DOCUMENT RESUME

ED 103 274 SE 018 790

AUTHOR	Helgeson, Stanley L., Ed.; Blosser, Patricia E., Ed.
TITLE	Investigations in Science Education, Volume 1 Number 1. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION	Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	74
NOTE	70p.
AVAILABLE FROM	Ohio State University, Center for Science and Mathematics Education, 244 Arps Hall, Columbus, Ohio 43210 (Subscription, $6.00, $1.75 single copy)
EDRS PRICE	MF-$0.76 HC-$3.32 PLUS POSTAGE
DESCRIPTORS	*Abstracts; *Educational Research; Research; Researchers; *Research Methodology; *Research Skills; *Science Education

ABSTRACT

This first issue of INVESTIGATIONS IN SCIENCE EDUCATION (ISE) is designed to provide a new perspective for viewing research articles and is intended to aid in the improvement of writing research reports. The analyses are intended to provide useful comments and suggestions and to serve as a device which might be used for training in the writing of research articles. Articles included in ISE are selected primarily from such sources as professional journals and reports of government-funded projects. All critical abstracts published in ISE are prepared by science educators. Abstracts included in this issue relate to topic areas such as effectiveness of various teaching strategies, learning theory and processes, socio-economic influences on scientific concept development in children, and verbal interaction in science classes. The format includes bibliographical data, purpose, rationale, research design and procedure, findings and interpretations as well as detailed notes offered by the abstractor. (Author/EB)

Vol. 1, No. 1 Autumn 1974

INVESTIGATIONS IN SCIENCE EDUCATION

Expanded Abstracts and Critical Analyses of Recent Research

National Association for Research in Science Teaching
ERIC Science, Mathematics, and Environmental Education Information Analysis Center
The Center for Science and Mathematics Education
The Ohio State University

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC Science, Mathematics, and Environmental
Education Information Analysis Center

Published Quarterly by

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

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 25¢ for Canadian mailings and 50¢ for foreign mailings.

NOTES . . .

from the Editor

This first issue of Investigations in Science Education marks the beginning of another cooperative effort by the National Association for Research in Science Teaching and the ERIC Science Mathematics, and Environmental Education Information Analysis Center.

Investigations in Science Education (ISE) is designed to provide a new perspective for viewing research articles and is intended to aid in the improvement of writing research reports. Expanded abstracts and analyses will be developed for each research article reviewed in Investigations in Science Education. This analysis is intended to provide useful comments and suggestions for the preparation of research reports and to serve as a device which might be used for training in the writing of research articles.

Articles to be included in ISE will be selected primarily from such sources as professional journals and reports of government-funded projects. Included articles which have been cited in Current Index to Journals in Education will carry the ERIC accession numbers in the form EJ 000 000; those articles which have been cited in Research in Education will carry accession numbers in the form ED 000 000. These numbers allow the reader to locate the original annotations or abstracts that appeared in those publications. In the case of non-journal articles, they also give the reader access to reproductions of many of the papers through the ERIC Document Reproduction Service (EDRS). The availability and prices of reproduction in hardcopy (HC) and microfiche (MF) are listed in the accession information section. For the convenience of the readers, each issue of Investigations in Science Education which contains documents available from EDRS will include an EDRS order blank preprinted with the available documents abstracted in that issue.

All critical abstracts published in Investigations in Science Education are prepared by science educators. The abstractors donate their time and talents so that ISE may be produced at minimal cost to the science education community. Your comments and suggestions for improving this publication will be appreciated.

Stanley L. Helgeson
Editor

Patricia E. Blosser
Associate Editor

EJ 046 111

Babikian, Yeghia, "An Empirical Investigation to Determine the Relative Effectiveness of Discovery, Laboratory, and Expository Methods of Teaching Science Concepts." Journal of Research in Science Teaching, Vol. 8, No. 3:201-209, 1971.

Descriptors--*Concept Formation, *Instruction, *Teaching Techniques, Discovery Learning, Educational Research, Laboratory Procedures, Scientific Concepts, Secondary School Science

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Gene Gennaro, University of Minnesota.

Purpose

The purpose of this study was to determine the relative effectiveness of discovery, traditional laboratory, and expository methods of teaching eighth grade classes six science concepts having to do with the principle of bouyancy (Archimedes' Principle) as measured by the following criterion learning tasks: overall achievement of students, verbalization, recognition, transfer to new situations, application to numerical problems, discovery, and the retention of concepts. The six science concepts were:

1. The volume of a liquid displaced by an immersed object is equal to the volume of the object.

2. The object loses part of its weight in a liquid.

3. The weight lost by an object in a liquid is equal to the weight of the liquid displaced.

4. If the density of an object is less than the density of a liquid, the object floats on the liquid.

5. The weight of the liquid displaced by a floating object is equal to the weight of the object.

6. The part of a floating object under the surface of a liquid is equal to the volume of the liquid displaced.

The hypotheses tested were that when the six concepts were taught, no significant differences would be observed: (1) among the groups taught by discovery, laboratory, and expository methods; (2) between high I.Q. and low I.Q. students; (3) between the boys and girls; (4) between the effects of the teaching methods upon the achievements of high and low I.Q. groups; (5) between the effects of the teaching methods upon the girls and boys; (6) between the effects of intelligence upon the girls and the boys; and (7) between the interactions of the methods and I.Q. upon the achievement of the girls and the boys.

Rationale

Many science educators advocate that, in the teaching of science in the elementary and secondary schools, concepts be taught by discovery rather than conventional laboratory and traditional expository methods. This experimenter believes that before this methodological approach is universally advocated for all teaching situations, empirical evidence should be provided. The work of Ausubel and others on expository organizers would be relevant to the present study.

Research Design and Procedure

The sample consisted of 216 eighth grade students assigned to three science teachers. The students of each teacher were divided into three homogeneous classes: "fast," "average," and "slow." The three classes assigned to each teacher were instructed by the experimenter during six class periods using one of the three experimental methods: the expository lesson plan (the teaching of the six bouyancy concepts was entirely verbal, except for occasional use of the chalkboard, the students were provided with worksheets); the laboratory lesson plan (each student had a laboratory instructional worksheet and with the aid of laboratory equipment was asked to verify the six bouyancy concepts); the discovery lesson plan (this plan utilized the same equipment as the laboratory lesson plan but provided only the procedure for the discovery of the concept by the student and the elimination of the statement of the concept involved in the experiment.) Students were pretested and posttested. On the basis of the results from the pretest and student files, it was determined that the experimental groups did not differ significantly from one another in I.Q., knowledge of the introductory concepts of weight, volume, and density and their preknowledge of the learning task. A test of 38 items was administered. The results of the posttest were analyzed by a 3-factor randomized design.

Findings

No significant differences were observed in respect to discovery of concepts, retention of concepts, and the interaction of the variables. Significant differences were observed in respect to the following criterion measures: overall achievement, verbalization of the concepts, recognition of the concepts, application of the concepts to numerical problems, recognition of the concepts to numerical problems; and in each of these cases, results favored the expository and the laboratory lesson plans.

Interpretations

1. The expository and the laboratory methods are significantly more effective than the discovery method for teaching science concepts to eighth grade students, in respect to the following criterion measures: overall achievement, verbalization of concepts, recognition of concepts and the application of concepts to numerical problems.

2. Irrespective of the method by which students in the upper I.Q. groups are instructed, they achieve significantly better than those in the lower I.Q. group in respect to the following criterion measures: overall achievement, verbalization of concepts, recognition of concepts, and the application of concepts to numerical problems.

3. Irrespective of the method used to instruct boys, they achieve significantly better than girls in respect to the following criterion measures: overall achievement, recognition of concepts, and transfer of concepts.

Although the superiority of the expository method over the discovery method might have been partly due to the particular expository lesson plan used and to the lack of experience of the subjects in discovery methods, the investigator feels that the investigation provided enough evidence to recognize the importance of expository methods in science education.

Abstractor's Notes

The study reported here is concerned with the study of concepts associated with buoyancy. Further research of teaching methodology needs to be performed in achievement, verbalization, transfer, application, discovery, and retention of concepts in other areas of the physical sciences and in the life and earth sciences as well. Since concepts associated with Archimedes' Principle are difficult concepts for students of all age levels to verbalize and understand, it may be that similar difficult concepts are best taught by expository-laboratory methods and not by discovery methods. On the other hand, it may be that a guided-discovery approach is needed, in which the teacher stops at crucial points and asks students to share pertinent data with one another so that students are getting the necessary data for further parts of the experiment. Or it may call for post-laboratory discussion so that students discuss their data and have a chance to grapple with inducing the relevant concept. If questions arise as to which data are accurate, students can return to the laboratory to check findings. It may be that this approach takes longer but, ultimately, is more efficient for learning.

Some concepts in science are certainly not as difficult for youngsters to understand as are the buoyancy principles associated with Archimedes' Principle. Perhaps these easier concepts are the ones that should be handled in a discovery fashion by the teacher. Students in a chemistry class can certainly inductively arrive at the generalization that soluble bases, acids, and salts conduct electricity by a discovery approach. On the other hand, trying to get students in a biology class to arrive at certain concepts associated with population genetics using a discovery approach might be futile. The former requires good observation, technique, and some ability to generalize from data; the latter may require brilliance plus a great amount of background. Hence, it may be that when one looks at what is needed to understand buoyancy principles that the necessary background is missing; from studies which have been done, it appears that students do not master density concepts even after these concepts have been explained to them and after having worked density numerical problems. Unless these concepts are understood, it would be difficult to teach concepts associated with Archimedes' Principle using a discovery approach.

It may be that major concepts need to be taught in one fashion and subsumed concepts could be taught using a discovery or guided-discovery approach. One would need to determine major concepts and identify subsumed concepts. Students of different ages, ability, and background would have to be tested using various approaches to teaching and then deciding which approach appears to be "best" for which students.

Also, it would be well to compare students several weeks or months after having had exposure to a concept that was taught in expository fashion with students taught using a discovery or guided-discovery approach to see if differences exist. Advocates of the discovery approach have suggested that retention is better using a discovery approach with certain concepts.

It's important that the approach one uses with students in teaching certain science concepts be one that not only allows them to achieve understanding of certain concepts but also makes them like what it is they are learning. The science teacher has them for at most an hour a day, but students' thoughts about a subject go on beyond the hour of class study. If the teacher captures students' interest, the amount of time that students think about certain concepts is increased and hence exposure time to these concepts is, in fact, increased. Our selection of books and articles to read, TV programs to watch are undoubtedly partially based on past favorable recollections.

As suggested by the author, other cognitive and manipulative skills such as critical thinking, ability to design experiments, ability to analyze data, accuracy of measurements, setting up of apparatus need to be measured as well as those tested for in this experimental design. Certainly, as the author also suggests, the lack of prior experience with discovery methods may be a contributing factor.

EJ 088 502

Beisenherz, Paul C., "A Comparison of the Quality and Sequence of Television and Classroom Science Questions With a Proposed Strategy of Science Instruction." Journal of Research in Science Teaching, Vol. 10, No. 4:355-363, 1973.
Descriptors--*Educational Research, *Elementary School Teachers, *Questioning Techniques, *Teacher Behavior, *Television Curriculum, Elementary School Science, Science Education, Teacher Characteristics, Televised Instruction, (Research Reports)

Expanded Abstract and Analysis Prepared Especially for the I.S.f. by William S. LaShier, Jr., University of Kansas.

Purpose

The purpose of this study was to analyze questions asked by studio (TV) teachers and classroom teachers in four school districts in the Seattle, Washington area. The classroom teachers were randomly assigned to four treatment groups representing varying degrees of utilization of TV science and regular school district science.

The first hypothesis contrasted the proportions of questions (modified Gallagher-Aschner system) asked by studio (TV) teachers during the exploration, invention, discovery, and review phases of the TV script. These four phases comprised the Instructional Strategy Category System (ISCS). The data presented dealt only with telescripts used with 3rd and 4th grade classes.

The second hypothesis dealt with significant differences in the proportions of questions in the Gallagher-Aschner categories among the four treatment groups of teachers.

The third hypothesis was concerned with identifying significant differences among the four treatment groups when the data from the Gallagher-Aschner System were viewed in terms of a pre-TV lesson, a post-TV lesson and a non-TV lesson. These three types of lessons comprised the Science Lesson Category System (SLCS).

The fourth hypothesis was somewhat similar to hypothesis one. The three treatment groups utilizing television instruction in grades three and four were contrasted in terms of the proportion of Gallagher-Aschner question categories asked in each of the four phases of the ISCS.

Rationale

The researcher investigated separately the questioning strategies of TV Studio teachers and the questioning strategies of four groups of teachers, each group reflecting differing degrees of utilization of both TV science and non-TV science. The investigator selected the Science Curriculum Improvement Study (SCIS) sequence of exploration, invention, and discovery lessons as the basic strategy within which to examine the questioning styles of the studio teachers and the classroom teachers.

The Gallagher-Aschner System of questioning categories was modified in this study by the addition of an observation category. This system generally reflects the structure of intellect as defined by J. P. Guilford.

