all four tests in science and mathematics. However, a significant effect was not evident for the covariate, namely, teacher achievement on the NTE on student achievement. Nevertheless, it was concluded from two planned orthogonal contrasts that students of teachers at the high-school level who had high institute attendance performed significantly better than those who attended only one or two institutes. This significant relationship was not evident at the junior-high-school level.

Interpretations

From the analyses and findings, the investigators indicated that although the study was not a true experiment, the consistent results, coupled with similar findings by other researchers, were sufficient to conclude that a real institute effect was present. The significance of results at the high-school level and nonsignificance at the junior-high-school level was interpreted as being a function of the subject matter taught and population of students between the two levels. It was implied that the high-school courses are more likely to be elective than those in the junior-high-school and consequently, those who take science in high school may be a higher ability group. In final, it was recommended that science and mathematics teachers continue to attend workshops, institutes and inservice courses.

ABSTRACTOR'S ANALYSIS

One can hardly question the merits of the purpose of the study. The vast amounts of money spent for teacher education certainly warrant scrutiny of the cost benefits. The support of such scrutiny may be inferred from a statement by Howard J. Hausman, who was intimately involved in NSF funding of science and mathematics education, in his article entitled, "Influence of Funding by the United States Government on the Teaching of Science in the Elementary and Secondary Schools" (1979). He states:

> Exactly how institutes actually benefited teachers, and in what ways these benefits might be measured, are exceedingly difficult questions. Many studies of institutes appear in the literature (the ERIC system catalogs such studies), offering varied approaches of different investigators to the problem. Objective data are hard to obtain, and measures of student improvement attributable to institute attendance are equivocal. Many subjective reports by participants and observers have formed the basis for positive conclusions on the effectiveness of the institute mechanism. For the most part, attempts to assess the impact of institutes on science teaching have been forced to fall back on self-reporting by teachers, or observations by more-or-less experts. Even when a listing of topics or details of topic treatment can be obtained from participating teachers as an index of institute effect, relationships to student achievement are attenuated and in the end the investigator's value judgment on strengthening the course must be imposed.

Despite these assertions concerning the lack of "hard data," Hausman concludes that "an inquiry into the impact of NSF's institutes programs on secondary schools must start with the vast scope and duration of the effort, together with the great preponderance of favorable reactions among participants and other parts of the educational structure." In brief, positive benefits are indicated.

There have been a number of concerns about benefits of institutes and other programs supported by the NSF. Some of these are documented by the abstractor in an article he prepared entitled, "Some New Perspectives on an Issue in Doubt" in the 1979 AETS Yearbook. But, despite the legitimate concerns that were raised, he stated, "Without regard for the ultimate merit of these Institute Programs, no other single activity has ever had a greater impact on American Education." The comment, however, was analogous to those of Hausman in that it was not based on hard data. Thus, the study by Willson and Garibaldi is timely and saluatory in that it was an effort, among many subjective and quasi-objective reports, to provide some hard data.

The investigators recognized that the optimal type of experimental design could not be implemented because of the nature of the mechanism for funding NSF institutes and selecting participants. Consequently, they selected what appeared to be the best pool of <u>post hoc</u> data available and in general, used defensible, standard statistical techniques to process them.

There are, however, certain questions with respect to the source of the data, and the techniques for handling them that need examination.

1. The source of the data used in the study was not stated categorically and consequently it was necessary to contact the senior author to verify what might be inferred as the source.
2. The data used by Willson and Garibaldi were the same as that used by Gullickson and Welch from five regions of the United States. However, there is some question as to how representative

the sample was, considering only two major urban centers were involved. The extent to which findings based on data from these sources could be generalizable to the entire population of teachers and students of science and mathematics is a matter of doubt.

3. It was indicated that the teachers and students who were subjects were selected by random techniques. However, with 81 percent of schools sampled for science and 91 percent of those for mathematics in small towns and cities under 50,000 population, the sample could hardly be considered parametric. It follows, therefore, that a random selection of teachers and classes from what may be a nonparametric sample could hardly result in parametric groups.

4. The disproportionality and unequal variances of the science data and the procedures used to compensate for these factors are matters of debate. The techniques used making the design proportional are at best specious unless some adequate defense is provided. There is no indication of how many students were removed to reduce oversized cells and the creation of a larger size cell by adding cell means, raising an N of 4 to an N of 7 needs a statistical justification. One may wonder whether such procedures really support the two planned orthogonal contrasts.

5. Considering the number of teachers who have participated in NSF Institutes, the Ns of the cells are difficult to view as other than small. This further raises the question of generalizability of effects.

6. There is one major anomaly that appears in the study. One analysis of covariance leads to the conclusion that the teacher's science ability is not apparently related to student achievement. Yet, another set of analyses seems to indicate that cumulative participation in NSF institutes has an effect on student achievement. Does this mean that NSF programs do not increase the scientific knowledge of teachers, one of the goals of these NSF-supported programs? The only conclusion that seems reasonable is

that they acquire greater ability to impart knowledge to students as a result of participation, despite the fact that their knowledge is not increased.

7. A final comment involving the four analyses of covariance with teacher achievement as the covariate seems worthy. Although the investigators conclude that student achievement is a function of cumulative NSF participation and differences are statistically significant, the class means are estimates since not all students took the achievement tests. Also visual inspection indicates that differences among the class means, even if significant, are hardly consequential.

In summary, it is difficult to assume that the investigators have contributed consequential "hard data" that support the benefits of institutes. The data that are provided contribute some evidence to what is available, but the "bottom line" is still illusory.

REFERENCES

Gullickson, A. R. and W. W. Welch. "Applying Experimental Designs to Large-Scale Program Evaluation." Research Paper No. 16. Minneapolis: Minnesota Research and Evaluation Project, University of Minnesota, June 15, 1972.

Hausman, Howard G. "Influence of Funding by the United States Government on the Teaching of Science in the Elementary and Secondary Schools." Pp. 214-238 in Abraham, Michael R. and Fred W. Fox (eds.) Science Education/Society: A Guide to Interaction and Influence, 1979 AETS Yearbook. Columbus: ERIC Clearinghouse for Science, Mathematics, and Environmental Education, The Ohio State University, November 1978.

Mallinson, George G. "Some New Perspectives on an Issue in Doubt." Pp. 180-189 in Steiner, Robert L. and Fred W. Fox (eds.) Science Education: Past or Prologue, 1978 AETS Yearbook. Columbus: ERIC Information Analysis Center for Science, Mathematics, and Environmental Education, The Ohio State University, November 1977.

