DOCUMENT RESUME

ED 170 175 SE 027 731

AUTHOR	Blosser, Patricia E., Ed.; Steiner, Robert L., Ed.
TITLE	Investigations in Science Education, Vol. 5, No. 3, 1979. [Expanded Abstracts and Critical Analyses of Recent Research].
INSTITUTION	Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	79
NOTE	68p.
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State University, 1200 Chambers Rd., 3rd Floor, Columbus, Ohio 43212 (Subscription $6.00, $1.75 single copy)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Attitudes; *Educational Research; Inquiry Training; *Preservice Education; *Questioning Techniques; Research; Research Methodology; Science Education; Teacher Behavior; *Teacher Qualifications

ABSTRACT

This issue of "Investigations in Science Education" contains 9 articles related to teacher education and a single study reporting the investigation of questioning styles used in high school biology textbooks. Each abstract, written by a science educator, includes bibliographical data, research design and procedure, purpose, research rationale, and the abstractor's analysis of the research. (PB)

ED170175

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC® Clearinghouse for Science, Mathematics, and Environmental Education

Published Quarterly by

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

Subscription Price: $6.0 year. Single Copy Price: $1.75
Add 50¢ for Canadian mai and $1.00 for foreign mailings.

Volume 5, Number 3, 1979

NOTES FROM THE EDITORS

This issue of INVESTIGATIONS IN SCIENCE EDUCATION contains a cluster of articles focused on teacher education and a single study. The single study relates to a topic that should be of interest to teachers: the styles of questions used in high school biology textbooks.

Within the teacher education cluster are found several articles written by the same author(s). The first three articles appear to have resulted from a single research effort involving different subpopulations of the science education community. No doubt in order to meet space limitations imposed by various professional journals, the investigators have reported their findings in different articles, each concentrating on data related to a subpopulation. Two of the reports appeared in the same issue of a journal, Science Education, and have been analyzed by one reviewer. The third article, appearing in Journal of Research in Science Teaching, has been analyzed by a different reviewer. The two reviewers have slightly different perspectives concerning the articles but they also share common concerns.

In addition to the first three articles about necessary teaching skills, there is a related article. It is related in that it focused on competencies (for "competencies," substitute "teaching skills"). The respondents involved in this research study were drawn from within a single state rather than from membership in professional groups that cut across state lines.

There is another cluster-within-a-cluster in the next several articles. These share a common concern: inquiry teaching as well as investigators/authors (Lee, Lazarowitz). Two of the articles, on teachers' attitudes and inquiry,

have been treated in a single review. One of these articles relates prima-
to instrument development and the other, to reporting on attitude change a
measured with the instrument which was developed. Third article also focuses on inquiry teaching strategies but the educational group involved in the study was that of preservice elementary education students rather than inservice secondary school science teachers (as reported in the at[illegible]de change article).

The two remaining articles in the teacher education cluster relate t[illegible] teacher behaviors. One of these, by Cotten *et al.*, investigated the use [illegible] an inquiry/process skills package. The other, by Tamir and Zoor, report on the formative evaluation of a large-scale implementation of an innovative curriculum project.

Patricia E. Blosser
Editor

Robert L. Steiner
Associate Editor

ΞACHER EDUCATION

1/2

Chiappetta, Eugene L. an. Alfred T. Collette. "Secondary Scien. Teacher Skills Iden::ified by Science Supervisors." Science Education, 62(1): 6'-71, 1978.

Descriptors -- *Educational Research; *Performance Based Teacher Educat. Science Education *Science Teachers Secondary Educ. ; *Secondary School Science; *Skills Surveys; Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Geral H. ...ckover, Purdue University.

Purp...

The purpose of Chiappetta and Collette's study was "to use the perception of science supervisor ... determine those competencies that all secondary school science te... rs should possess" (p. 67). Once these competencies were identifi... they could then be used as part of either a pre-ser... or inservice program for secondary science teachers.