Research Design and Procedure

Studio (TV) teachers presented 15-minute weekly lessons to students in grades one through four for 13 weeks. The 54 classroom teachers and their intact classes were randomly assigned to one of the following groups:

I. Television science only, materials for pre- and post-TV activities were provided to this group.

II. Television science and the regular district science program.

III. Television science only.

IV. Regular district science only.

All questions in the 13 television scripts were analyzed using the modified Gallagher-Aschner system. The 54 teachers audio-tape recorded all science taught during the 13 week period. Six lessons were selected from those recorded by each teacher in treatment groups I, II, and III. The sample size of questions for teachers in Treatment IV was based on the mean number of questions asked by teachers in the other three groups.

The proportions of questions derived from the Gallagher-Aschner System were analyzed within the context of the Instructional Category System. Next the proportions of questions were analyzed using the Science Lesson Category System (SLCS).

The high consistency among the coders of the modified Gallagher-Aschner system was indicated by an average percent agreement of 91.8 between each of the three coders on the three tapescripts analyzed at three different intervals. The mean proportions of question types

across grade levels, treatment groups, and instructional contexts were analyzed by the coders. Levels of significance between these proportions were determined by the use of Fattu's Nomograph.

Findings

In hypothesis one, 94 percent of the questions asked by TV teachers occurred during the exploration and discovery phases of the ISCS. Furthermore 78 percent of questions asked in all four phases were convergent questions.

In hypothesis two, 42 percent of the questions asked by teachers in the four treatment groups were included in the convergent category and 39 percent in the memory category. Group II asked a significantly higher proportion of memory questions than groups I, III, and IV considered separately. This corrected statement is based on an apparent error in Table II on the comparison of groups II x IV. The correction should read II > IV. A further correction in the convergent column of Table II should read III > I. The corrected finding of this hypothesis points to Group III asking significantly higher proportions of convergent questions than Groups I, II, or IV considered individually. Group I asked a higher proportion of convergent questions than did either Groups II and IV.

In hypothesis three the Gallagher-Aschner system data from all four treatments were recast into three science lesson categories (SLCS). However, the reader should recognize that only Group II had questions tallied in all three SLCS Categories. The associated Table III perhaps should show only the seven possible combinations rather than the 18 implied comparisons.

The findings in hypothesis three indicated that 80 percent of the TV related questions were asked during the post-TV period. Further scrutiny of this period indicated that Group III asked a higher proportion of convergent questions than either Group I or II. Also, during the post-TV period, significant differences were observed in the proportion of memory questions asked with Groups II > I > III.

In hypothesis four it appeared that 64 percent of the questions asked by classroom teachers in grades 3 and 4 were categorized into the discovery phase. In this phase Group III asked a significantly higher proportion of convergent questions than either other group.

Interpretation

One of the intents of the TV science series was to invent key science concepts and pose convergent questions to be developed in the post-TV series. The author therefore paired Groups I and III post-TV and contrasted their strategy with Groups II and IV non-TV periods.

Apparently on the basis of proportion of convergent and memory questions asked rather than levels of significance, it was concluded that Groups I and III asked more convergent and less memory questions than did Groups II and IV. The implication is that the intent of the TV series was being met.

Abstractor's Notes

The three phase model of SCIS plus the review phase was apparently superimposed on an existing science TV series for grades 3 and 4 rather than the TV series being pre-planned to emphasize one or more of the stages. This limitation gives rise to the question of how the various TV lessons were logically separated into the four phases of the ISCS. Stated differently, could the three coders establish a high consistency in separating an audio-taped lesson into the SCIS categories?

This reviewer also was interested in the ground rules established for the categories of the modified Gallagher-Aschner System. The categories of convergent and observation seemed to be somewhat overlapping.

In future studies, one might wish to re-examine the manner in which the modified Gallagher-Aschner question categories are emphasized in the exploration, invention, discovery and review phases of a science lesson. In a description of a typical TV studio lesson, the author emphasized that convergent questions were used during the exploration and discovery. Both the TV studio teachers and treatment group teachers used only three percent divergent questions. A future questioning model might include a high proportion of divergent questions in the exploration stage, followed by some convergent questions leading up to the invention of the concept and during the discovery phase some initial divergent questions followed by convergent questions.

In the present study the Group III teachers (television science only) asked significantly more convergent questions than did any of the other groups, including Group I that had all the supporting material for pre- and post-TV activities. A follow-up study might search for significant relationships between what the TV studio teacher does and the subsequent questioning behavior of the teachers in a post-TV lesson.

EJ 037 864

Fulton, Harry F., "An Analysis of Student Outcomes Utilizing Two Approaches to Teaching BSCS Biology." Journal of Research in Science Teaching, Vol. 8, No. 1:21-28, 1971.
Descriptor--*Biology, *Evaluation, *Group Instruction, *Individual Instruction, Instruction, Secondary School Science, (Biological Sciences Curriculum Study)

Expanded Abstract and Analysis Prepared Especially for I.S.E. by David R. Stronck, Washington State University.

Purpose

The purpose of this study was to compare the effectiveness of the following two approaches to teaching the content of the BSCS textbook Molecules to Man to students enrolled in eighth-grade biology at University High School at the University of Iowa: (1) the traditional group approach by which every student in the classroom was expected to perform and progress at a pre-determined rate established by the teacher, and (2) an individualized approach by which each student progressed at a rate commensurate with his abilities. The effectiveness of the two approaches was analyzed by comparing the student outcomes in each approach in the following five areas: (1) achievement in biology, (2) understanding in science, (3) critical thinking ability, (4) attitude toward science, and (5) evaluation of the teacher's ability to make the material understandable.

Rationale

This study accepted the validity and reliability of these eight instruments which were used: (1) the BSCS Comprehensive Final Examination, (2) the Nelson Biology Test, (3) the Test on Understanding Science, (4) Facts about Science Test, (5) Watson-Glaser Critical Thinking Appraisal, (6) Silance Attitude Scale, (7) Prouse Subject Preference Survey, and (8) the Performance Scale for High School Biology Teachers. The researcher assumed that these instruments would provide an adequate means for comparing the effectiveness of the two approaches in the five areas listed above.

This study did not depend on any previous research beyond that which was involved in the development of these instruments. The researcher assumed that the BSCS textbook Molecules to Man was equally appropriate to the different approaches. He also assumed that the different use of excursions had no significant effect on the student outcomes. The researcher explained that excursion activities were provided for the students of the individualized approach who had completed the basic core of materials before the end of the school year. For those in the group approach, such options were available during the

school year. The researcher gives no data on the use of these excursions. Nor does he describe the possible Hawthorne effect which is the tendency of subjects in some experiments to respond favorably to almost any change. Those in the group approach used the textbook as the basis for class discussions, testing, and laboratory investigations. On the other hand, those in the individualized approach worked with an adaptation of the textbook which was converted into a loose-leaf notebook.

Research Design and Procedure

The researcher begins by describing his design and procedures as those of the pretest-posttest control group. A sample of twenty students was selected randomly each year from the students enrolled in the eighth-grade biology program at University High School at the University of Iowa. The group approach used in the year 1967-68 may be interpreted as the one providing the control group. The individualized approach of 1968-69 provided the experimental group receiving the treatment of self-pacing in their instruction.

Although random selection of the students is an acceptable method of assuring similar samples, this researcher determined the background of the students by use of the Iowa Cardpac System of Educational Accounting Pupil Inventory and the Iowa Tests of Basic Skills. The Iowa Tests did not reveal any significant differences between the samples in composite scores, vocabulary scores, reading comprehension, work-study skills, or arithmetic skills. The Iowa Cardpac demonstrated very similar backgrounds in the marital status of the parents, occupation and education of the parents, students' attitudes toward studying, teachers, and additional years of education. This analysis of the students' backgrounds establishes the sufficiency of using the "t" test to compare the student outcomes in various areas.

The researcher not only did comparisons by use of the "t" test but also continued his statistical analysis by using the "F" test in an analysis of covariance on all pretest and posttest scores in which the pretest measure was treated as a covariate. In this analysis of covariance, posttest scores were adjusted on the basis of the pretest scores. The adjusted mean scores were used to compare the various scores between the two groups. This use of statistics implies the nonequivalent control group design of research.

Findings

At the five percent level of significance, six null hypotheses were rejected. These rejections consistently demonstrated the greater gain by the students in the individualized class. They were superior in scores on (1) the BSCS Comprehensive Final Examination, (2) the Test on Understanding Science, (3) the Facts about Science Test, (4) the Watson-Glaser Critical Thinking Appraisal, (5) the Prouse

Subject Preference Survey, and (6) the Performance Scale for High School Biology Teachers, Question No. 2. The last scale revealed that the students exposed to different approaches to teaching BSCS Biology did vary in their evaluation of the teacher's ability to make the material understandable.

Interpretations

The researcher simply observed that this study showed a consistently greater gain by the students in the individualized class in all of the areas under consideration. He did not attempt to generalize beyond the one teacher involved in the study and the single course taught at one high school.

Abstractor's Notes

This study uses many well accepted instruments of evaluation. The researcher has thoroughly established his interpretation that students in the individualized class had superior outcomes in comparison with those of the previous year in the "group approach." This abstractor is led to the conclusion that the one teacher involved in the study certainly did a superior performance of teaching the second year with the "individualized approach."

There are at least five possible reasons which might explain the superior outcomes with the individualized approach. The researcher did not discuss any of these possible explanations in his article. First of all, the individualized approach may have been needed simply because the BSCS textbook Molecules to Man (Blue Version) may not be an appropriate textbook for eighth-grade students. Many have described the 1964 edition of the Blue Version as the most difficult of the three BSCS textbooks designed for use by students in the tenth grade. Some have suggested that this textbook should be used only by students who have completed a course in high school chemistry. The difficulty of this textbook for these students may have demanded extraordinary techniques in its use, e.g., the individualized approach.

A second reason for the superior outcomes may have been the excellence of the adaptation of this textbook for the individualized approach. The content of the textbook was placed in a loose-leaf notebook providing space in the reading material for the student to consider any question presented to him in the adaptation. Unfortunately the loose-leaf notebook system was used only with those students in the individualized approach.

A third possible explanation for the superior outcomes of the individualized approach may be the Hawthorne effect. Students in the group approach received a relatively traditional method for instruction in biology. On the other hand, the use of the loose-leaf notebook and

the absence of the usual lecture and discussion methods may have created an atmosphere of great novelty and experimentation. Such changes are certainly capable of producing superior attitudes and outcomes.

A fourth reason for the superiority of the individualized approach may be found in the personality of the teacher. This study considered the work of only one teacher. The Performance Scale for High School Biology Teachers used in this study demonstrated that he made the materials more understandable in the individualized approach. His superior performance may have been because of his inexperience as a teacher during the previous year, or his inadequacy as a public speaker, or his enthusiasm for teaching with individualized materials.

Another possible explanation for the difference in outcomes may have been the quality of the students in the samples. Each year the sample consisted of only twenty students. Although the Iowa Tests of Basic Skills did not show significant differences between the groups, nevertheless, the mean scores revealed a definite trend for superior skills in the individualized group in composite score, vocabulary score, and arithmetic skills. Because of the small size of the samples, this trend may have become significant in the later tests which were used in this study.

At the present time there is a strong emphasis in our nation for the increasing use of individualization in instruction. Unfortunately this study does not provide any strong generalizations which support arguments for increasing individualization. Because only one teacher in one school was involved in this research, the study remains a relatively anecdotal discussion of his improved teaching. His circumstances seem very unique because of the use of Molecules to Man as a textbook for eighth graders. Moreover, University High School at the University of Iowa is probably not a typical American school.

The students exposed to two different approaches to teaching BSCS Biology in this study on the Performance Scale for High School Biology Teachers revealed a significant difference in their judgments on the ability of the teacher to make the material understandable. This difference becomes the most obvious explanation for the superior outcomes of the students in the individualized group. Unfortunately this difference obscures the direct comparison between the group approach to instruction and the individualized approach. Those who are interested in this important comparison can find in the conclusions of this study only that one teacher did a superior job with individualized instruction in comparison with his more traditional teaching of the previous year. If another study revealed that many teachers rather consistently were able to make the materials more understandable when they used an individualized approach, then there would be a strong argument in favor of the use of the individualized approach. This suggested additional study should also involve the use of materials which are truly designed for the specific grade. The materials should also be equivalent in format, e.g., loose-leaf notebooks. Otherwise the lack of sufficient controls may again fail to isolate the impact of individualization.

EJ 045 403

Good, Ronald G., "A Study of the Effects of a 'Student-Structured' Laboratory Approach to Elementary Science Education Methods Courses: Affective Domain." Journal of Research in Science Teaching, Vol. 8, No. 3:255-262, 1971.
Descriptors--*Attitudes, *Changing Attitudes, Educational Attitudes, Methods Courses, Scientific Attitudes, Scientific Enterprise, Student Centered Curriculum, Teacher Education, Tests

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Burton E. Voss, University of Michigan.

Purpose

The purpose of this study was to develop an instrument that could assess college students' beliefs and attitudes toward the objectives of a specific elementary science methods course.

Rationale

The basic assumption underlying the methods course was that a person's beliefs tend to dominate his behaviors. A change in beliefs concerning the nature of science, the nature of scientists, the relationship of science to society, the nature of elementary science and children would therefore give an indication of a corresponding change in subsequent behaviors.