INSTRUCTION

Bowyer, Jane B. and Marcia C. Linn. "Effectiveness of the Science Curriculum Improvement Study in Teaching Scientific Literacy," Journal of Research in Science Teaching, 15(3): 209-219, 1978.

Descriptors--Curriculum; Educational Research; Elementary Education; *Elementary School Science; *Elementary School Students; *Instruction; *Program Evaluation; Science Course Improvement Project; Science Education; *Scientific Literacy; *Test Development; Tests

Expanded abstract and analysis prepared especially for I.S.E. by Gerald Skoog, Texas Tech University.

Purpose

This study sought to identify the relationship between the curriculum goals of the Science Curriculum Improvement Study (SCIS) and student achievement in terms of knowledge of basic concepts in science and an understanding of the nature of science. Also, the study included an assessment of gender differences in the development of scientific literacy.

Rationale

An important question for the developers and users of any curriculum project is whether the goals and objectives of the project are achieved through the instructional procedures and patterns and students experience. SCIS, the curriculum project involved in this study, is a six-year elementary science curriculum with the goal of developing scientific literacy in terms of knowledge of basic concepts in science and an understanding of the nature of science.

Previous studies had measured the impact of individual SCIS units on the student's ability in areas such as conservation tasks, serial ordering, and classification. Data resulting from investigations of the impact of the total SCIS curriculum on student achievement as related to the goals of the curriculum have been almost nonexistent.

Data on differences in achievement of science among sexes are needed to help explain the low level of participation of females in the study of science at the high school, college, and professional level. This study was designed to secure such data.

Research Design and Procedure

A scientific literacy test was developed and administered to 312 sixth-graders in two rural Michigan schools that had used the SCIS program for six years. The six-year SCIS program had been studied in its entirety by 196 or 62 percent of the 312 students in the experimental group. The 219 sixth-grade students in the control group attended nearby schools, spent 1.5 hours a week on science as had the students in SCIS, and were closely matched to the experimental group in terms of ability level, age, and socioeconomic status.

Nine evaluation tasks specifically designed for this study and tested in a pilot study were used to examine the children's thinking regarding the basic processes of science and their understanding of major content-oriented concepts. The problems posed by the tasks could not be explained satisfactorily with a memorized answer or guess. The validity of the test tasks was determined by matching the objectives of the curriculum to the test instrument. The reliability of the test tasks was determined through the testing of 24 children who had varied experiences with SCIS. The correlation coefficients ranged from .85 to .97.

A standardized set of directions was developed for administering the test tasks, answering student questions, and scoring the tests. Scorer reliability was determined to be at the 92 percent level of agreement.

Multivariate analysis of variance was used to analyze the data. The nine task scores were treated as dependent variables; gender and SCIS experiences were the independent variables.

Finding

Overall, the experimental group performed better on the test tasks than did the control group ($F = 13.7$, $df = 9$ and 519, $p < .001$). Most of the difference in performance was in five tasks. Students with experience in SCIS did better than the control group in two of the four process or skill-oriented tasks (variables and analyzing experiments) and three of five content-oriented tasks (relative position, energy transfer, solution and evaporation). The direction of significance favored the control group in the process skill area of histograms.

Gender was not a factor in the performance of students in the experimental group which had studied SCIS ($F = 1.2$, $df = 9$ and 519, $p < .28$). However, the combined results of the females in the control and experimental groups together indicated the females outperformed the males on the test tasks ($F = 4.07$, $df = 9$, and 519, $p < .001$). Three test tasks dealing with variables, analyzing experiments, and histograms accounted for most of the gender differences. These were the most verbally demanding test items administered and the differences may have resulted because of superior verbal skills rather than superior scientific reasoning.

Interpretations

The data supported the conclusion that students involved in the SCIS program assimilated some fundamental concepts of science that contributed to the development of scientific literacy. Also, the data suggested that the direct science experiences the students had in SCIS as they worked through the learning cycle of exploration, invention, and discovery were effective for learning both concepts and processes of science.

Curricula or programs involved in comparison studies seldom have similar goals. Student ability in naming variables and analyzing experiments were objectives for both the developers of the SCIS program and the authors of the science textbooks used by control groups. SCIS students outperformed the control students in these areas.

Noncognitive factors may be responsible for the low level of participation of females in the study of science at the high school, college, and professional level inasmuch as the data in this study indicated that there were no differences in the ability of elementary aged males and females to learn science concepts. Differences in tasks requiring scientific reasoning may have been the result of the verbal requirements of the test.

Interpretation of the data also indicated that SCIS students did poorly on the test tasks concerned with energy transfer despite much study in the area. SCIS students were very successful with the test tasks concerned with solution and evaporation. Finally, the performance of SCIS students on the tasks concerned with analyzing experiments and relative position indicated that the curriculum effects in these two areas of logical thinking can be detected for at least two years after being taught.

Inasmuch as the students involved in this study are from a small rural area, the results are limited in terms of the population to whom they can be generalized.

ABSTRACTOR'S ANALYSIS

This was a conventional research study where two groups of students were matched and tested to determine whether two individual variables, SCIS experience and gender, influenced scores on a battery of evaluation tasks. The goals of the study were defensible. Appropriate research methods and procedures were used. The conclusions did not go beyond the data and the interpretations were reasonable. The data presented should be useful to those involved with SCIS as well as to other science educators, curriculum specialists, and learning theorists.

The students in the experimental group of this study were from two elementary schools where the teachers in all 19 classrooms taught science 1.5 hours a week and 196 or 62 percent of the students had completed the six-year sequence of SCIS units. The authors also reported the teachers were using SCIS with enthusiasm. In an era when the science curriculum projects of the 1960s are declining in use, less money and time are being spent on science instruction in elementary schools, and textbooks and content learning dominate, information on the assumptions, forces, and factors that contributed to the adoption and continued use of SCIS by all the teachers in these two schools might be more important to know than the difference in student achievement. Today, it is painfully obvious that curriculum specialists in science education have not given enough consideration to the social, political, and educational considerations involved in achieving improvements in schooling.

Scientific literacy and gender differences in science achievement, which were the focus of this research, are two areas where additional research and dialogue must occur. What skills and knowledge are necessary for literacy in science? Despite some similar objectives, the science material studied by the experimental and control students in this study represented different views of what is required for literacy in science. The authors concluded that students in the experimental group assimilated "some fundamental concepts of science which contribute to the development of scientific literacy." However, were the students in the experimental group, who outperformed the control group in five of nine test tasks, making significant progress toward attaining literacy in science? Were they more literate in science than the students in the control group? Or, was their literacy just different and of another dimension? Overall, this research stimulates questions concerning the dimensions of scientific literacy and how it is achieved. However, the use of the words "Teaching Scientific Literacy" in the title may be misleading or at least presumptuous.