Rationale

The authors have attempted to identify initial competencies for use in teacher education programs for secondary science teachers. They used the perceptions of science supervisors for four reasons:

(1) Science supervisors have a close association with secondary science teaching.
(2) They are instrumental in selecting science programs and materials.
(3) Science supervisors observe science teachers in a variety of situations.
(4) They are often consulted in the hiring and firing of science teachers (p. 67).

Research Design and Procedure

Three stratified samples of 100 members from the National Science Supervisors Association (N > 600) were selected to participate in the research. Each sample included a random sample of 8 state science supervisors plus a random sample of 72 other members. Most of the state supervisors in the 50 states were invited to take part in the project. Each sample participated in only one of the three parts of the study.

The procedures in this study were borrowed in part from the "Delphi method." They deviated from the "original Delphi" by providing feedback to different samples of the same population, rather than the same individuals. The modifications were utilized to increase generalizability of the findings to the target population, prevent the loss of respondents over rounds, and eliminate the possibility of manipulating consensus (p. 68). The authors made these modifications to help improve the reliability and validity of the results.

The research was accomplished using a three-round procedure. In round one competency statements were generated by participants. In the second round the participants selected levels on the Taxonomy of Cognitive Objectives that they felt each competency statement should specify. In the third round the participants rank-ordered the competencies in terms of importance (p. 68).

After round one, a panel of five judges grouped the responses into three categories: 1) cognitive competencies, 2) affective competencies, and 3) personality characteristics. The cognitive competencies were the focus of the remainder of the study. After round two, one-sample X^2 tests were used to determine the taxonomic levels consistently chosen by the respondents to represent the skills. In round three the science supervisors ranked the cognitive competencies in order of their perceived importance.

Rank order was determined by the sums of ranks assigned to each competency. The Kendall coefficient of concordance (w) was computed to determine the degree of consistency of the ranking of competencies.

Findings

After round one. [illegible]4 competencies (skills, knowledge, and attitudes) were received [illegible] 45 science supervisors. Then a panel of five judges (two science education professors, a science supervisor, a supervisor of secondary science teachers, and a science teacher) grouped the responses into the three categories previously mentioned. Eighteen cognitive competencies (34 competency statements) were the focus for the remainder of the study.

In round two, 57 science supervisors responded and identified a level on the Taxonomy of Cognitive Objectives at which secondary science teachers should be able to demonstrate a given skill (competency statement). The frequency of responses for each of the 34 competency statements across six taxonomic levels were analyzed using one-sample X^2 tests. Thirty-one competency statements showed significance ($p \leq 0.05$) in the analyses, while three did not. Thus, three competency statements were eliminated and the number of cognitive competencies was reduced from 18 to 15.

The round three supervisors ranked the 15 cognitive competencies in order of importance. Rank order was determined by the sums of ranks assigned to each competency. The Kendall coefficient of concordance (w) was 0.29 ($p \leq 0.001$) for the ranking.

The first eight ranked competencies appear to be important and necessary for science teachers. They are:

1) formulates instruction that emphasizes the inquiry and process approach to teaching science,
2) plans and organizes appropriate instruction,
3) provides for individual differences,
4) employs a variety of instructional strategies and techniques,
5) evaluates student progress and success,
6) demonstrates effective management of the science laboratory,

7) interprets the psychology of learning as it relates to science instruction,

8) organizes a relevant science curriculum.

Interpretations

The eight competencies previously stated appear to be important and necessary for science teachers. These competencies represent a core of essential skills that should be the focus of inservice and preservice science teacher training. The authors also state that it would be significant if the profession could facilitate the achievement of the eight competencies by a majority of secondary science teachers.

ABSTRACTOR'S ANALYSIS

Chiappetta and Collette's study provides an excellent foundation for future research and discussion with regard to the identification of teacher attributes related to effective teaching. However, one questions whether science supervisors are the most competent to identify the skills needed by the secondary science teacher. Furthermore, is a sampling of members of the National Science Supervisors Association a random sample? The response rate of usable returns from the science supervisors for the three rounds was only 45 percent, 57 percent, and 55 percent respectively with over 90 percent of the respondents being males. Also one can infer from the data presented that over 50 percent of the respondents had not taught science after 1971.