The author further stated that a course permitting students to discover or structure their own learning was conducive to the development of the desirable attitudes he was attempting to measure.

Research Design and Procedures

The assessment instrument was developed by the author. It consisted of 40 items. Twenty were related to the nature of science, the nature of scientists, and the relationship of science to society. Twenty were designed to assess beliefs about the nature of science education for elementary school children, including the role of the teacher.

The instrument was field tested by giving it to 100 prospective elementary teachers. Initial results showed a need for revision. The most notable changes were the reduction in the number of statements related to characteristics of scientists and the role of science in society.

A preliminary version of the test was given to students in the Fall of 1969 and further revisions were made in the instrument. Most of the questions related to the nature of scientists and the relationship of science to society were deleted, thus leaving the orientation of the assessment instrument to questions more directly related to the nature of children and the role of the elementary school science teacher.

Further revisions of the assessment instrument were again made and a final form of the instrument was administered as a pre- and posttest to 109 students in the Winter Quarter, 1970, and to 55 students in the Spring Quarter, 1970. Major efforts were made to reduce the teacher variable in the different sections of the course.

Findings

The student was asked to respond to each statement on the assessment instrument according to a five point scale ranging from (1) strongly agree, (2) agree, (3) no opinion, (4) disagree, to (5) strongly disagree. Analysis of variance was used in determining the level of significance in differences between pre- and posttest scores. There were no "right" or "wrong" responses to the test. There was, however, a preferred pretest-posttest direction of responses to statements if the course was having a "positive" effect.

The individual item analysis testing for significant difference (also indicating change in directionality) showed students made significantly different choices on 25 of the 30 items from pre- to posttest during the Winter Quarter. The results of the Spring Quarter testing were very similar.

Interpretations

The author concluded the student structured course was having an influence on favorable attitude development of students since there was a significant directional gain consistent for both experimental groups on at least nine of the 30 statements on the assessment instrument. The major inference drawn from the study was that actual participation by students in the student structured process seemed to be effective in changing beliefs about the nature of elementary school science, the nature of children and the role of the teacher.

Abstractor's Notes

This article has a great deal of merit in that it attempted to work with an elusive, yet critical aspect of teacher education -- teacher attitude development.

One concern about the study is that statements about science, scientists, and the role of science and society were removed from the assessment instrument with little justification. Teachers' attitudes concerning these aspects are important, but as the author states, "how much can be accomplished in a course 10 weeks in length?"

The abstractor has also worked on a science attitudes test for junior high students. The same technique of analysis was used in interpreting changes in students' attitudes from pre- to posttest. The direction of change on each statement was analyzed. The statisticians associated with the study wanted a student test score on the test. This could have been done, but the individual statement analysis provides a more powerful method for interpretation.

The author could have done a split halves reliability to obtain some indication of the reliability of the assessment instrument.

The main thrust of the study seemed to move from the purpose of developing the assessment instrument to an analysis of student-structured learning situation in an elementary science methods course. Major inferences were made about the apparent effectiveness of the approach without using an appropriate control group or groups. The study did, however, provide a very interesting insight into elementary science teacher education.

EJ 038 719

Harke, Douglas J., "Hierarchical Analysis of the Randomized Multiple-Choice Format." Journal of Research in Science Teaching, Vol. 8, No. 1:29-35, 1971.
Descriptors--*Cognitive Processes, *Evaluation, *Multiple Choice Tests, *Physics, *Test Construction, College Science, Organization, Problem Solving

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Rodney L. Doran, State University of New York at Buffalo.

Purpose

The study compared the performances of students to a free-response item and a set of multiple-choice items on the same science concept.

Rationale

One reason science teachers have been reluctant to use machine-scored, multiple-choice tests is that many questions require several steps to be solved correctly. Partial or total credit is assigned depending on the "degree of correctness of the solution." Nedelsky suggested using several objective test items to test student understanding of a problem instead of the essay format. In this study, Harke compared student responses on multiple-choice items to free-response items testing the same science content.

Research Design and Procedures

Six physics problems were analyzed to determine the sequence of steps necessary to solve the problem correctly. For each of the steps in this hierarchy, a separate multiple-choice item was written. These items when randomly sequenced and administered in a test constitute the Randomized Multiple-Choice (RMC) format. Students complete the free-response format item and use it to aid them in answering the RMC items.

"Direct comparisons, using correlations, did not reveal whether the students used the same mental processes to solve the same problems on the two test formats." Therefore, an indirect approach using two methods of hierarchical analysis was employed to compare the mental processes used by students to solve the problems on the two test formats.

A Consistency Ratio, similar to that used in the Science - A Process Approach evaluation, was calculated for each dependency relationship in each problem. Based on the hierarchy of steps for each problem, the following kinds of transfer relationship exist among the RMC items for a given problem:

1. Mastery of both the upper level item and lower level items.

2. Failure on the upper level item and failure on one or more of the lower level items.

3. Failure on the upper level item but mastery of all of the lower level items.

4. Mastery of the upper level item but failure on one or more of the lower level items.

As only the first three types of transfer relationship are consistent with the hierarchy, the consistency ratio was defined as the number of transfer relationships of type 1, 2, and 3 divided by the total number of transfer relationships. This ratio was calculated for each dependency relationship in each problem. High consistency ratios on all dependency relationships in a hierarchy would support the validity of the hierarchy. As guessing on the RMC items could change transfer relationships from type 2 to type 4 and thus reduce the consistency ratio, Harke used 0.85 as the criterion rather than 0.90 which was used in situations involving free-response questions.

"The consistency ratio method considered group responses as a whole on the complete hierarchy and group responses on segments of the hierarchy. The pattern analysis technique developed by Rimoldi and Grib was used to analyze the responses for the complete hierarchy on a subject by subject basis." Students' response patterns were examined to determine whether they corresponded to one of the "expected patterns" based on the hierarchy of RMC items. These expected and observed response patterns were compared using the Rimoldi and Grib procedure resulting in an index of agreement for each of the six RMC problems.

Findings

The average consistency ratio of four of the six problems was above the established criterion, 0.85. All the consistency ratios in the other two problems were below .85.

The indices of agreement obtained through the pattern analysis were .789, .640, .812, .705, .939, and .819. The lowest indices of agreement were obtained with the problems that had low consistency ratios.

Interpretations

The two methods of analyzing the student responses produced similar results. Harke concluded that these analyses indicated that "students in solving the RMC problems used sequences of mathematical and physical concepts similar to the sequences they would have used if the same problems had been presented in free-response form."

Abstractor's Notes

The fact that the same two problems failed to satisfy both validation techniques suggests that there may have been something in error with the hierarchy or the items developed from the hierarchy. One major question when developing hierarchies is whether to use logical or psychological models. Harke assumed "These sequences can be logically determined from knowledge of physics." A psychological rationale for a hierarchy would entail using a sequence of student behaviors or operations. In reality, many hierarchies may be composed of steps that are sequenced on the basis of both content and student behavior (e.g., student can state rules for volume). Harke's dependence on the logical structure of the content of each problem could have overlooked what the students have to do with the content to proceed along the hierarchy.

Once one has established a hierarchy, the construction of items assessing each unique step can still be a source of error--this is especially true for items that are designed for student behaviors above the recall level. The students' mental process used to answer an item may or may not be the same as the researcher intended. This possible discrepancy could be a function of many variables such as the background of the students, clues with the item, clues from other items ordered to the same problem. In his article, Harke included a sample physics problem with its five RMC items. However, the hierarchy represented by this problem was not included nor was the specific objective measured by each item. Replication of this study would need such kind of information.

The author has obtained evidence to support the use of a measurement procedure that has great potential for science teachers. It is hoped that replication and extensions of this research will be forthcoming to further substantiate the value of this measurement procedure. This RMC technique can become one of the measurement tools all science teachers have at their disposal. Such an analysis of the steps in solving specific science problems would also make teachers more aware of the student behaviors requisite to solving specific problems. This study is one of few research studies that produces something of immediate value to science teachers, in addition to investigating a problem of interest to researchers.

EJ 038 274

Johnson, Paul E., Thomas E. Curran and David L. Cox., "A Model for Knowledge of Concepts in Science." Journal of Research in Science Teaching, Vol. 8, No. 1:91-95, 1971.
Descriptors--*College Science, *Cognitive Processes, Evaluation, Learning Processes, Models, Physics, Scientific Concepts

Expanded Abstract and Analysis Prepared Especially for I.S.E. by H. Craig Sipe, State University of New York at Albany.

Purpose

The purpose of the study was to examine the relationships between interrelated science concepts and the psychological relationships by which these are learned and understood.

Rationale

Neither the logical models of science (e.g. Kuhn, Nagel, or Pearson) nor the psychological models (e.g. association and similarity) have independently been particularly useful in explaining how science is learned and understood. Perhaps the two sets can be used jointly. If so, knowledgeable students will be able to respond to stimulus words in a way characteristic of association, overlearning, mediation, and similarity due to adjacency in structure. The subject matter used in the test exercise is from physics and deals with six relational concepts (power, work, force, acceleration, velocity, and momentum) and three operational concepts (mass, distance, time). Figure 1 shows the interrelationships, e.g. time relates power and work, mass relates force to acceleration. If these relationships exist in the mind of

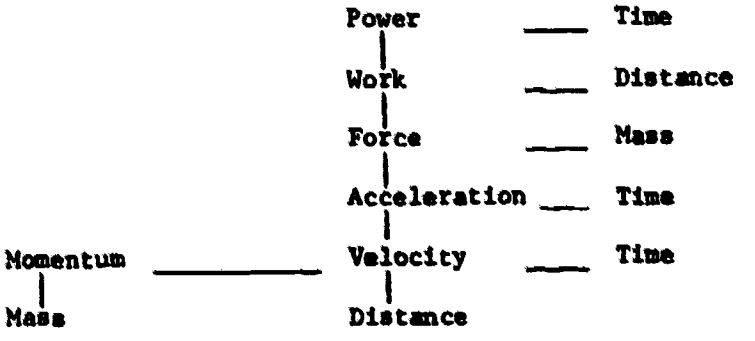

Figure 1. A model for concepts in mechanics

the student, then they can be inferred by the investigator from the responses of the student to stimulus words on a free-association test, a constrained-association test, and a similarity rating test. Johnson

has reported earlier related work in a series of papers in Journal of Educational Psychology (55:84-88, 1964; 56:217-224, 1965; 58:75-83, 1967; and 60:32-40, 1969).

Research Design and Procedure

A three-section test was administered to 59 male physics majors at the University of Minnesota, all of whom had completed at least their junior year. The data from the free-association tests (FA) were reduced and displayed as a 9 x 6 matrix of observed frequencies when subjects were given one minute each to respond to stimulus words (concepts in Figure 1) placed in random order, one to a sheet, in a test package. The constrained association test (CA), untimed, asked the subjects "to make a list of all nouns you would use in defining physics concepts represented by the words," the stimulus words were the same as the FA test. The similarity rating test (SR), also untimed, consisted of 36 possible pairs of nine words (again those in figure 1) to be rated on a seven-point scale of similarity (one for the highly dissimilar to seven for the highly similar). The data from the SR test were reduced to and displayed as "average similarity ratings." To facilitate the comparison of data, (1) rank order correlations were computed between the stimuli words (relational concepts) and the response words on the FA and CA tests and (2) rank order correlations were computed between the "relatedness coefficients" (after Garskoff and Houston, Psych Rev. 70:277-278, 1963) for the pairs of concepts on the FA and CA tests and the average ratings for pairs of concepts on the CR tests.

Findings

1. The associations of stimulus words are consistent with the relationships assumed to be within the conceptual framework (Figure 1). Stimulus words elicit mostly the words adjacent to them as responses. Words at the end of the hierarchy elicit fewer words than those in the middle. Like effects ($P = 0.90$ with $p < .01$) are observed for the six relationally-defined concepts when FA and CA responses are compared.

2. The presumed conceptual network can account for observed similarity relationships reflecting position in the structure ($P = 0.82$ with $p < .01$ for "average similarity ratings,") compared with hypothesized similarity/dissimilarity in the network.

Interpretations

The high agreement between SR test and FA test ($P = 0.85$ with $p < .01$) and between SR test and CA test ($P = 0.79$ with $p < .01$) indicates that "association and rated similarity may be simply alternative ways of indexing the same underlying process." The value of the instruments used depends on the success found in diagnostic evaluation rather than as predictors of success in complex tasks such as problem solving.

Abstractor's Notes

The investigator makes an implicit assumption that instruction is based on verbal association and the subsequent ordering of verbal content. The population used is more typical of those who have succeeded in meeting the demands of a logically structured system than of those who are learning conceptual relationships for the first time. Although no claim is made about the universality of the findings, specialists in instruction need exercise caution in extrapolating the findings to discovery learning modes, to the concrete operational pupils, to task analysis instructional designs, and to situations calling for either application of meaningful verbal learning constructs or of creative problem solving. One must take seriously the investigator's comment on the diagnostic value of the tests and the inherent nature of complex mental tasks.

Three relatively simple procedures are provided to explore hierarchical conceptual relationships. The evidence in the example chosen is reasonably tight. Engaging in further study with less-well defined material in physics and other disciplines would do much to establish the boundary limits for using such promising measurement procedures.