This study provided useful data concerning gender differences in science achievement. It is obvious that many questions remain to be answered as researchers study how political, social, and educational forces shape decisions made both by males and females as they choose among interests, courses, and careers.

Hall, J. R. "A Study of the Teaching of Elementary Chemistry." Journal of Research in Science Teaching, 13 (6): 499-507, 1976.
Descriptors--*Chemical Reactions; *Chemistry; Educational Research; Elementary Education; *Elementary School Science; Science Education; *Teaching Methods; *Instruction

Expanded abstract and analysis prepared especially for I.S.E. by Ann C. Howe, Syracuse University.

Purpose

The purpose was to compare two methods of introducing the idea of chemical change to students beginning the study of chemistry. The methods compared were (a) the inductive method advocated in the Nuffield O-level Chemistry, a curriculum widely used in Great Britain, and (b) a method of study employing a "growth model" developed by the author.

Rationale

The Nuffield Chemistry curriculum was developed in 1966 with the support of the Nuffield Foundation, much as the various science programs in this country were developed in the 1960's with the support of the National Science Foundation. The Nuffield Chemistry curriculum advocates the use of an inductive, discovery method based on student practical work. The student is provided with examples and non-examples and is expected to attain a given concept with very little guidance.

The author had argued in a previous paper (Hall, 1971) that this approach to the idea of chemical change was mistaken on the grounds that the idea of chemical change was a higher-order concept derivable from "logico mathematical" rather than "physical" experience and "the inductive process was appropriate only for lower-order concepts derivable from directly perceived characteristics of exemplars" (p. 499). He proposed an alternate approach, which he called the "growth model," based on the idea of continuous elaboration and development of a concept. In this view, a concept is not attained at one point in time when the parts of a puzzle fall into place but is built up through a sequence of experiences that stimulate gradual growth of understanding.

It was hypothesized that a teaching method based on the growth model would lead to (a) a better grasp of the concept of chemical reactions, (b) improved integration of ideas, and (c) a better understanding of particular reactions.

Research Design and Procedure

Two parallel classes of approximately 30 upper level 11- and 12-year old students were taught a series of 70-minute lessons, one per week, for 11 weeks. The same teacher taught both classes. The groups had approximately equal numbers of boys and girls who scored very close together on an intelligence test. The classes were probably as well matched as is ever possible in a normal school setting. The total number of students was 62.

Class A received carefully planned lessons based on the growth model; Class B received the Nuffield materials as designed.

After the 11 weeks of instruction students were given an oral test to measure grasp of the chemical reaction concept and written tests to measure recall of information and learned principles. An observer also kept anecdotal records. It was expected that Class A would (a) give more conservation responses in the oral test, (b) show better recall of facts and principles and (c) give less evidence of lack of integrative thinking.

An oral test of conservation was given to a "representative" sample of 18 students from each class. Each student was asked questions about three experiments and was categorized as conserver, non-conserver or neither on mass, identity, and composition.

Short written group tests were used to assess recall of facts and principles.

Non-parametric statistical methods were used to test for significance of differences.between classes.

Findings

There were no significant differences between classes in the number of students giving conservation responses to the questions of the oral test. There were, however, differences in the reasons given for these responses; in two cases the chi-square test showed a significant difference between classes.

A somewhat similar result was obtained on the written tests. The t-test showed a significant difference in only one of the three group tests, but further probing showed significant differences among classes on specific items.

Evidence for lack of integration was sought by examination of both the oral and written tests for instances of acceptance of contradictory statements or failure to apply learned principles. Both classes provided examples of lack of integration but somewhat more cases (significantly more in one instance) were found among Class B (Nuffield) students.

Interpretations

Results are interpreted as lending no support to the hypothesis that Class A students would give more conservation responses but providing some evidence in favor of the other two hypotheses.

Reconsideration of the difference between the instructional methods used in the two classes prompted the author to suggest that the contrast in terms of growth and inductive models might not be as appropriate as a comparison in terms of what each series of lessons appeared to say to students. He reinterpreted the results by assuming that Class A received a series of statements about combination (decomposition) and

Class B received a series of statements about reversible/irreversible change. Viewed in this way, responses of both classes could be interpreted in terms of a growth model and the appropriate use of the two schemes in different circumstances.

The author concluded that the growth model could provide a useful key in understanding the development of a complex concept and that the results supported his original hypotheses that the growth model provides a better way of introducing basic ideas than the inductive approach.

ABSTRACTOR'S ANALYSIS

Carefully executed, small-scale, in-depth studies of this kind have been reported more frequently from researchers in Britain than from those in this country. The chief value of this type of study, as the author indicated, lies in the insight provided into teaching and learning processes. Concomitant with this is the danger that reported results may be unjustifiably generalized.

The small total number of students involved in this study and the even smaller number (36) of those who received the oral test do not allow generalization of specific results. (Since only two classes were used there are some purists who would argue that the proper N is 2, but we will ignore that argument.) The study is nevertheless suggestive and instructive in several ways.

First, it is a study that could have been (though this was not) carried out by a classroom teacher. If more studies of this kind were conducted and replicated in classrooms, the results would lead to better, more thoughtful instruction and a gradual accumulation of generalizable knowledge about teaching and learning. College and university-based science educators might have more influence in their field if they more often sought ways to aid teachers who wish to carry out well planned studies of simple design in their own classrooms.

A second way in which this study makes a useful contribution is in the author's insight into the false dichotomy of the original conceptualization of the two teaching methods as inductive and non-inductive ("growth model"). Perhaps because the study was small and not intended to produce reliable quantitative results, the investigator could afford to rethink his original assumptions about the differences in teaching methods. Those who spend time observing in classrooms are aware that the act of teaching is often quite different from what was in the mind of another person who wrote out the lesson plan, and that the distinctions we make between inductive and deductive or between discovery and didactic do not always come through that way to students.

Although the crucial aspects of the lessons may not be what the teacher or researcher intended, interpretations of outcomes can only be valid if they are made in relation to what came through to students rather than to what the teacher or curriculum designer had in mind.