While the modified Delphi method was used to improve the validity and reliability of the results, the results do not seem to support this. Furthermore it is disturbing to note that, while the "judges" grouped the responses into three categories 1) cognitive competencies, 2) affective competencies, and 3) personality characteristics, the authors chose to ignore two of the three categories and only focused upon cognitive competencies, thereby negating two-thirds of the categories and approximately 300 statements. Their reasons for ignoring affective competencies and personality characteristics are glaringly weak and distract from the usefulness and credence of the study.

Even concentrating in the cognitive competencies and utilizing the Taxonomy of Cognitive Objectives, the authors made an additional "assumption." The authors assumed that the six taxonomic levels have an equal probability of selection (p. 69). Taxonomic classification indicates that this assumption is not valid and that nearly 80 percent of the cognitive selection is in the bottom three levels.

The Kendall coefficient of concordance (w) was 0.29 ($p \leq 0.001$), which suggests a rather low degree of consistency among the ranking of the 15 competencies (p. 69). Furthermore, the authors state that the first eight competencies appear to be important and necessary for science teachers (p. 70). They base the dividing line because of a "gap" of 50 points between competencies eight and nine. If one examines the data table, other "gaps" can also be detected such as a 27-point "gap" between competencies two and three, and a 115-point "gap" between competencies four and five. The authors have failed to identify which competencies really are the most important. Are there really eight or just two or fifteen or none?

If one examines the first two competencies which are separated by one point they are: 1) formulates instruction that emphasizes the inquiry and process approach to teaching science, and 2) plans and organizes appropriate instruction. One wonders if these are truly "cognitive" competencies as delineated by the Taxonomy of Cognitive Objectives. If these are cognitive competencies, what did the affective and personality competencies look like? It is debatable that these or up to one-half of the competencies listed are really "cognitive" ones.

Future studies should focus upon the identification of secondary science teacher skills utilizing all three domains of learning (cognitive, affective, and psychomotor). Furthermore, researchers must be careful not to structure their studies so that much of the useful data is discarded. In addition, studies involving rank ordering should utilize more sophisticated research procedures for the analysis of the data collected.

Chiappetta, Eugene L. and Alfred T. Collette. "Secondary Science Teacher Skills Identified by Secondary Science Teachers." Science Education, 62(1): 73-78, 1978.
Descriptors--*Educational Research; *Performance Based Teacher Education; Science Education; *Science Teachers; Secondary Education; *Secondary School Science; *Skills; Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Gerald H. Krockover, Purdue University.

Purpose

The purpose of Chiappetta and Collette's study was "to solicit the perceptions of secondary school science teachers to determine those competencies that all secondary school science teachers should possess" (p. 73). It was hoped that, once the competencies were identified, they could be used as part of the secondary science teacher education program.

Rationale

The authors have attempted to identify the needed science teacher skills as perceived by practicing secondary science teachers. The authors recognize that different groups (science supervisors, science educators, science teachers, science students, etc.) place different priorities on the desirable attributes of science teachers.

Research Design and Procedure

Three random samples of 100 secondary science teachers from New York State who teach in grades 7-12 were invited to participate in the study. Each sample was involved in only one of the three parts of the study.

The procedures utilized a modified Delphi method and the research was accomplished using a three-round procedure. In the first round competency statements were generated by participants. In the second round, the participants selected levels on the Taxonomy of Cognitive Objectives at which they felt each competency statement should be specified. In the third round, the participants rank-ordered the competencies in terms of importance.

After round one, the judges (two science education professors, a science supervisor, a supervisor of secondary science teachers, and a science teacher) grouped the responses into three categories: 1) cognitive competencies, 2) affective competencies, and 3) personality characteristics. After round two, one-sample X^2 tests were used to determine the taxonomic levels consistently chosen by the respondents to represent the skills. In round three, the science teachers ranked the cognitive competencies in order of their perceived importance. The Kendall coefficient of concordance was computed to determine the degree of consistency in the ranking of the competencies.