Unresolved is the issue of an appropriate model for developing a knowledge of conceptual interrelationships in science. The investigator chose a situation wherein the constraints on the population assured reasonable conformity to expectation. Little is known about the methods for the initial teaching of the concepts, the main problem that besets the teacher-scholar.

EJ 040 691

Klein, Carol A., "Differences in Science Concepts Held by Children from Three Social-Economic Levels." School Science and Mathematics, Vol. 71, No. 6:550-558, June, 1971.
Descriptors--*Problem Solving, *Scientific Concepts, *Socioeconomic Influences, *Socioeconomic Status, Curriculum Planning, Educational Research, Elementary School Science, Research, Science Education, Student Evaluation

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Alan M. Voelker, Northern Illinois University.

Purpose

This study was designed to determine whether children from three social-economic levels differed in their understanding of selected science concepts and in the methods they would suggest to find answers to questions associated with the concepts.

Rationale

As elementary school science continues to increase in quantity and quality more care must be given to the selection of instructional materials and the use of these materials in facilitating a science curriculum. The study was conducted to verify that the social-economic level of the learner has an effect on what he can derive from the use of specific instructional materials. If social-economic level affects what children can and do learn, provision must be made for maximizing the probability of a productive aptitude-treatment interaction. Using the same materials for all children is not consistent with theory, practice, or research findings.

Research Design and Procedure

Test items were constructed to measure the understanding of concepts selected from the basic text used in the participating school system. The items were piloted in a school system other than the school system where final testing occurred. The children involved were from lower and middle social-economic levels. Data were used to determine the clarity and discriminating power of the items. A prototypic question was:

> Which of these three animals is the grown-up? (picture choices)
> Why did you choose this one?
> If you did not know which was the grown-up, how could you find out?

Pilot study data were also used to determine whether the test format should be oral or written. The items were first administered to 60 third grade children and their responses to the questions were audio-taped. Later, 40 of these 60 children took the same test in written form. Questions on the written form were read aloud to minimize reading difficulties. Administration of the two forms of the test was used as a test-retest procedure for establishing reliability. A correlation of .87 to .90 between the two forms of the test was considered an adequate basis for justifying the use of the written form in the main study.

The population for the main study consisted of the third grade children in fifteen elementary schools in St. Paul, Minnesota. Each school served a single social-economic level and included no racial minority groups. These schools were classified as serving high, middle, or low social-economic levels based on property values, percent of deteriorated houses, average room value, education of parents, occupations of parents, statistics from the Bureau of Health, and ratings supplied by principals and teachers.

All the children used the same science text and had the same amount of science instruction per week. No one level had all new or experienced teachers.

The final sample consisted of 310 fourth grade students in fifteen classes in nine schools.

Test reliability was estimated by use of a Chi-square test of interdependence and Hoyt internal consistency. Significant differences (.05) were found between the three groups on 11 of the 15 items with the high level classes answering all items correctly more often than the other classes. An internal consistency of .76 was obtained.

Findings

A total of five hypotheses were tested.

One-way ANOVA was used to test the null hypothesis of no significant difference between the mean I.Q. scores (The Lorge-Thorndike Intelligence test administered one month prior to the science test). The hypothesis was rejected (.01) and the Student-Newman-Keuls test results indicated that each group mean was significantly different from each other group mean.

The null hypothesis of no significant difference in the mean level of understanding of the concepts by the children from the three groups was also rejected (.01, One-way ANOVA). Each mean was found to be significantly different from each other group mean (Student-Newman-Keuls).

The third hypothesis was of no difference in the mean level of concept understanding when the scores were adjusted for I.Q. differences. This hypothesis was rejected (.01, ANCOVA). The mean for the low social-economic group was significantly lower than that of both the middle and high group but there was no significant difference between the middle and high groups (Student-Newman-Keuls).

Percentages of referral were used to test the "hypothesis" that there would be no difference in the ways students in the three groups would suggest for verifying their answers or seeking answers to the questions about the concepts. Observation and experiment were suggested most often by all three groups and books and newspapers were the second choice for each of the three groups. These two choices were the highest for the high group but there was no regular pattern of decrease from the high to the middle to the low group and the actual percentages varied widely.

A Chi-Square test of independence was used to test the hypothesis of no difference in the number of correct answers given by the three groups to the individual items on the concept test. This hypothesis was rejected (.05) for 13 of the 15 items. The other two questions were answered correctly more often by children in the middle and high groups.

Interpretations

Social-economic level is a factor that influences children's learning of science concepts. This factor needs to be taken into account when planning an elementary school science curriculum and the related instruction. Utilization of the same curriculum materials for all classes in all schools within a school system produced different learning outcomes. This approach is not consistent with the results of research. Children in low social-economic groups appear to be less dependent on observation and experiment for answers to their questions, and their attainment of selected science concepts is less than for children of higher groups where the materials consisted primarily of textbooks. Activities-oriented programs and more selective use of materials were suggested for helping alleviate this problem.

Abstractor's Notes

Sequencing the various procedures employed in the study would have enhanced the readability of the report. It was difficult to determine the exact sequence of events and see how data were used to make decisions about continuing various parts of the study. It was also difficult to determine whether one or two pilot studies were conducted. All comments to be made about the pilot tests, etc. should be presented together and should precede discussion of the main study. Maintaining consistency of language throughout would also help establish the main

points of the report. Concept and understanding are used interchangeably, but there is no explanation that indicates whether they are interchangeable. This should be done in a study of concept learning.

Reference is made to 36 studies which dealt in some way with methods of determining concept development and how children think they get their information for concept formation but they were not cited. How was the output of these studies used in making decisions for setting up this study? How did these former findings influence the construction of a test to account for the nature of concept development in children?

Discussions of test construction and establishing test validity and reliability would have been enhanced by indicating the size of the initial item pool and the procedures and criteria for reducing the pool to 15 items. And can we justify test validity by saying that "the items were based on the science concepts found in the science textbook?" Coefficients of test reliability were reported for the pilot study but not for the main study. Pilot test data indicate whether it is advisable to proceed but they are not adequate for determining whether the final study data have credibility. There are many factors which could completely change the test reliability from pilot to final study.

Including more details about selecting children for testing would make it easier to determine what credibility to assign to the results. For example, provision was made to use I.Q. as a covariate but there is no indication whether other variables which influence concept attainment have been accounted for by random selection of students for testing.

A total of 15 questions was administered to the students in the main study. But there is no indication of whether there is a 1:1 correspondence between questions and concepts. And there is no indication of the types of concepts included in the study. Were all concepts classificatory concepts, or was there a mix of classificatory, relational, and theoretical concepts? One can only infer from referrals (metamorphosis of a butterfly, motion of the earth around the sun, living things, influence of heat on the expansion and contraction of materials) that there was a mix of concept types. Was this accounted for in designing the study and preparing test items? A list of the concepts should have been included in the report.

The author emphasizes that we should select materials appropriate for various social-economic levels. How critical is this factor? The design of the study does not account for teacher expectations and behavior. How were the existing materials being used? Could the significant differences be attributed to things other than social-economic level?

The implications of the study are extrapolated beyond the findings. The study is a status study. Yet extensive time is devoted to discussing activities-oriented science instruction in the introduction and the concluding remarks. The curriculum of the participating school system was text oriented. Don't children from lower social-economic levels profit as much from text oriented science as children in higher levels? The ability to profit from school oriented experiences is a function of pre-classroom and non-classroom experiences. Can we generalize so much from a status study? Many factors must be considered and controlled in selecting the highest potential for student treatment interactions. Certainly social-economic level is a major factor to consider, especially with concept learning so dependent on a breadth and depth of quality experiences with phenomena.

The study addresses itself to major lines of science education research, concept learning and aptitude-treatment interactions. It further verifies that a careful matching of instructional materials and the stage of development of the learner must be pursued. Shotgun therapy isn't always productive and changing one prescription for another won't help if we guess at the symptoms and ignore the underlying causes of the malady.

EJ 046 113

Kline, Arlyn A., "A Study of the Relationship Between Self-Directed and Teacher-Directed Eighth Grade Students Involved in an Open-Ended Supplementary ESCP Laboratory Block." Journal of Research in Science Teaching, Vol. 8, No. 3:263-271, 1971.
Descriptors--*Discovery Learning, *Instruction, *Laboratory Procedures, Attitudes, Earth Science, Educational Research, Science Course Improvement Project, Student Centered Curriculum (Earth Science Curriculum Study)

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Rolland B. Bartholomew, The University of Texas at Austin.

Purpose

The purpose of this study was to ascertain whether or not the open-ended laboratory block on soils could be learned as efficiently by self-directed students as by teacher-directed students, and to determine students' attitudes and interests toward learning science through a laboratory-block discovery approach.

The six hypothesis tested in the study are: There will be no significant difference in

1. cognitive understanding of the laboratory block between the two groups of students.

2. post laboratory achievement between students in the first session and those in the second session as a result of time difference between the two sessions or added insight gained during session II.

3. interest shown toward the block between the teacher-directed and self-directed students.

4. attitudes exhibited by teacher-directed and self-directed students toward involvement in future laboratory blocks.

5. difficulty experienced in reading comprehension of the block between teacher-directed and self-directed students.

6. difficulty experienced in mathematical calculations within the block between teacher-directed and self-directed students.

Rationale

The researcher supported his study by quoting from the work of Schwab, Bruner, Sachman and Butts. The basic assumptions underlying the study are:

1. students are different in many ways, and
2. instruction must be tailored to meet the different needs of the students.

Research Design and Procedure

The sample consisted of 97 junior high students using the ESCP text -- Investigating the Earth. The study began after all students had completed the first 12 chapters in the ESCP text. The students were randomly assigned to two groups -- a self-directed group of 49 students and a teacher-directed group of 48 students. Each group was further divided into two subgroups for convenience of class size and scheduling. All groups met for 45 minute periods for 20 consecutive days. The researcher taught the two teacher-directed groups and served as a consultant to the two groups of self-directed students.

The researcher designed the laboratory block materials used in study (soils). At the completion of the laboratory block, each student was given a post-laboratory test designed by the researcher. One day later, each student completed a 10-item questionnaire.

Data gathering instruments included:

1. the post-laboratory test containing 50 four-item multiple choice questions, and
2. a student questionnaire containing 10 questions.

The researcher used standard procedures to determine reliability, item difficulty, cognitive levels of test items, and a weighted response for each question in the questionnaire. Data were obtained for evaluating each of the six hypotheses.

Findings

The findings were:

1. No significant differences were found between the two groups on the post-laboratory test.
2. No significant differences were found between groups in session I and those in session II.

3. No significant differences were found in responses on the student questionnaire between the two groups in either interest or attitude.

4. A significant difference was found in the degree of reading difficulty between the two groups. The self-directed students had the most difficulty.

Interpretations

The researcher states, "...it appears feasible that similar blocks could be developed, not only to supplement the ESCP text, but also for other areas of junior high school science." Also, the researcher felt that the high percentage of students (96 percent found the block interesting and 90 percent indicated they would like to do similar work) indicated that the block provided a challenge and freedom and flexibility for each student.

Abstractor's Notes

In this study the researcher made a strong case for the necessity to develop teaching materials that would fit the individual needs of students. However, he set up a laboratory-block that all students were required to take. What provision for individual differences were made within the teaching of the laboratory block were not described nor tested and were clearly not considered in his analysis of his finding. Thus, the study is built around a premise that is not examined. The reviewer is left with the opinion that what the researcher wanted to examine was not studied.

The analysis of the experimental results begs many serious questions. For example, since no significant differences were found between the two groups -- except for reading comprehension -- is it logical to propose that similar blocks could be developed? Would more blocks prove anything? What is gained by one teaching procedure over another? This study indicates no gain in favor of either self-directed study or teacher-directed study.

Obviously, the researcher worked hard on this study and he produced some challenging materials. However, after it's all over, the reviewer is left with the impression that the researcher's original question was in error. What he set out to do could not be done with the tools he used.

EJ 038 125

Koran, John J., Jr., "A Study of the Effects of Written and Film-Mediated Models on the Acquisition of Science Teaching Skill by Preservice Elementary Teachers." Journal of Research in Science Teaching, Vol. 8, No. 1:45-50, 1971.
Descriptors--*Classroom Observation Techniques, *Questioning Techniques, Evaluation, Instructional Films, Models, Preservice Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Judy C. Egelston, University of Rochester.

Purpose

This study compared two methods of teaching questioning techniques to preservice teachers with the hypothesis that the method using a film of model questioning behavior would be superior to the reading of a script of the film.

Rationale

When attempting to change and improve teaching behavior, research results suggest that observational learning may be more efficient and effective than the traditional approaches of listening to a lecture or reading information. According to Koran, the film-mediated modeling procedure where the preservice teacher observes the actual filmed performance of the desired teaching skill has been shown superior to written models for some tasks. Is observation-classification questioning skill one of these?

Research Design

Variables High school student raters evaluated questions written by three groups of science methods students before viewing the models and after each of two exposures to the models. The independent variable included a third treatment group of students as a control since they did not experience any model. The dependent variable was the total number of appropriate observation and clarification questions asked.