A third useful contribution of the study is the finding that many of the students in the classes studied failed to integrate the knowledge gained from the series of lessons. Since the students who participated would have been at sixth and seventh grade levels in the United States, this result will not be surprising to many science educators. There is a growing awareness that complex, higher-order concepts are often introduced without sufficient preparation and before students have the mental maturity to bring together separate facts and ideas to form an integrated whole. The "growth model" that the author advocates would lead students from one level of understanding to another in such a way that the concept to be learned would be expanded and elaborated through a gradual developmental process. More work is needed to translate this idea into classroom practice.

Much more work is also needed to take the insights gained from modest studies of this kind and use them to plan and execute more ambitious studies that may lead to deeper understanding of, and generalizable knowledge about, teaching and learning sciences.

Herron, J. D.; E. Agbebi; L. Cottrell; and T. W. Sills. "Concept Formation as a Function of Institutional Procedure or: What Results from Ineffective Teaching." Science Education, 60(3): 375-388, 1976.

Descriptors--*Chemistry; *Concept Formation; College Science; Educational Research; *Higher Education; *Instruction; Learning Processes; Remedial Programs; *Science Education; *Teaching Methods.

Expanded abstract and analysis prepared especially for I.S.E. by Richard J. Bady, Mt. Senario College.

Purpose

The study was designed to confirm experimentally what the authors had noticed previously: that different methods of teaching a concept may not only result in a better or poorer learning of the concept, but may result in the learning of different concepts. Specifically the authors contrasted the concept learned from a verbal definition with the concept learned by presentation of examples and counter examples.

Rationale

The study was carried out in the framework of what the authors feel is that of most evaluative research; that is, a concept is taught to two different groups, using two different methods, and the effectiveness of the two methods is compared by using a posttest. What the authors wish to demonstrate is that such an approach often reveals only quantitative differences and that qualitative differences may be more important. For instance, if the analysis is inadequate, it may be that the groups learned different concepts.

"Concept" as used in this study refers to the traditional psychological use of that term: a finite collection of attributes which must be present for the concept to be appropriate.

Research Design and Procedure

The subjects were students in a remedial college chemistry class. They were randomly assigned to one of three treatment groups. Data were collected and used for 42 students in each treatment group. Subjects were tested two days after receiving about ten minutes of instruction.

Treatment groups. Subjects in treatment group A were given a verbal definition of "mib," the concept to be learned. The definition was: "A mib is a right triangle with a segment perpendicular to the shortest side." A single example of a mib was shown, then students were asked to draw 26 examples of mibs.

Students in treatment group B were given the same verbal definition, but no example. They were presented with 26 figures and asked to circle those that were mibs.

Students in treatment group C were not given a definition, but rather a classic concept formation exercise in which 26 figures were presented, each followed by feedback as to whether the figure was or was not a mib.

Posttest. All three groups took the same test two days later. Twenty figures were presented. Subjects were instructed to circle those figures that were mibs. In addition they were asked to write a definition of mib.

Test results were analyzed in three ways. First, performance of the three groups on each item of the test were compared. Second, the definitions given by the three groups were compared, both in terms of their correctness and in terms of the consistency between the definition each subject gave and the way he/she applied it in the test items. Third, the three groups were compared as to how well they had learned the concept. Implicit in the experimental design is the assumption that this third question is ambiguous--that how well a concept is learned is a meaningless question unless it is clear what concept is being learned.

Findings

The rate of success was different for the three groups on 10 of the 20 items (chi-square test, $p < .05$).

A score was assigned to each student's definition and the means of the groups compared using ANOVA. The F ratio was significant ($p < .05$) and the scores for both groups A and B appeared to be greater than that of group C, although post-hoc tests were not reported.

Scores for consistency between definition and use were derived. Again, ANOVA revealed differences among the three groups. No post-hoc tests are reported but it appears that group C was most consistent in applying their definitions.

Using the definition the authors initially intended as the "correct" answer, group C was found to be poorer than A and B in terms of number of "correct" responses.

Interpretations

The authors conclude that while it might be tempting to conclude that one teaching method is superior to another, their analysis demonstrates that, in this case, <u>different</u> concepts were learned. Students in group C included some attributes that were not intended by the authors, but which are quite reasonable in the actual materials used. For example, the perpendicular segment of each "mib" always pointed outward, although this was not specified in the authors' "correct" definition. Thus what the "correct" answer was was not clear. The materials were ambiguous.

Since the groups given a verbal definition remembered that definition, but were not consistent in applying it, the authors concluded that knowledge of a verbal definition is not evidence for understanding it or being able to apply it.

ABSTRACTOR'S ANALYSIS

The article demonstrates clearly that the analysis of a concept and the materials designed to teach it is a crucial but difficult part of any evaluation research. The ideas in this article are useful to anyone trying to sort out the difficulties and contradictions in the results of such research. The authors have effectively called into question the assumption that two different teaching methods, such as "learning by discovery" vs. "verbal learning" can be easily and unambiguously compared.

Three criticisms of the study are in order, the first of which is quite ironic. In demonstrating the ambiguity of the idea of the "right concept," the authors say that "Specifically as the concept was presented to group C, it might be inferred that an illustration is a mib only if the perpendicular segment is located near the center of the short leg of the triangle and extends outward from the triangle." However, this itself is a typical mistake of logical analysis. All the examples given fit this description, but this does not imply they are necessary for mibhood! No negative examples are given. For example, all the figures are printed in black ink, but we do not conclude that a figure must be black to be a mib.

A minor statistical criticism is that the authors conclude which of the three groups are superior to which on the basis of a significant ANOVA test. Post hoc tests are required for this although inspection of the data indicates that their conclusions would probably be verified by such tests.

Finally, a criticism that should be mentioned is about the ease with which the authors generalize from this experiment, in which the concept is fairly explicit (but still difficult) to concepts of science in general. Even in their example the concept analysis is problematic. It is quite possible that the complex concepts we must teach do not fit the simplistic model of concepts used in this study.

This study teaches an important lesson to those who would perform a simple experiment to show that teaching method A is better than method B. An analysis of the concepts used is not just a necessary prerequisite to such a study, it should be a goal of the study as well. Understanding concepts from the learner's point of view requires continuous revision of our own ideas. Concept analysis before and as part of a study can aid in in this.

Mathis, Philip M. and John W. Shrum. "The Effect of Kinetic Structure on Achievement and Total Attendance Time in Audio-Tutorial Biology." Journal of Research in Science Teaching, 14(2): 105-115, 1977.

Descriptors--*Academic Achievement; *Biology; College Science; *Higher Education; *Instruction; Science Education; *Teaching Methods

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Joel J. Mintzes, University of North Carolina at Wilmington.