Conditions A total of 31 undergraduate female elementary education majors, who were members of the author's science methods class, were randomly assigned to one of three groups. One group spent 14 minutes viewing a film of a teacher uring appropriate observation-classification questions with a class of first graders who were working on a unit from Science - A Process Approach. A second group was allowed 14 minutes to read a script of the film, and a third (control) group received unspecified placebo materials for 14 minutes.

Procedures

Testing All 31 subjects were given the materials from a kit from the S-APA program for a first grade unit and were directed to write questions which might be used with a first grade class. Three high school student evaluators were given the questions by random assignment to evaluate as appropriate for observation or classification questions for the kit of materials. These data were planned as a pretest, but they could not be used as a pre-treatment covariate because of lack of parallel regression lines. Hence, there were no tests for between group differences carried out.

The three groups spent 14 minutes exposed to their respective models and were again asked to generate written questions, which were again scored by the secondary student evaluator to form the data for trial two. A second exposure and follow-up question writing session comprised trial three.

Analysis Only within group data were investigated, using a one-way ANOVA (trials by subjects design). The changes in frequency of questions over the three trials were displayed in a bar graph.

Results The film-mediated and written models produced significant ($p < .05$) gains in questioning skill from the pretest to trial three. The film-mediated model produced significant gains over the control group. By the third trial the written modeling group wrote significantly more questions than the control group.

Interpretations and Generalizations

The written model may be particularly effective in training individuals who initially do poorly in writing observation-classification questions, while those trained in questioning skills with a film-mediated model reach a ceiling of performance with higher level responses such as classification questions. Written models, since they are more practical, less expensive, and as effective as film-mediated models, would seem the better recommended method. Both methods appear promising to teach questioning skills to preservice teachers.

Abstractor's Notes

Unfortunately, it would be impossible to replicate Koran's study because the procedure is not adequately described. He never explains exactly what the 31 subjects do in the three trials. Some information may be inferred from the tables, but the report does not explain the tables. One may also wonder why there were two exposures to the models. Perhaps one was used for observation questions and the other for classification questions. The control of timing the exposure to the written model would seem a sham. Even if the length of the film (14 minutes)

included times when no verbal interaction took place, it does not seem reasonable to assume that subjects would be able to perform as well if they did not finish reading the script in the 14 minutes. Or alternately, the written model group may have had enough time to read the script more than once.

The reader is not informed about the relationship of the three high school students to the subjects. It is also impossible to tell from the report how the raters were trained. It is never clarified whether the raters were blind to the treatment of the subjects whose questions they were rating. If the experimenter hypothesized that those subjects who saw the film would perform better, then raters may have been biased by this information, in spite of random assignment of questions.

Koran appropriately chose not to test the between group comparisons as planned because of lack of parallel regression lines. However, they could have been tested by using an ANOVA model with multiple contrast procedures contrasting cell means (2 way model).

The author investigated differences in performance in questioning skill between trial one (pretest) and trial three. It would seem interesting and more critical to examine the differences between trials one and two, since the data in the bar graph indicate that this is where most of the behavior change occurred. Examination of Figure 1 would also lead one to question whether Koran ran a check on the adequacy of randomization of subjects to groups. It certainly appears that the poor questioners were lumped together. The way Figure 1 appears also leads the reader to believe that the relative gain of the two groups is equivalent. It might be suggested that the experimenter perform a 1-way ANOVA on the pretest scores to look at the gain scores between groups.

Koran hypothesizes a hierarchy (control < written < film) but never tests it or discusses it. While it is recognized that analysis of gain scores has limitations, it is better than nothing, especially in light of his original hypothesis. The most appropriate analysis would be a split plot ANOVA. Koran's study was interesting and well planned. Unfortunately he never got at where the differences between groups or within groups lie.

EJ 074 390

Lawlor, Francis X. and Elizabeth P. Lawlor, "Teacher Expectations: A Study of Their Genesis." Science Education, Vol. 57, No. 1: 9-14, 1973.
Descriptors--*Ability Identification, *Preservice Education, *Student Ability, *Teacher Characteristics, *Teaching Skills, Educational Research, Elementary School Teachers, Science Education, Teacher Attitudes

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Donald E. Riechard, Emory University.

Purpose

The three-part problem was defined as: Finding out

1. Whether in fact teacher expectations (judgments of a child's ability) would be consistent despite a brief observation period.

2. Whether these judgments would be formed with an absolute minimum of verbal cues.

3. Whether there were some salient behavioral cues which were used to form these expectations.

Rationale

This study is broadly grounded in the notion that teacher expectations of student performance function as self-fulfilling prophecies. The research and writing of Rosenthal and Jacobson, Thorndike, Good and Brophy, and Rowe are referenced. The authors attribute their specific interest to the fact that undergraduate student teachers working on science lessons in microteaching situations consistently expressed evaluations of student ability after a 15-minute contact with the children.

Research Design and Procedure

The subjects were 72 college undergraduates enrolled in four sections of an elementary school science methods course. The stimulus was a video tape of two 10-minute science lessons involving a bulbs-and-batteries exercise with second graders from an inner city school. The first lesson involved five children; the second lesson involved four children.

The subjects viewed the tapes in groups of twenty and were asked to make written anonymous responses. There were two viewings of each tape. On the first viewing, subjects were asked:

1. To rank order the children or at least to identify the top and the bottom student in terms of ability.

2. To list the clues used in making the ability judgments.

Subjects were told that if they did not have sufficient information to make judgments, they should indicate this fact.

In the second viewing, the tape was stopped every two minutes or upon demand of a subject. At this point each viewer was asked:

1. To suggest a specific teaching action that might be appropriate.

The data were organized in three tables and two graphs. Results were also presented in written (sentence) form.

Findings

This abstractor summarized the results as follows:

1. Five subjects (14 percent) did not rank the children. Two of these stated that there was not sufficient information on the tape.

2. On lesson 1, 72 percent of the subjects agreed on the highest ability child and 80 percent agreed on the lowest.

3. On lesson 2, 50 percent of the subjects agreed on the highest ability child and 64 percent agreed on the lowest.

4. Approximately 20 percent of the subjects gave no reasons for their rankings.

5. The 248 "reasons for rankings" that were given were classified as follows:

    a. There were 149 "unsubstantiated inferences."

    b. There were 109 "factual observations."

6. Subjects listed a total of 171 teaching actions (teaching moves).

    a. Twice as many teaching actions were directed at the "most able" children than at the "least able."

    b. Forty-seven percent of the teaching actions suggested were vague and general despite instructions to be specific.

Interpretations

The abstractor summarized from the authors' conclusions as follows:

1. Apparently, the function of the teacher as the judge and evaluator of pupils is so much a part of the teacher role that the student teacher is very easily pushed into making these judgments even when the evidence for judging is extremely skimpy.

    a. There is most agreement among judgments on the high and low ends of the ability scale.

2. Apparently, it is difficult to link ability judgments with specific observed behaviors.

    a. The use of "speed of task accomplishment" as a basis for judgments in the absence of verbal cues is consistent with the use of timed intelligence tests, the criteria for the College Quiz Championships and the brief "Wait Time" exhibited by teachers at all levels.

3. When a teacher must decide what moves to make in order to fulfill her instructional role, she is anxious to help those with low ability and to recognize the accomplishments of those with high ability. The end result of these two tendencies is to direct a greater number of moves toward the high ability children and to place the low ability child in a role subordinate to both his teacher and his more able peers. This may be a factor contributing to the "Pygmalion Effect."

Abstractor's Notes

In order to preserve the research report's integrity and the authors' intent, much of the above is quoted directly from the article. In several instances, however, the authors' comments have been summarized and/or reorganized so as to make them appropriate for this abstract.

The most important of the concerns expressed below have to do largely with the "tightness" of the study -- that is, with the preciseness and clarity of the problem statement and with the direct lines of relationship leading from that problem statement through the procedures, results, and conclusions.

The Problem and Procedures It might be useful to comment first on the clarity of the problem statement. The first part of that statement deals with, "Whether in fact teacher expectations (teacher judgments of a child's ability) would be consistent despite a brief observation period." It does not become clear until later in the paper that "consistent" really means the number or percentage of the subjects who would agree on an individual child's ability rank. Another point of confusion is the phrase "...despite a brief observation period." Does this imply that teacher expectations are "consistent" over a "long" observation period?

The second part of the problem was to find out, "Whether these judgments would be formed with an absolute minimum of verbal cues." What does "absolute minimum" mean? The authors state that, "A very noisy soundtrack on the tape limited the number of audible verbal cues." The problems of interpreting such statements relative to verbal cues could have been eliminated by simply shutting off the audio portion of the tape.

The formal problem statement did not include reference to teaching actions among the three things the authors sought to find out. However, data were collected on teaching actions (teaching moves). Those data were presented and discussed in the paper. Granted, in the general discussion of the problem, the authors did question, "How do these judgments relate to teaching strategies?" If, however, that question were to be one of the formal questions investigated (and it evidently was), then it would most appropriately have been enumerated along with the other three parts of the problem.

The sample of 72 subjects seems quite adequate. A question might be raised, however, about the number of children available for ranking by the subjects in each video taped lesson. With five children (first lesson), each child to be ranked represents 20 percent of the total; in the second lesson with four children, each child represents 25 percent of the total. Therefore, the "chance alone" probability of a child being placed at any particular rank is quite high.

The Results Results of the study are presented effectively. There is some duplication of effort in that the graphs of Figures 1 and 2 are simply different ways of presenting the same data found in Tables I and II, respectively. It is questionable if the dual presentation is necessary in this particular research report but it (dual presentation) is certainly a valid technique when data are difficult to interpret.

It might be noted here that tables should always carry descriptive titles. Table III in this study was not titled (most probably due to a simple oversight).

The Conclusions As a matter of style, conclusions are more easily understood if referenced directly to the problem statement. This can be accomplished with re-statements of research hypotheses or questions so that each answer, as determined by the investigation, can be stated explicitly in close proximity with its corresponding hypothesis or question. Conclusions in this study were presented in a general discussion form without direct reference to the research questions or how those questions should be answered.

Conclusions in this investigation seem mainly to be reiterations of the results of the study. While this is quite legitimate, it is generally helpful when additional interpretations and/or implications are presented. For this study, such questions as the following come to mind:

1. What are the implications of this research for classroom teachers?

2. What are the implications of this research for teacher preparation programs?

3. What are the implications of this research for future investigations?

Speaking specifically to the stated conclusions in this research report, the abstractor presents the following comments and questions:

1. It is stated that subjects were easily pushed into making judgments on childrens' abilities even when the evidence for judging was extremely skimpy. Question--When the instructions were given to the subjects, were they (subjects) made to "feel," inadvertently, that making a judgment was "better" than not making a judgment?

2. The terms "teacher" or "teachers" are used in parts of the conclusions. Questions--Do the authors imply generalizing from "72 undergraduates" to "teachers" as a whole? Would the results of this study differ if the subjects had been experienced teachers?

3. It is stated that basing ability judgments on "speed of task accomplishment" is consistent with several practices including "...the brief 'Wait Time' exhibited by teachers on all levels." Question--What is "Wait Time"?

Finally, the authors are commended on the caution used in interpreting their data relative to the self-fulfilling prophecy ("Pygmalion Effect") referred to in the opening paragraph of the study. Their investigation concerned student-teacher judgments about children's abilities and did not treat the effects that such judgments might have on children or teachers. Writers sometimes fail to distinguish between empirical findings and hypothesized relationships. The authors' caution is exemplified as they interpret a portion of their results with the statement, "This factor may be a factor contributing to the 'Pygmalion Effect.'" (Italics added by the abstractor for emphasis.)

EJ 037 866

Mackay, Lindsay D., "Development of Understanding About the Nature of Science." Journal of Research in Science Teaching, Vol. 8, No. 1: 57-66, 1971.
Descriptors--*Attitudes, *Evaluation, *Scientific Enterprise, Integrated Curriculum, Secondary School Science, Scientific Literacy

Expanded Abstract and Analysis Prepared Especially for I.S.E. by George O'Hearn, The University of Wisconsin at Green Bay.

Purpose

Mackay designed this study "to use TOUS (Test on the Understanding of Science) to measure the changes that occur in students' understanding of the nature of science as they study integrated science courses in Grades 7 to 10 of Victorian secondary schools (Victoria, Australia)."

Rationale

Understanding the nature of science is cited as "one of the most commonly stated objectives for science education" and "probably the most important purpose of science teaching." Supporting statements on the place of the nature of science in developing scientific literacy are presented from the American Association for the Advancement of Science, the Commission on Science Education, the International Conference on the Education of Professional Physicists and the Organization for Economic Cooperation and Development. Mackay does not attempt to define the nature of science but, rather, relies on the Cooley and Klopfer analyses of the nature of science as represented in the TOUS test, Form W. Previous research on the nature of science and assessment of student understanding of the nature of science is cited and includes the published research by Welch and Pella who used the Science Process Inventory and Kimball who used the Nature of Science Scale. In addition, it was reported that several forms of TOUS have been used by Schmidt, Miller, Broadhurst, Trent, Crumb and Klopfer.

No indication is given on the emphasis which the nature of science receives in the integrated science course curriculum which is indicated in the schools in the study. No information is available on the relationship of the content or scales of the TOUS Test with the established curriculum.

Research Design and Procedure

Using the TOUS Test's Items and Sub-scales as a criterion reference, Mackay obtained data from a random sample of three state high schools in Melbourne, Australia. The test was administered to students studying Science in Grades 7, 8, 9 and 10 at the beginning of the 1968 school year, and after the final examination at the end of the same year. The initial population included 1,556 students.

No attempt was made to identify or control variables such as curriculum variation, school size or socio-economic factors. It is noted that Science is compulsory in Grades 7 and 8, but it is an elective in Grades 9 and 10.

The basic design of the study provided for status assessment, using the TOUS Items as the criterion reference. No attempt was made to establish controls or manipulate variables.

Analyses of results proceeded on this basis. Results were obtained from both male and female students at each of the four grade levels. Data were analyzed, using statistical techniques, including "t" tests, estimates of reliability and Chi Square. Gain scores were calculated for each grade for males and females. Frequency gains were examined on individual items, using data from the start of the year for Grade 7, and from the end of the year for Grade 10.

Findings

Gain scores were partitioned into "mean gains in test-taking ability" and "mean gain due to improved performance on TOUS items," using a calculation based on the number of "students correctly answering questions on the second occasion which had not been reached on the first occasion." From this analysis, it was concluded that "a higher proportion of students completed TOUS on the second testing than on the first testing and, for some groups, this accounted for as much as 50 percent to 60 percent of the increase in mean scores on TOUS from the first to second testing."

Significant gain scores were reported at each grade level when the end-of-the-year scores were compared with the starting scores. The "null hypothesis was rejected in favor of the alternative hypothesis that there had been an increase in the probability of correctly answering TOUS Items between the first and second testing for the group of males in Grade 7 and females in Grades 7, 8, 9 and 10."

"The proportion of students correctly answering TOUS Items in Grade 10 was significantly higher than for the Grade 7 group for both male and female students on 28 items." Further analyses of these data resulted in the identification of the following "common deficiencies in students' appreciation of the nature of science."

Area I: Understandings About the Scientific Enterprise

Students commonly showed a lack of appreciation of (1) the influence of the people and government on scientific activity; (2) the role of creative endeavor in science; (3) the dynamic, expanding and on-going nature of the scientific enterprise; (4) the functions of scientific journals; (5) the contributions of past findings in scientific research; (6) the functions of scientific societies; (7) the contribution of scientists in many countries in advising United Nations agencies.

Area II: Understandings About Scientists

Common misunderstandings were that scientists (1) do not need to be creative; (2) use "scientific attitudes" when they were doing almost anything; (3) spend every possible minute in the laboratory; (4) have honesty and self-criticism as personal characteristics; (5) are born with a special scientific aptitude; (6) are serious and intelligent people.

Area III: Understandings About the Methods and Aims of Science

Students commonly exhibited a lack of appreciation (1) of the functions of scientific models; (2) of the roles of theories and their relationships to research; (3) of the distinctions between hypotheses, laws and theories; (4) of the relationship between experimentation, models and theories and absolute truth; (5) that science is not solely concerned with collection and classification of facts; (6) of what constitutes a "scientific explanation"; (7) of the difference between science and technology; (8) of the close interrelationships between and interdependence of the different branches of science.

Interpretation

Mackay concludes that there is a need to develop teaching materials to improve the students' understanding of the nature of science and that "until such time as these materials exist, one of the most important objectives of science teaching will largely fail to be achieved."

Abstractor's Notes

This is an assesament of TOUS scores on a population of students in three Melbourne, Australia high schools. A research design is not provided and key variables are not controlled. Characteristics of the populations in the three selected high schools are not discussed and the curriculum is not mentioned, so the reader is left to speculate about the relevance of the TOUS Test to the curricular objectives and emphases. Cultural and linguistic variations are not considered and the philosophical bases for defining the nature of science are not discussed.

The intent of assessing students' understanding of the nature of science is laudable and the use of an instrument such as TOUS to obtain bench mark data is defensible.

Mackay does use these data in an attempt to demonstrate the deficiencies in students' understanding of the nature of science. The statistical techniques employed are of questionable value and, perhaps, of questionable validity, given the myriad of uncontrolled variables and assumptions. For example, there is no statement concerning the homogeneity of the school populations while this is a necessary assumption of the employed statistical procedures. Additional information is also needed on the procedure employed in partitioning gain scores. Finally, analyses of gains on individual items imply a substantial level of item validity. Additional information on this and the other points would have been welcomed by the reader. The many questions raised by Mackay are significant but the heavy dependence on statistical processes and the lack of control over serious variables obfuscates an otherwise important study.

EJ 038 273

Phillips, Darrell G., "The Development of the Concept of Displacement Volume: A Hierarchial Model and Its Partial Testing Under Two Methods of Presentation." Journal of Research in Science Teaching, Vol. 8, No. 1:9-19, 1971.
Descriptors--*Conservation (Concept), *Cognitive Processes, *Elementary School Science, *Learning Processes, Learning, Models, Organization

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Jerry G. Horn, The University of South Dakota.

Purpose

The primary purpose of this study was to investigate the possible existence of a sequential set of conservation concepts which lead to the conservation of displacement volume. A secondary purpose was to examine two different methods of task presentation (graphic presentation and object presentation). The hypotheses reflecting the stated purposes and tested in this study were:

(1) H_0: Success on the conservation tasks utilized in this study is not related to a hierarchial sequence in the order predicted by the model.

(2) H_0: There will be no significant difference of mean scores on the conservation tasks between subjects in the Graphic presentation and subjects in the Object presentation.

Rationale

The "rediscovery" of the work of Piaget (1) by American scientists and educators has produced a remarkable amount of interest in Piaget - type cognitive research, and replication and training studies based upon Piaget's work have been reported in profusion, but investigations attempting to explore the substructure or prerequisites of certain concepts have been relatively few in number. Curriculum development, especially in science and mathematics, has also reflected the impact of Piaget's work, but often the results have been disappointing. At first glance it may appear that a well-defined science or mathematics curriculum can be derived directly from Piaget's stages and substages of intellectual development, but such is not the case. The levels of development, as outlined by Piaget, are extremely broad and lack the prerequisite sequencing necessary for curriculum development.

The possible existence of a sequential set of conservation concepts which lead directly to the conservation of displacement volume and suggested by Piaget's model is of particular interest in this

study because of the doubt raised in a study by Lunzer (5). Other research findings by Piaget (4), Lunzer (5), Lovell (6), Lovell and Ogilvie (7), and Uzgiris (8) raise questions about the idea that interior volume is conserved before displacement volume.

The literature indicates that there is at least a three- or four-year time lapse between the conservation of interior volume and the conservation of displacement volume. There is some disagreement, however, as to the proportion of eleven- and twelve-year-old subjects who have attained conservation of displacement volume. As Uzgiris (8) has pointed out, this discrepancy may well be caused by the type of materials used and the methodology employed in the task protocols.

The data from the research literature and a pilot study by this author prompted the postulation of a hypothetical model leading to attainment of conservation of displacement volume. The model follows as a table.

Proposed Hierarchical Model of Conservation Concepts Leading to Conservation of Displacement Volume

Level	Conservation Concept	Materials	Transformation
1.	Continuous quantity	Water	Shape
2.	Discontinuous quantity	Blocks	Subdivision
3.	Interior volume	Blocks	Position
4.	Interior volume	Blocks	Shape
5.	Displacement volume	Clay in water	Position
6.	Displacement volume	Blocks in water	Position
7.	Displacement volume	Clay in blocks	Position
8.	Displacement volume	Blocks in blocks	Position
9.	Displacement volume	Clay in water	Shape
10.	Displacement volume	Blocks in water	Shape
11.	Displacement volume	Clay in blocks	Shape
12.	Displacement volume	Blocks in blocks	Shape

Research Design and Procedure

This study employed the individual interview technique with two methods of task presentation: (1) Object Presentation, i.e., equipment and standard protocol; and (2) Graphic Presentation, i.e., line drawings, no equipment, but the same protocol as object presentation. Levels 1, 3, 6, 8, 10, and 12 of the model were chosen for examination. Level 1 was investigated in Task I, and Level 3 was investigated in Task II. Since Levels 6 and 10 deal with the displacement of water,

they were combined into Task III; in like manner, Levels 8 and 12 (displacement of blocks) were combined into Task IV.

The experimental sample was composed of 120 children selected from two Tallahassee, Florida, public schools. Three constraints were imposed upon the selection:

1. Forty children were selected from each of three grades (3, 5 and 7).

2. An equal number of males and females were selected from each grade.

3. All subjects had an I.Q. score within the range of 100-200 as measured by the California Short-Form Test of Mental Maturity.

After the sample was selected, each subject was randomly assigned to either the Object Presentation Group or the Graphic Presentation Group. Each subject was interviewed individually by the experimenter, and all questions and responses were oral and were recorded by a concealed audio tape recorder. All subjects were given the same selection task and the same four conservation tasks, but the order of the conservation tasks was randomized. These tasks are listed below, and a brief description of the selection task is provided.

A. Selection Task - This task, presented first, was designed to make certain the subjects used in the study knew two basic facts: (1) that when an object is placed in a container of water, the water level rises, and (2) a large object causes a greater rise in water level than a small object. The task involved the subject's prediction about the comparative displacement of water by two cubes of obviously different volumes in beakers of water. If a child could not predict correctly, he was deleted from the study.

B. Task I: Conservation of Continuous Quantity (water).

C. Task II: Conservation of Interior Volume.

D. Task III: Conservation of Displacement Volume - Continuous Medium Displaced (water).

E. Task IV: Conservation of Displacement Volume - Discontinuous Medium Displaced (blocks).

Each of the four task protocols was composed of two subtasks and required that the subject give separate choice-reason responses. In order to receive a passing score on a subtask, a subject has to state a reason based upon one or more of these categories: (1) logical necessity, (2) compensation, and (3) reversibility.

Findings

Using a modified Guttman (16) scalogram analysis, the sequential dependency of the tasks was determined. This statistical procedure allowed determination of a coefficient of reproducibility (R_{rep}), and the minimum marginal reproducibility (M_m).

The percent of subjects passing the tasks varied from 5 to 100. The order of the four tasks as obtained from the scalogram analyses was not the anticipated order of I, II, III, and IV, but I, II, IV, and III. Since the R_{rep} is a true measure of the scalability of a group of scores when no item is passed or failed by over eighty percent of the subjects, the data for Task I were eliminated and a second set of scalogram analyses was computed for Tasks II, III, and IV. The obtained order for these tasks was II, IV, and III. Using the same procedures and eliminating the data from Task I, scalogram analyses were computed to determine the scalability of the subtasks. The order of the subtasks obtained from the analyses was IIa, IIb, IVb, IIIa, IVa, and IIIb. A test of significance as reported by Chilton (22) indicated that the obtained values of R_{rep} could occur by chance in only five out of one hundred such presentations.

An analysis of variance was computed for Tasks I, II, III, IV, and all four tasks combined. Significant F ratios (.05 level) were obtained between grade levels, but not between types of presentations.

Correlation coefficients were computed to examine possible relationships between total score on tasks and I.Q. score. The correlation coefficients ranged from 0.01 for grade 5 to 0.19 for the combined data from grades 7, 5 and 3.

Interpretations

Relative to the model, the hypothesis that the tasks used in this study are not related to a unidimensional scale in the order predicted was not rejected. The tasks were related to a unidimensional scale, but the obtained order of Model Levels was 1, 3, 8, 10, 12 and 6. The type of presentation (graphic or object) was not found to be a significant variable, but the acquisition of these science related concepts was shown to be significantly different (.05) between grade levels.

The investigator suggests research to further test all levels of the model and the model itself should be elaborated and extended. Also, additional presentation techniques, such as paper-and-pencil and other audio visual methods, need to be investigated. From an application point of view, the hierarchical sequence and the two modes of presentation open several possibilities for teaching the concept of displacement volume.

References

1. Piaget, J., "Cognitive Development in Children." Journal of Research in Science Teaching, Vol. 2:170-186, 1964.

2. Raven, R. J., "The Development of the Concept of Momentum in Primary School Children." Journal of Research in Science Teaching, Vol. 5:216-223, 1967-1968.

3. Belanger, M., "Learning Studies in Science Education." Review of Educational Research, Vol. 39:377-395, 1969.

4. Piaget, J., B. Inhelder and A. Szeminska. The Child's Conception of Geometry. New York; Harper and Row, 1964.

5. Lunzer, E. A., "Some Points of Piagetian Theory in the Light of Experimental Criticism." Journal of Child Psychology and Psychiatry, Vol. 1:191-202, 1960.

6. Lovell, K., The Growth of Basic Mathematical and Scientific Concepts in Children. London: University Press, 1961.

7. Lovell, K. and E. Ogilvie, "The Growth of the Concept of Volume in Junior School Children." Journal of Child Psychology and Psychiatry, Vol. 2:118-126, 1961.

8. Uzgiris, I. C., "Situational Generality of Conservation." Child Development, Vol. 35:831-841, 1964.

9. Elkind, D., "Children's Discovery of the Conservation of Mass, Weight, and Volume." Journal of Genetic Psychology, Vol. 98: 219-227, 1961.

10. Elkind, D., "Quantity Conceptions in Junior and Senior High School Students." Child Development, Vol. 32:551-560, 1961.