Purpose

This study examined the effects of verbal communication kinetic structure and verbal ability on cognitive achievement and total attendance time in college-level, audio-tutorial biology. Six hypotheses were tested:

H_1: that students receiving high kinetic-structure (HKS) audio-taped communications would achieve more than students receiving low kinetic-structure (LKS) communications.

H_2: that high verbal-ability students would achieve more than middle verbal-ability students and that the latter would achieve more than low verbal-ability students.

H_3: "...that the degree of kinetic structure would become an increasingly important determinant of achievement as student verbal-ability decreased." (interaction)

H_4: that students receiving HKS communications would spend more total time in the audio-tutorial center (ATC) than students receiving LKS communications.

H_5: that high verbal-ability students would spend less time in the ATC than middle verbal-ability students but more time than low verbal-ability students.

17

H_6: "...that high verbal-ability subjects receiving LKS communications would spend more total time in the ATC than high verbal-ability subjects receiving HKS communications." (interaction)

Rationale

Postlethwait's (1972) audio-tutorial approach was used as a vehicle to study several predictions based largely on O. Roger Anderson's kinetic structure theory of verbal learning (1969, 1971). Although the study was not the first to examine relationships between communication structure and achievement, this effort added to existing knowledge by: (a) using college-age subjects, rather than junior high-age subjects; (b) employing kinetic structure levels that approximate those found in actual science lessons; (c) integrating visual and other illustrative material into the audio communications, and (d) examining previously unstudied subject matter content (cell structure and function and movement of materials).

Hypotheses which postulated the effects of kinetic structure on time spent in the audio-tutorial center were rooted in Waetjen's (1965) work on motivation and cognitive dissonance.

Research Design and Procedure

One hundred twenty-five (125) students who enrolled in a first year, introductory biology course at Gainesville (Georgia) Junior College served as subjects. Data shrinkage resulted in analyzable information for 101 subjects.

The experimental design was similar to Campbell and Stanley's Design 4 (1963):

$$\begin{array}{cccc} R & O_1 & X_1 & O_2 \\ R & O_3 & X_2 & O_4 \end{array}$$

in which students were assigned randomly to two treatment groups—low kinetic structure audiotapes (X_1) and high kinetic structure tapes (X_2). Multiple choice pretests (O_1 and O_3) and posttests (O_2 and O_4) were administered and data on verbal ability (SAT-V scores) and time spent in the audio-tutorial center were obtained.

A 2 x 3 factorial design was used to examine the effects of kinetic structure (2 levels) and verbal ability (3 levels) on achievement and total attendance time. Two separate univariate analyses were performed. Effects on achievement were examined by covariant analysis using total time spent in the audio-tutorial center as the covariate. Effects on total time spent were studied by analysis of variance.

The two levels of kinetic structure were established by varying the sequence of "discourse units" within the audio-taped presentations. These tapes were then integrated into the normal course routine. The entire treatment period consisted of a two-week segment during which established course procedures were followed.

Findings

Findings of the study can be summarized as follows:

(1) Differences in cognitive achievement ($p < .05$) were found among treatments favoring students receiving high kinetic structure communications.

(2) Differences in cognitive achievement ($p < .01$) were found among verbal-ability groups. Post hoc analysis (Duncan's Multiple Range) showed the high verbal-ability group mean to be different from the low verbal-ability group mean ($p < .05$).

(3) Interactions between levels of kinetic structure and verbal ability in cognitive achievement were not detected.

(4) No differences were found in total attendance time among treatment groups nor were differences found in total attendance time among verbal-ability groups.

(5) Interactions between treatment and verbal-ability in total attendance time were not detected.

Interpretations

The authors conclude that their findings vis-a-vis the effect of kinetic structure on achievement support previous work by Anderson and others. Their failure to find significant interactions between kinetic structure and verbal ability in cognitive achievement was interpreted as evidence that, "...the advantage of HKS communications over LKS communications is not restricted to subjects within a narrow range of verbal ability."

Failure to find differences in total attendance time by students receiving HKS and LKS communications was not reconciled by cognitive dissonance theory. Additionally, the nonsignificant differences in attendance time attributed to verbal-ability groups was seen as evidence, "...that the theoretical advantage of the A-T instructional approach may not be fully realized in actual practice."

Finally, the authors suggest two major implications for the teaching of college-level biology:

(1) that kinetic structure analysis might be a good way of organizing and ensuring the effectiveness of topics or units of study presented in A-T biology courses,

(2) (in apparent contradiction to previous statements) "...that the flexible scheduling arrangement commonly employed in A-T courses may be an effective way of minimizing individual differences in achievement." (based on large variability in attendance time within verbal-ability

and significant correlation between time spent and achievement).

ABSTRACTOR'S ANALYSIS

The audio-tutorial (AT) approach has been used in college biology teaching for almost two decades. During this period a significant number of studies have examined the effectiveness of this instructional strategy and the results seem to verify early claims that the method is at least as effective as conventional approaches and in many cases more effective (Fisher, 1976). In addition, attitudinal studies invariably show that students like courses that use the AT format. Second generation studies have shown significant relationships between biographical, intellective and personality variables and achievement in AT courses (Mintzes, 1975).

The present effort is one of a number of third generation studies which uses the AT technology as a vehicle for examining the general predictive validity of cognitive learning paradigms. The emphasis has shifted, therefore, from the teaching method itself to questions of a more fundamental nature. Many science educators would agree that the trend toward theory-based research marks the emergence of a maturing discipline. Some, including Novak (1977), have recognized the usefulness of the AT approach (with its emphasis on sequencing and integration) as a powerful tool in understanding variables affecting human learning.

It is worthy of at least passing note that both the Cognitive Assimilation Theory advocated by Ausubel (1968) and Novak (1977) and the Kinetic Structure Theory advocated by Anderson (1969, 1970) are concerned largely with the sequencing of instructional materials. Ausubel's theory addresses the sequencing of conceptual "information bits" while Anderson is concerned with the sequencing of verbal "discourse units." Apparently the tightly programmed nature of the AT format enables researchers to manipulate these potentially important sequencing variables thereby studying their effects on learners with differing cognitive styles.

Anderson's work has had a significant impact on research in science education over the past few years. Much of this work has been done with adolescent and elementary school-age children. Perhaps the most important contribution of the present study is that it extends the work on verbal kinetic structure to college-age subjects.