11. Piaget, J. and B. Inhelder, Le De'velopement des Quantitie's Chez l"Enfant. Neuchatel: Delachau_ et Niestle, 1941.

12. Phillips, D. G.,"An Investigation of Possible Hierarchial Dependency of Four Piaget-type Tasks Under Two Methods of Presentation to Third-, Fifth-, and Seventh-Grade Children." Unpublished doctoral dissertation, Florida State University, 1967.

13. Karplus, E. F. and R. Karplus, "Intellectual Development Beyond Elementary School." School Science and Mathematics, Vol. 70: 398-406, 1970.

14. Bussis, A. M., "From Theory to Classroom." Educational Implications of Piaget's Theory. Waltham, Mass.: Ginn-Blaisdell, pp. 328-344, 1970.

15. Masako, F., J. F. Campbell and J. S. Helmick, "Piaget for First-Grade Teachers: Written Exercises for Assessing Intellectual Development." Educational Implications of Piaget's Theory. Waltham, Mass.: Ginn-Blaisdell, pp. 324-328, 1970.

16. Guttman, L. A., "A Basis for Sealing Qualitative Data." American Sociological Review, Vol. 4:179-190, 1944.

17. Goodenough, W. H., "A Technique for Scale Analysis." Educational and Psychological Measurement, Vol. 4:179-190, 1944.

18. Edwards, A. L., Techniques of Attitude Scale Construction. New York: Appleton-Century-Crofts, 1957.

19. Wohlwill, J. R., "A Study of the Development of the Number Concept by Scalogram Analysis." Journal of Genetic Psychology, Vol. 97:345-377, 1960.

20. Cox, R. C. and G. F. Graham, "The Development of a Sequentially Scaled Achievement Test." Journal of Educational Measurement, Vol. 3:147-150, 1966.

21. Kofsky, E., "A Scalogram Study of Classificatory Development." Child Development, Vol. 37: 191-204, 1966.

22. Chilton, R. J., "Computer Generated Data and the Statistical Significance of Scalogram." Sociometry, Vol. 29:175-181, 1966.

23. Piaget, J. and B. Inhelder, The Child's Conception of Space. London: Routledge and Kegan Paul, 1963.

Abstractor's Notes

The review of the literature, including the work of Piaget, Uzgiris, Elkind, Lowell, Ogilvie and others, has raised some real questions about cognitive development, particularly related to the development of conservation concepts. The casual reader may easily read one source and fail to pick up the contradictions and/or potential contradications in the findings of another person's work. Phillips seems to have been particularly acute at identifying points for further study. A survey of the listed references reveals that most of the literature offered for review was published within a relatively short period of time, ten years or less. Was there no significant research work on this topic prior to 1960 except one by Piaget and Inhelder?

The hierarchical model proposed for testing of its validity held up remarkably well in this study. Phillips indicates that the model was postulated from his review of the literature and the data from a pilot study. Since the pilot study was referenced as his unpublished

dissertation, one might suspect that it was rather carefully done and was scrutinized by several persons, either as the research topic of the dissertation or as a preliminary study to the dissertation.

The research design in this study is clean and focuses on the two main purposes of the study. The relationship between the levels and the tasks is clearly identified, and the exclusion of Levels 6, 8, 10, and 12 is mentioned, but the specific reason(s) for exclusion is missing. The statistical treatment of the data relative to the two main hypotheses is well documented and seems very appropriate. The reported correlation coefficients for total score on the conservation tasks and IQ seem foreign to the stated hypotheses. Furthermore, the IQ range of 100-120 would seriously reduce the expectation of finding significant correlations between these two variables.

The strength of this research study lies in the investigation of a model that can and should be tested and retested to establish it as viable for procedures of application. As mentioned in the study, much curriculum development professes to be compatible with the development of the child. A simple model such as described would provide access to vast amounts of research through this hierarchical sequence.

The researcher mentioned that it is not uncommon to note almost word-for-word repetitions of reasons between subjects who received object and those who received graphic presentation. The subjects of this study were selected from a rather small geographic area and a narrow IQ range of 100-120. The generalizability of the finding to other areas and different cultural and intellectually capable children would be important information.

EJ 037 865

Schuck, Robert F., "The Influence of Set Induction Upon Pupil Perception of Effective Teaching, Achievement, and Retention in Units on Respiration and Circulation in the BSCS Curricula." Journal of Research in Science Teaching, Vol. 8, No. 1:51-56, 1971.
Descriptors--*Biology, *Learning Processes, *Secondary School Science, Achievement, Evaluation

Expanded Abstract and Analysis Prepared Especially for I.S.E. by William G. Holliday, The University of Calgary.

Purpose

The purpose of this study was to investigate a set induction training program designed for preservice biology teachers. The effects of the program were evaluated in terms of achievement gain scores of high school students taught by teachers trained in set induction compared with students taught by teachers receiving a placebo treatment.

Rationale

Set induction is a tendency in a learner to view and approach a situation (e.g., problem) in a predetermined manner as a result of initial instruction. In the present study this instruction purported to facilitate continuity "through the use of analogy rather than by simple association". It was previously demonstrated that pupils of preservice teachers who received set inductive training achieved higher gain scores. In addition, those teachers who incorporate set induction techniques into their instructional strategies were perceived as more effective teachers by their students.

Research Design and Procedure

Eighteen preservice teachers were randomly assigned to a set induction training treatment or a control group. Each teacher was also randomly assigned to 10 students who were randomly selected from a pool of 180. Student achievement (short and long term retention) pre- and posttests and a teacher achievement test represented the dependent variables. Teacher's knowledge of the material taught in the classes was evaluated to establish the lack of information bias gained in either treatment group. An introductory training session and three biweekly sessions were attended by all teachers during eight week experiment. The experimental group studied set induction and its application to the biology classroom. The control group received the relevant substantive parts of the BSCS biology curriculum. Students were taught two four-week units and evaluated at the end of each unit (short term retention).

Long term retention tests were administered subsequent to the teaching sessions five weeks later (respiration unit) and two weeks later (circulation unit).

Findings

Student gain scores on the short and long term tests covering the respiration and circulation units generally favored ($p < .01$) the set inductive trainee classes. The F ratios were unusually high. No significant differences were found between either the experimental or control groups in terms of teacher knowledge on the substantive biology material.

Interpretations

This study demonstrated that a set induction training program for preservice teachers can positively affect short and long student achievement. These results are generally consistent with similar set induction experiments in teacher education.

Abstractor's Notes

The effectiveness of set induction as a teaching skill has been established by previous teacher education research as cited by Schuck. However, many researchers have noted that set induction can have an unfavorable instructional effect on learners. For example, Luchin, and Birch and Robinowitz in classical experiments have demonstrated that a learner who is predisposed to view certain problem types in a particular way can be at a disadvantage in solving similar problems of subtle differences. The apparent trouble is that these learners are less likely to try a wider range of problem solving techniques. More recently, set induction techniques, such as interspersed questions, have been investigated by Rothkopf and Frase and are commonly referred to as mathemagenic activities (i.e., the study of those learning behaviors that determine the nature of the effective stimuli).

Schuck's research design and procedure were excellent for four reasons. First, he used student achievement (short and long term) as an indicator of the success of a teaching skill in an experimental investigation. Rosenshine and the Korans have constantly urged researchers in teacher behavior to follow this procedure. Unfortunately, research studies evaluating skills similar to Schuck's are correlational in nature and fail to use student outcome as a dependent variable. Second, Schuck's experiment was longitudinal in terms of the teacher training sessions, pupil instruction and the two retention tests. Similar paradigms are commonly encouraged by educators but seldom chosen. Third, the possibility of the Hawthorne effect confounding the experimental results was

diminished through the use of a placebo treatment and a comparative assessment of teacher achievement gain scores between the experimental and control groups. Fourth, teachers and students were randomly assigned to each other and to the treatments. A more common method of assignment prescribes the random assignment of teachers with their students to treatment groups.

Schuck's presentation and analysis of his student achievement scores constituted a weakness in his study. Gain scores derived from the pre- and posttest scores of the two treatment groups were compared using analysis of variance. Cronbach made a detailed examination of the psychometric logic of this technique as proposed by a number of statisticians. In short, he concluded that the use of posttest scores as a dependent variable and pretest scores as a covariate(s) will generally result in the most meaningful comparative information relative to alternative methods. Another analysis problem in Schuck's study was the decision to use the number of students in each treatment (n = 90) as representing the number of cases in the data analysis of student achievement scores. In this study, it would be difficult to argue against the development of confounding student-student and student-teacher interactions within the intact classes during the eight week instructional sessions. As a result of these interactions, a potential historical effect could have influenced student scores in some unknown fashion. Unfortunately, the statistics reported in the study preclude the re-evaluation of the data. The only reported statistics in the reviewed article were the number of teachers and students and the F ratios for the tested hypotheses.

In conclusion, Schuck's study was well conceived in terms of his theoretical framework and experimental design. However, his conclusions must be somewhat tempered because of the cited flaws in the data analysis and statistical omissions within the article.

EJ 046 112

Szabo, Michael and John F. Feldhusen, "Success in an Independent Study Science Course at the College Level as Related to Intellective Personality, and Biographical Variables." Journal of Research in Science Teaching, Vol. 8, No. 3:225-229, 1971.
Descriptors--*Achievement, *Biology, *College Science, *Instruction, *Student Characteristics, Autoinstructional Aids, Educational Research, Personality, Prediction (Audio Tutorial Instruction)

Expanded Abstract and Analysis Prepared Especially for I.S.E. by O. Roger Anderson, Teachers College, Columbia University.

Purpose

(1) To assess the contribution of pupil personality, scholastic achievement, and social-history variables to the prediction of academic success in an audio-tutorial independent study course (A-T) and a traditionally organized lecture-laboratory course (T-O).

(2) To predict achievement for individual students in each of the instructional modes.

Rationale

The author's rationale is more pragmatic than theoretic. Effective use of the many different instructional modes available necessitates empirical studies to determine what kinds of student variables are most compatible with a given instructional mode toward maximizing student academic success. Selection of student variables to be assessed was quasi-theoretical in that the authors selected available indicators of student characteristics that were most likely to be relevant psychologically to the treatment conditions and the learning criteria. They state that certain personality variables are more likely to be relevant to student performance in an independent study course than in a traditional lecture and laboratory course.

Research Design and Procedure

The subjects were 630 college freshmen and sophomores who voluntarily completed the necessary data forms and were enrolled in a one semester, introductory biological science course at the university level. A subset of the sample (N = 312) who had completed, subsequent to the A-T course, a traditionally-taught introductory biological science laboratory course served as a comparison to the A-T group.

The study used multi-variate techniques to assess the relative value of the independent variables in predicting academic success in the A-T and T-O treatments.

The independent variables were: (1) scores on the Guilford-Zimmerman Temperament Survey (general activity, restraint, ascendance, sociability, emotional stability, objectivity, friendliness, thoughtfulness, personal relations, and masculinity); and, (2) high school grades in science, mathematics, social studies, and English, high school graduation rank, College Entrance Examination Board and Scholastic Aptitude Test scores; and (3) biographical information secured with a questionnaire. The criterion measures were final grades in the A-T and T-O courses. Data were correlated using a build-up linear multiple regression analysis model. In addition to analyzing the total group of subjects, the sample was stratified into three achievement level subgroups of equal size (high, middle, low) using first semester predicted grade averages as a criterion.

Findings

The major findings are: (1) for the total sample, both verbal aptitude and restraint (tendency to be serious-minded and responsible or to be introverted) were significantly related to the criterion scores in both the A-T and T-O treatments. Differences between A-T and T-O modes were noted, in that prior achievement in mathematical reasoning skills and social studies were more often related to success in the A-T course while prior achievement in mathematics computational skills and science achievement were more often related to success in the T-O treatment; (2) for learners with high predicted achievement, mathematics reasoning skills and science achievement were significantly related to success in the A-T course while verbal aptitude, mathematical computational skills, and restraint were significantly related to success in the T-O course; (3) among the middle predicted achievement groups, mathematics computational skills, prior science achievement and restraint were significantly related to success in the T-O course. No intellective predictors exhibited a significant relationship with success in the A-T course; and (4) in the low achievement group, verbal aptitude was significantly related to success in the T-O course and achievement in social studies was significantly related to success in the A-T course. In summary, the restraint variable exhibited value as a predictor for academic success for the total sample and the low achievement subgroups in both the A-T and T-O treatments. The restraint scale was significantly correlated with academic success for the high and middle achievement subgroups of the T-O course, but not for the A-T course.

Interpretations

The findings provide guidelines for providing alternative modes of instruction that complement the student's intellective and personality characteristics. An efficient learning environment in science can take into account student characteristics and provide a range of instructional experiences that will maximize individual student success. The prediction system identifies personal characteristics of the student in which he is weak or strong and thus can be used to provide remedial counseling or instruction or to adapt the instruction to fit the student's strength or weakness.

Abstractor's Notes

This research addresses a general question of considerable significance in practice-oriented research in science education; namely, "What are the empirically testable variables that will predict academic success among two or more equally reasonable modes of science instruction?" It assumes a multivariate rationale that inherently suggests a value position in research: that the teaching-learning environment is a complex network of interacting variables and that it is reasonable to examine multiple relationships among a subset of these variables rather than identifying a single pair of variables in isolation from the test. Moreover, there is a tacit practical value orientation inherent in such a procedure; namely, that one method of teaching need not be considered totally or universally superior to another based on a single criterion measure. In the past, much research on teaching methodology has examined two methods of instruction with the assumption that one is totally superior to the other based on a philosophical or psychological rationale. These studies of "racing one course against another" as it were have frequently yielded few significant findings.

In some of these studies, the students were treated as though they were a homogeneous mass when in fact there was considerable within group variance that sometimes swamped the treatment variable effects. The multivariate rationale employed in this research recognizes student variation and accounts for some of it by assessing relevant variables as part of the research design. The net result is that much of the within groups variance, that otherwise could plague the investigator as unaccountable error variance, becomes marked and manipulated toward a larger information yield. From a practical stand-point, we learn more about individual student differences and possibly some ways of meeting needs resulting from these differences through flexible instructional programs organized to match student needs with appropriate learning experiences.

These multivariate investigations are not immune to the more pervasive problem of finding a suitable theoretical basis for selection of predictor variables that are clearly relevant to the criterion achievement variables. The investigators in this study have used a

set of variables selected from a curriculum-building model and from publications showing the predictive ability of the variables in previous research of this kind. This is a rational and systematic approach that at least builds on prior published findings and contributes toward an accumulation of related research. Ideally, some theoretical model will emerge with sufficient predictive ability that variable selection will become more systematic within a consistent frame of reference. The investigators reported that the predictor variables that they used accounted for only 40 percent of the criterion achievement variance, thus indicating there is a substantial amount of additional information to be gained. Additional variables should be identified that may increase prediction accuracy.

It is clear that in studies of teaching method the need for a concise yet complete description of the treatment variable is of the utmost importance. Obviously, one cannot say much in a practical way about the relative contribution of a teaching method if it is defined in imprecise and grossly general language. Multivariate research such as the kind reported here will be strengthened if instructional treatment variables are defined in precise conceptual and operational terms.

This study will have served well if, in addition to its primary findings, it also stimulates other researchers to pursue similar lines of multivariate analysis and to build on this study by extending it to include additional variables that account for more of the criterion variable variance. Other research reports that are relevant to this one are: (1) Gallagher, J. J., "A Comparison of Individualized and Group Instruction in Science: Effects on Third Grade Pupils." Journal of Research in Science Teaching, Vol. 7, No. 3:253-263, 1970; (2) Grobe, C. H., "A Regression Approach to Evaluating Instructional Programs in Science." Journal of Research in Science Teaching, Vol. 10, No.1:55-62, 1973; (3) Koran, J. J. and M. L. Koran, "Differential Response to Structure of Advance Organizers in Science Instruction." Journal of Research in Science Teaching, Vol. 10, No. 4:347-353, 1973.

EJ 037 542

Tisher, Richard P., "Verbal Interaction in Science Classes," Journal of Research in Science Teaching, Vol. 8, No. 1:1-8, 1971.
Descriptors--*Interaction Process Analysis, *Secondary School Science, *Teacher Characteristics, *Verbal Communication, Achievement, Classroom Observation Techniques, Cognitive Measurement

Expanded Abstract and Analysis Prepared Especially for I.S.E. by James R. Campbell, University of Pennsylvania.

Purpose

1. Identify the nature, distribution and patterning of verbal discourse within six secondary science classrooms in Australia.

2. Compare the results with data from classes in America and New Zealand.

Rationale

Conceptual framework - The study utilized the Smith-Meux Category System. This interactive system does represent a simple model for quantifying interactive data on a cognitive level. However, the system has more in common with an operational definition of cognitive verbal behavior than it does with any theoretical framework in this area.

Underlying Assumptions:

1. Language plays an important role in the development of pupils' understandings.

2. Classrooms in different countries with very different settings can be compared.

Related Research: This study is part of a larger study conducted by Tisher and is related to other studies performed by Smith and Meux (1962) and Nuthall and Lawrence (1965).

Research Design and Procedure

Research Design: Since no experimental variable was involved, no quasi-experimental design was used. The investigator used an ecological-field study approach by observing a set of Australian secondary classrooms over a period of time. (Six classes, thirty minutes of observations per class)

Variables: Verbal behavior as measured by the Smith-Meux Interactive System. (Dimensions: describing, designating, stating, defining, substituting, evaluating, classifying, comparing and contrasting, conditional inferring, explaining, classroom management, opinioning).

Inter-observer agreement r = .91. No specification of statistical method used to determine reliability or the number of tapes used to determine the correlation.

Sample: Four teachers in six different grade eight classrooms were utilized. All the teachers were "non-graduates" and did not exhibit any substantial academic training in the sciences. No description of the student population was provided. No details are provided as to the rationale used in selecting these samples.

Findings

1. The dominant verbal behavior for all groups involved the recall behaviors of describing and designating. Such behaviors accounted for 62 percent of all verbal behaviors recorded for the Australian teachers. Similarly, data from a sample of science teachers in Illinois (Smith and Meux, 1962) revealed that American teachers were found to have 48 percent of their verbal behavior in these categories.

2. If the number of incidents involving "stating" are added to these cumulations, then the percentages rise to 73 percent for the Australian teachers and 52 percent for the American sample. New Zealand data (Nuthall and Lawrence, 1965) show similar findings.

3. Verbal behaviors included in the categories Conditional Inferring and Explaining accounted for 20.9 percent of all dialogue for Australian teachers and 15.8 percent for the American sample. The author concludes that these percentages show that American and Australian teachers are similar in their approach.

4. Lower cognitive level behaviors involving designating, describing, and stating predominate in the Australian sample over higher cognitive level behaviors which involve substituting, classifying, evaluating and explaining. The latter account for 23 percent of all incidents.

5. Inquiry behaviors as evidenced by explaining, classifying, comparing and contrasting, evaluating and inferring are not prevalent in the Australian sample.

6. Teachers' questioning behavior ranged from 2/minute to 4/minute.

7. The amount of teacher talk for these Australian teachers was approximately 95 percent in sharp contrast to Flanders' 66 percent for American teachers.

8. Simple, short duration incidents predominated for this sample of classes. Complex incidents "involving the answering of subsidiary questions within the context of answering one major question" occurred rarely.

Interpretation

1. Lower cognitive level dialogue predominates in Australian, New Zealand, and American classrooms.

2. Inquiry behaviors occurred less frequently than expected.

Abstractor's Notes

This segment of the larger study needs more amplification. How much time was spent in training the observers with this system? How many tapes or typescripts were used to determine inter-reliability? how many observers were utilized, both for the inter-reliability and for the whole study? Was any intra-reliability determined?

The report did not describe the students involved. Were they average in ability? Heterogeneously or homogeneously grouped? What was the ability range of the students? These factors have been known to effect the interaction.

Over how long a period of time were the audiotapes collected? Were all tapes collected simultaneously? Were all teachers covering the same subject areas? How was the sample selected? Was there any attempt at randomization of students or teachers?

National Comparisons: Were the grade levels the same for the New Zealand and American studies? Since the New Zealand study did not contain any science groups, are the comparisons valid? Without adequate descriptions of the three samples from different nations, are the comparisons justifiable?

Inquiry Behaviors: The author states that inquiry behaviors "are by no means prevalent ones" in the Australian sample, however, these categories did account for almost 19 percent of the data. This accumulation does seem substantial. The data seemed to contradict the author's conclusion.

Teacher Questioning: The author determined that the teachers use a large number of questions per class. He then used this number to determine the number of questions youngsters were exposed to per school day (700) and the number per school week (3,000 - 4,000). These extrapolations assume teacher consistency in all subject areas. Perhaps different subject areas use different volumes of questions. Likewise the same teacher might use many questions one day and very few the next. Larger sampling would be needed to determine the quantity of questions asked of youngsters. Such extrapolations cannot be made without more data.

In spite of such short comings the study does contain some quantifiable data which will doubtlessly prove fruitful to other researchers in the field. Researchers in observational analysis have been more concerned with description of the complex interactive events occurring in classrooms than with combining these data into paradigms and educational models. Description must be considered the lowest level of the process. Researchers must now begin to utilize these data in larger and larger contexts.

EJ 037 434

Van Koevering, Thomas E., "The Distinguishing Characteristics of High Schools With High and Low Enrollments in Physics." Journal of Research in Science Teaching, Vol. 8, No. 1:37-39, 1971.
Descriptors--*Enrollment Influences, *Physics, *Secondary School Science, *Teacher Characteristics, Attitudes

Expanded Abstract and Analysis Prepared Especially for I.S.E. by John T. Wilson, University of Iowa.

Purpose

Do high schools with either high or low enrollments in physics exhibit any distinguishing, measurable characteristics?

Rationale

Even though the percentage of high school students enrolling in physics has been declining for more than half a century, little research has been undertaken to determine the reasons for the decline. (No other rationale was given).

Research Design and Procedure

Class A, B, C high schools (classes not defined) in Michigan were rank-ordered by percent of seniors enrolled in their physics classes. Physics and chemistry students, physics teachers, and guidance counselors were then selected from those high schools found in the upper or lower quartile. Data were then collected by means of visits and questionnaires to each of the 48 schools which included 51 physics teachers, 1429 physics students, and 2338 chemistry students.

Questionnaires were given physics teachers which focused on personality factors, training and experience, and opinions about physics. Students in physics were questioned concerning their opinions and perceptions about the learning environment, vocational plans, grades and other coursework, and achievement in physics. Chemistry students were questioned about their future plans before and after graduation, especially their intentions toward physics enrollment and the reasons for their decision. Guidance counselors were questioned concerning prerequisites for physics, attitudes, and how enrollments could be increased.

Data were analyzed using a t-test and a Fisher Exact Probability test. (No other explanation about design or analysis was given).

Findings

While several differences between schools with high and low enrollments were reported, only one was identified as significant. Schools with high physics enrollments contained significantly more ($p < .05$) non-science students in the physics class than schools with low physics enrollments. Non-significant findings included differences between the physics achievements in the two types of schools and differences in the numbers of teachers using PSSC in the two types of schools. Summaries of other findings were also reported but no indication of significance or non-significance was stated. These included differences in personalities of the teachers, prerequisites for physics, differences between students' physics grades and other grades in school, and student achievement in physics.

Interpretation

Three conclusions were presented. (1) Neither the personality of the physics teacher nor the classroom environment, as measured by the instruments utilized, seem to have any major influence on physics enrollments. (2) The mathematics prerequisite of three years of high school mathematics is directly linked to low enrollments. According to the results of the achievement scores in physics, this prerequisite seems unwarranted. (3) Many students are not aware of what physics entails and consequently do not enroll in the course.

Abstractor's Notes

In this research effort, desired outcomes were concerns for decreased enrollments in high school physics. However, why fewer students are taking physics, per se, has little generalizable value unless the findings are viewed within a broader rationale. One rationale could be the changing demands for scientists within society and the scientific enterprise. This rationale emphasizes certain social and descriptive variables. Another plausible rationale considers various factors in the learning environment and their relationship to the successes of learners. This rationale suggests yet another set of variables. Both of these are major concerns in science education, not only in high schools with decreasing enrollments but also in colleges.

This research effort provided no citations of literature, which implies that no such review was performed; hence, no rationale was constructed to support the problem presented. Certainly some research has been done which may apply in a general sense; if not, the void should be so reported. Without an adequate framework, findings are limited in interpretation to the specific circumstances of their collection. There are, for example, theoretical notions which place high value on teacher personality and claim that variances here greatly influence learning. Furthermore, the chance for learners to be

successful in a given learning environment, such as physics is considered by these notions to be a fairly good enticement for enrolling. Other notions are no doubt relevant to this study also.

The generally brief problem description is accompanied by sketchy information about the design, procedures for analysis, and an absence of any data tables. The design resembles a "one-shot case study"; no treatments or controls were described. A total of eight questionnaires or published instruments were used; different instruments were administered to each sample group (teachers, counselors, physics students, and chemistry students). Students NOT enrolled in physics were omitted entirely. If common instruments had been used with all students (including a non-physics sample), shifts in attitudes between students in chemistry, in physics, not in physics would be potentially revealing. Some questions should also be raised concerning the statistical analysis. It is not clear how the t-test was applied to the data. Furthermore, in view of the many t-tests, probably performed, only a few were evidentally significant. If a large number of t-tests were performed, then at .05 level of probability, a few could by random error alone be significant. Respectability of data sets and significance of findings require more than selecting respectable instruments and statistical tests. Both the instruments and the analyses must be congruent with the problem and rationale. The absence of clarity in the analysis, and the omission of any data or statements of assumptions, along with the other weaknesses cited above, raise serious questions about the credibility of the research findings.

With the many weaknesses of this particular research effort, perhaps the single reason for its review is that it does examine some variables relative to the school learning environment to see what influence they have. Too often research of this type collects only obvious descriptive data any untrained observer would collect. Poor designs encourage inadequate analysis. A lack of rationale contributes to myopic interpretations. Credibility and generalizability, as a result, are questionable. However, research of this type should not be abandoned; further efforts should be made. Such efforts would, however, benefit from a re-examination of some research design fundamentals, measurement and evaluation guidelines, and a thorough search of the literature relevant to the general domain of the program. The resulting product can only increase the credibility and generalizability of the research effort.