Although the experimental design of the present study was basically sound, several questions concerning instrumentation and data analysis persist. For example, the KR-20 reliability of the cognitive posttest was calculated and reported (0.84); however, no information on content validity was provided. How was the validity determined? What levels of cognitive difficulty were assayed? These are important questions when one is attempting to determine the effects of some independent variable (such as verbal kinetic structure) on cognitive achievement.

A second question concerns the choice of analytic procedures. The analysis of covariance technique was used to examine the effects of kinetic structure (and verbal ability) on achievement. A pretest (items randomly selected from posttest) was administered. The low scores and (presumably) nonsignificant differences between treatment groups "...indicated that at the outset of the experiment, subjects possessed little knowledge relative to the topics considered..." Normally in this situation one proceeds to an analysis of variance on posttest data. However, in this study, the authors chose to use the "amount of time spent in the ATC" as the covariate in an analysis of covariance. Although there may be good theoretical reasons for choosing time as a covariate, the experimental design seems to require that an ANOVA be done first.

The choice of time spent in the ATC as a covariate in the first analysis is somewhat confounded by its use as a dependent variable in the second analysis. What further mystifies the reader is why time was chosen as a covariate in the first analysis when no differences were found among treatment groups with respect to time in the second analysis.

The generalizability of the findings might further be questioned in terms of the length of the treatment period. The entire experimental treatment consisted of a two-week segment of an on-going course. One wonders whether the results reliably reflect those which might be found in a typical semester course. This weakness was readily acknowledged by the authors.

Despite these procedural difficulties the study was well thought out. The research was appropriately presented within the context of previous work and the written report was clear and concise. And finally, the apparent contradictions concerning the effectiveness of the A-T strategy for minimizing individual differences in achievement were at least partially resolved.

REFERENCES

Anderson, O. R. Structure in Teaching: Theory and Analysis. New York: Teachers College Press, 1969.

Anderson, O. R. Quantitative Analysis of Structure in Teaching. New York: Teachers College Press, 1971.

Ausubel, D. P. Educational Psychology, A Cognitive View. New York: Holt, Rinehart and Winston, 1968.

Campbell, D. T. and J. C. Stanley. "Experimental and Quasi-experimental Designs for Research on Teaching," in Handbook of Research on Teaching, N. L. Gage (Ed.). New York: Rand McNally, 1963.

Fisher, K. M. "A-T Science Teaching: How Effective Is It?" BioScience, 26(11): 691-697, 1976.

Mintzes, J. "The A-T Approach 14 Years Later—A Review of Recent Research." Journal of College Science Teaching, 4(4): 247-252, 1975.

Novak, J. D. A Theory of Education. Ithaca, New York: Cornell University Press, 1977.

Postlethwait, S. N.; J. D. Novak; and H. T. Murray. The Audio-Tutorial Approach to Learning Through Independent Study and Integrated Experience, 3rd ed. Minneapolis: Burgess Publishing Co., 1972.

Waetjen, W. B. "Learning and Motivation: Implications for the Teaching of Science." The Science Teacher, 32: 22-26, 1965.

Mintzes, Joel J.; David L. Littlefield; David P. Shaub; Richard Crockett; Robert W. Rakitan; and Ronald Crockett. "Studies on Individualized Instruction: Biology." School Science and Mathematics, 76(8): 675-686, December 1976.

Descriptors--*Achievement; *Biology; Biological Sciences; *College Science; *Educational Research; Higher Education; *Individualized Instruction; *Prediction; Science Education; Secondary Education; *Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by Robert E. Yager, University of Iowa.

Purpose

Five separate studies are reported. The reports represent a coordinated effort to investigate relationships among various student characteristics and learning outcomes in five attempts with individualizing biology teaching at the high school, junior college, and university levels. The studies focused on the identification of potential self-learners in individualized biology. A wide range of student characteristics was studied as possible predictors of student achievement.

Rationale

Although the virtues of individualizing instruction have been extolled in many quarters, the investigators state that empirical studies designed to validate these claims have been few. Furthermore, they report that most studies which have been performed employ the typical comparative methods paradigm and add little to our knowledge of students or the teaching-learning process.

Like many phrases in educational jargon, the term "individualized instruction" has been abused by teachers and professional educators alike. The abuse of the term stems, in part, from a proliferation of diverse programs each claiming to provide "individualized" instruction. The authors define an individualized program as one which possesses at least four of the following five characteristics.

1) Students are permitted and encouraged to proceed through instructional material at a pace commensurate with their interests and abilities.

2) Heavy emphasis is placed on self-instructional approaches.

3) Students select learning activities which prepare them to master a set of instructional objectives.

4) Students determine the amount of time devoted to studying material and how this time will be allocated.

5) The role of the teacher is that of "advisor" and "facilitator" rather than "information-giver."

One of the distinguishing features of most individualized programs is that they are purported to tailor instruction to the personal needs, abilities, and aptitudes of each student, thereby permitting students to progress through conceptual material at their own pace. By shifting the responsibility for completion of coursework from the teacher to the student, these programs are said to de-emphasize the importance of intelligence differentials and place greater stress on the amount of time the student allots himself for study.

Experience with many individualized programs has shown, however, that some students have difficulty adjusting to the freedom afforded by these self-instructional teaching modes. As a result, overall achievement is often depressed because of procrastination or lack of self-discipline on the part of students.

Because of these and other problems, it is becoming apparent to many educators that presentday "individualized" programs may be appropriate instructional approaches for some students but inadequate for many others. If teachers had a reliable way of identifying potential self-learners at an early stage, they could provide more time to students in need of personal attention either in groups or on an individual basis.

Research Design and Procedure

Following is a listing of the five programs studied and a listing of distinctive features of the individualized program there:

Program and Its Location	Distinctive Features
Individualized Learning (IL) Glenbrook North High School (Northbrook, Illinois)	Learning concepts Audio-equipped carrels Testing Center Open Laboratory Optional teacher-presentations
Open Laboratory (OL) Lake Forest High School (Lake Forest, Illinois)	Unscheduled lab experiences Large lectures
Computer-Assisted Instruction (CAI) Kennedy-King College (Chicago, Illinois)	CAI terminals Open Laboratory
Mastery Learning Mode (MM) Wright College (Chicago, Illinois)	Behavioral objectives Alternative diagnostic exams
Audio-Tutorial Modules (AT) University of Illinois (Chicago Circle Campus)	Audio-equipped carrels Open Laboratory Behavioral objectives Large-group lectures Small-group discussions

Data were collected on the following student characteristics with the instrument indicated for each of the five programs.

Summary of Data on Student Characteristics

Student Characteristics	Mode of Assessment	IL	OL	CAI	MM	AT
			Individualized Biology Programs			
critical thinking ability	Watson-Glaser Critical Thinking Appraisal (WGCTA)	WGCTA	WGCTA	—	—	—
understandings about science	Test On Understanding Science (TOUS)	TOUS	TOUS	TOUS	TOUS	TOUS
biological knowledge	Nelson Biology Test (NBT)	—	NBT	NBT	NBT	NBT
attitude toward science	Scientific Attitude Inventory (SAI) Likert-type item—locally constructed inventory (SA)	SAI	SAI	SAI	SAI SA	SAI
scholastic aptitude	Iowa Tests of Educational Development (SchA)	—	SchA	—	—	—
personality factors	16 Personality Factor Questionnaire (16PF) High School Personality Questionnaire (HSPQ)	HSPQ	—	16PF	—	16PF
intelligence	Otis-Lennon Test of General Intelligence (INT) Black Intelligence Test of Cultural Homogeneity (BITCH)	—	INT	BITCH	—	—
motivational factors	School Motivation Analysis Test (SMAT) expected course grade (CGE) aspired educational level (AEL)	SMAT	CGE AEL	CGE	CGE AEL	CGE
socioeconomic level	parents' educational level (FEL-father) (MEL-mother) approximate family income (FI)	—	FEL MEL	FEL MEL	FEL MEL	FI
others	locally-constructed inventories	—	sex age # older siblings # science courses	sex age major GPA number siblings	sex age major GPA # older siblings # younger siblings # science courses	sex age major GPA H.S. math average # college math crses H.S. type # siblings

The sample size and the analytic procedures used in the studies are summarized as follows:

Program	Number of Subjects: Individualized	Number of Subjects: Conventional	Analytic Procedure
IL	306	--	multiple discriminant function
OL	48	48	linear discriminant function
CAI	54	59	multiple linear regression
MM	58	38	linear discriminant function
AT	40	41	canonical variate analysis

Findings

The analytic procedures indicated previously were employed. All characteristics which predicted, discriminated, or correlated at the .05 significance level or better were defined as significant characteristics. The results of the studies are summarized as follows:

Program	Significant Characteristics: Individualized	Significant Characteristics: Conventional
IL	critical thinking ability (WGCTA) "obedience" (SHPQ-E) science attitude (SAI) school attitude (SMAT-9) superego (SMAT-8)	
OL	biological knowledge (NBT) socioeconomic level (FEL) scholastic aptitude (SchA) # younger siblings	biological knowledge (NBT) school motivation (AEL) # younger siblings sex
CAI	understandings about science (TOUS) age major cultural homogeneity (BITCH) "humility" (16PF-E)	school motivation (CGE) "creativity" (16PF-M) # siblings "intelligence" (16PF-B) major
MM	biological knowledge (NBT) # younger siblings understandings about science (TOUS) science attitude (SAI and SA)	science attitude (SA) biological knowledge (NBT) school motivation (AEL) understandings about science (TOUS)
AT	biological knowledge (NBT) understandings about science (TOUS) science attitude (SAI) "intelligence" (16PF-B) "sobriety" (16PF-F) major GPA # college math courses	"intelligence" (16PF-B) "creativity" (16PF-M) major # college math courses H.S. type H.S. math average sex

Interpretations

The authors report that the results of these studies are important from both a practical and a theoretical standpoint. In terms of the application of these findings for the improvement of educational practice, they provide a set of rough guidelines for teachers interested in the assignment of students to optimal teaching modes. From a theoretical standpoint, the findings help identify several clusters of variables that are related to student performance in both individualized and conventional biology programs. These clusters include: measures of prior knowledge in science/biology, measures of intellectual ability, and measures of school/science motivation.

In every analysis except one (in which the factor was not studied), some measure of prior knowledge in general science or biology was shown to be related to cognitive performance. These measures were pre-treatment scores of the Nelson Biology Test and the Test On Understanding Science as well as the students' college major (science/non-science). The authors see this finding as supportive of the cognitive learning paradigm of David Ausubel which holds that single most important factor in determining student learning is "what the learner already knows."

A second recurring factor in several of the analyses related to the intellectual ability of the student. This cluster includes scores on the Watson-Glaser Critical Thinking Appraisal and the "intelligence" subscale of the 16 Personality Factor Questionnaire. In addition, scores on the Black Intelligence Test of Cultural Homogeneity were shown to be related to knowledge acquisition in one analysis. This is related to the work of Novak (at Cornell) regarding the importance of "analytic ability" in knowledge acquisition. The authors do caution the reader regarding the relationships among the terms analytic ability, critical thinking ability, and intelligence.

A third cluster concerned with the scientific interest and general school motivation of students was found to be related to performance in both individualized and conventional biology programs. This cluster includes scores on the Scientific Attitude Inventory and the School Motivation Analysis Test as well as measures of educational aspirations and course grade expectations. The authors see once again the relativity of this finding to the Ausubel-Novak paradigm which alludes to the importance of affective factors in cognitive learning.

ABSTRACTOR'S ANALYSIS

Relationship to Matrix of Other Studies. Although the authors see the study as supporting the cognitive learning paradigm of Ausubel as "refined and empirically tested" by Novak and his co-workers at Cornell, they do not specify any real relationship with specific findings. They are content with stating that their results are "supportive" in three general areas, namely prior knowledge, intellectual ability, and motivation. There is no real matrix of studies identified except by name (i.e., Ausubel/Novak).

The authors do not clearly state a problem nor do they review specific research related to a problem. The "other studies" are identified only as a facet of the discussion of results.

New Conceptual Contributions. The article does not provide any new conceptual contribution. Instead the report tends to offer a refinement of the concept of individualization and multiple ways of reviewing its effect upon learners. Its stress upon the multifaceted nature of learning outcomes and the varied effect of given instructional strategies upon individual learners are two important contributions of the study(ies).

New Methodological Contributions. The idea that results are more generalizable if a given problem is approached from a variety of perspectives and settings is a good one. The use of five settings, three academic levels, and some variation in the concept of and approach to individualization represents a kind of methodological contribution. Although this is not a new experimental method, the varied design and the "set" of controlled experiments is not frequently found in educational research literature.

Validity of the Study. The descriptions of the individual studies include many value judgments. Indeed there is little rationale provided, and at times, little to see for making some of the observations and measurements for assessing student outcomes in the five attempts at individualization. Although there is some control used within the separate studies, it is not always apparent. The differences in course, in level, in instruction, in methodology, in measurement from program to program suggests problems for which inadequate information is provided for the reader.

Comments on Research Design. Although individualization is defined, it is a loose definition. Further, the authors do not carry the concept beyond an approach to content which teachers have organized, conceived, and defined. The use of varied measures, analyses, and designs for the separate experiments are not discussed and/or explained. Too little information and/or rationale is provided to the reader for making a serious assessment on experimental design.

Adequacy of Written Report. The written report is clear while being much too general in a research sense. It should be emphasized, however, that the journal in which the study is reported serves primarily a teacher audience. Hence there is justification for brevity and general descriptions of the experiments conducted. It is unfortunate, however, that many value judgments are made without reference and without qualification. For example, biological content is a given and individual approach seems to emphasize timing, instructional objections, and programming. It is assumed that some students do not respond well to such "individualization." In one sense no learner ever learns unless it is as an individual. Is there such a thing as "group learning"?

Assessment of Current State of Research. Essentially there is no assessment of current research--either with respect to individualization and/or the Ausubel/Novak hypothesis concerning learning. There is an attempt to relate the studies which are reported to each--in a very global sense. The article could be improved as a research contribution if the setting had been better described, if current research had been analyzed, if more precise questions had been asked, if more rationale for choice of certain test instruments, of some discussion and/or interpretation of "significant" characteristics had been provided.

Suggestions of Future Research Directions and Efforts. Certainly more information is needed upon the use of given individualized approaches to teaching or learning. However, experiments designed for seeking more information need to be designed carefully and specifically. Although general studies are of value in pointing out directions and need for other experiments, this report may be too general and too loose to be of real value. More work is needed in defining a rationale for biology teaching--at various academic levels. Once a program is established, information can be sought concerning the relative effectiveness of approaches for meeting such objectives.

Useful research in the general area of individualization strategies must be more controlled. What is meant by open laboratory, the use of CAI, learning contracts, audio-equipped carrels, teaching centers, alternative diagnostic exams? What do given instruments measure? In what situations? Since it is widely known that students vary in interests, in studying/learning styles, in motivation, and in many other ways, can a general study such as reported here result in useful information? If the results support a learning paradigm such as the Ausubel/Novak one, cannot readers expect researchers to analyze and exemplify such support more specifically?

Although a series of related experiments can be a desirable direction for research in science education, enough control, description, similarity in terms of content, approach, and measurement are needed if useful results are to be expected.

Smith, Albert B.; Gray Ward; and Joseph S. Rosenshein. "Improving Instruction by Measuring Teacher Discussion Skills." American Journal of Physics, 45: 83-87, January 1977.

Descriptors--College Science; *Educational Research; *Higher Education; Instruction; *Interaction Process Analysis; *Physics; Questioning Techniques; Science Education; *Teacher Improvement

Expanded abstract and analysis prepared especially for I.S.E. by Herbert A. Smith, Colorado State University.

Purpose

The study was designed to determine the effectiveness of interventions to improve the questioning skills of graduate teaching assistants in lower division college physics. More specifically, an inservice training program for teaching assistants was designed to modify their behavior with respect to questioning techniques and to promote the use of higher order questions by the assistants. Statistical tests were applied to measure the effectiveness of the procedures.

Rationale

The basic concept of the study is that student achievement can be improved by improving the questioning skills of the teacher. The study builds upon the considerable body of research related to taxonomic description of objectives and formalized systems for observation and classification of teacher classroom behavior. It is assumed that higher quality teacher performance will be reflected in higher quality student achievement.

Research Design and Procedure

The plan of the study required a pre- and post-videotaping of graduate teaching assistants and a comparative analysis of the two performances of each teacher. The focus was on the questioning skills of teachers in posing higher order questions and on student responses to such questions. Between the pre- and post-videotaping sessions, an inservice program was conducted over a five-week period and included such activities as lecture, discussion, peer observation by another teaching assistant and interaction analysis including use of instruments by Flanders and Gallagher. Eighteen TA's who were assigned to the open laboratory of the undergraduate general physics course (900 students enrolled) were involved in the study. The study reported appears to have extended over one term at the University of Florida.

Findings

Teaching assistants did ask more questions and more higher order questions during the post-training period in which they were videotaped. Student responses were also found to be on a higher level.

Interpretations

The authors concluded that the systematic use of observational systems could be used to improve college instruction. They found most teaching assistants responded favorably and desired more opportunity to study the tapes as a means of self-improvement.

ABSTRACTOR'S ANALYSIS

As previously indicated, the study builds on the considerable research and literature relating to teacher behavior and its analysis drawing on the work of Bloom, Flanagan and others.

The study can be described as a beginning effort and, it is hoped, will be followed up by more sophisticated and better conceptualized efforts. Although the reviewer believes the direction of the research is sound and represents a field with considerable potential for research and development, he considers the study to be seriously flawed in several respects. Some of the more important points include the following:

a) It is grossly misleading to report an increase of 92 percent in higher order questions asked by teachers and a 71 percent increase in higher order responses of students. A more legitimate comparison would appear to be the relative proportion of higher order questions in the total questions asked in the pre- and post-evaluations. These ratios were .54 in the pre-evaluation and .59 in the post-evaluation for teacher's questions for an increase of 5 percent. Comparable student responses were .60 and .61. The authors appear to recognize this defect but still reported the misleading percentages.

b) The change in number and/or proportion of higher order questions and responses may not necessarily be the important dimension. In sound teaching it may sometimes be necessary to raise a number of factual questions to be certain that a student has a command of the basics, is on the right track or in need of redirection. It is the quality of the questions and responses, and perhaps neither the total number or proportion of higher order questions, that is important. This aspect is not addressed.

c) We are not informed as to the experience and training level of the teaching assistants. Were they first semester graduate students or advanced doctoral candidates?

d) Only 10 of 18 teaching assistants were included in the final analysis because of "poor tape quality." This is unfortunate and raises questions of sampling, care with which the study was monitored and similar questions. The integrity of the study depends on the skill, insight and judgment of the individual(s) making a classification of questions and responses in the analysis of the videotapes. It seems unfortunate that the analysis was entrusted to a graduate assistant, however, competent, rather than being done by one or more of the principal investigators.

e) No effort was made to assess the actual effects of the questioning interventions on the overall achievement of students. Increased level of student responses (very slight on a proportionality basis) does not necessarily indicate higher achievement in general college physics. Only evidence of better achievement in the objectives for the course can provide the necessary validity criterion of the success of the experimental procedures.

The investigators point out some additional limitations in their study which are well taken.