Findings

After round one, 370 competencies were received from 48 secondary science teachers. The panel of judges grouped the responses into the three categories previously mentioned. Fifteen cognitive competencies (34 competency statements) were the focus for the remainder of the study.

In round two, 52 science teachers responded and identified a level on the Taxonomy of Cognitive Objectives at which secondary science teachers should be able to demonstrate a given skill (competency statement). One-sample X^2 tests were used to determine the taxonomic levels consistently chosen by the teachers to represent the skills. Skills or competency statements with associated X^2 values significant at $p \leq 0.05$ were written at the specified levels. Thirty competency statements showed significance in the analyses, while four did not.

In round three, 54 participants ranked the 15 cognitive competencies in order of importance. The Kendall coefficient of concordance was computed to determine the degree of consistency in the ranking of the competencies and a concordance value of 0.26 ($p \leq 0.001$) was obtained.

Competencies ranked numbers one, two, and three appear to be most important to the secondary science teachers. They place human relations first, inquiry teaching second, and employing a variety of instructional techniques third. Competencies numbered 4-12 appear to occupy a moderate degree of importance for secondary science teachers. The areas related to these competencies are the psychology of learning, organizing instruction, the laboratory, evaluating achievement, science and society, a general education, classroom communications, and classroom discipline. Those competencies ranked 13-15 seem to have the least importance. The science teachers apparently see little need for improving teachers' background in the sciences. At the "bottom" of the list of priorities is the area of professional growth and professionalism.

Interpretations

Three of 15 competencies, according to the authors, appear to be most important to science teachers. The authors also expressed surprise that science teachers rank human relations first and professionalism last. The authors also speculate that what is taking place in secondary science classrooms is psychologically stressful for students and teachers. They also state that "science classrooms, in general, probably lack an environment which promotes science as a human endeavor" (p. 76). They conclude that "science educators will have to improve science teachers' attitudes toward becoming more active professionally before secondary science teacher skills can be significantly influenced."

ABSTRACTOR'S ANALYSIS

Many of the criticisms directed at Chiappetta and Collette's science supervisor study also hold for this study. This study asks more questions than it answers and the authors seem surprised by the results. The response rate of returns from the sample for the three rounds is similar to that for the supervisors study (48 percent, 50 percent, and 54 percent respectively). Also over 80 percent of the respondents were male and all were from New York State. Is New York a typical state? The authors define secondary science teaching as grades 7-12 while many other states define secondary science teaching as grades 9-12 and junior high/middle school as grades 5-8. Did the authors lose something in trying to cover grades 7-12 rather than 9-12? Were all grades equally represented in the sample? The authors failed to report this.

The validity and reliability of the modified Delphi method is just as questionable as in the supervisor study along with the authors' use of categories and limitation to cognitive competencies. In this study the authors also assumed that the six taxonomical levels have an equal probability of selection. The Kendall coefficient of concordance (w) was 0.26 ($p \leq 0.001$) which is even lower than the value reported in the science supervisor study.

The grouping of competencies is also as questionable as in the supervisor study. There is a "gap" of 46 points between the two "top" competencies and yet the authors group them together. However there is "only a 51-point gap between competencies 2 and 3." Again, the authors have failed to identify which competencies are really the most important.

As in their previous study, the authors call these competencies "cognitive." How can "providing for a humanly supportive learning environment" (competency 1) be classed as cognitive? Are the competencies identified by either of these studies really cognitive?

The authors were surprised that the secondary science teacher ranks human relations a solid first as opposed to science supervisors who rank inquiry and process approach to teaching science as number one. This indicates the values held by each group.

Even though these two studies have raised more questions than they have answered, they are valuable additions to the science education literature reservoir. They should serve to continue to promote and provoke science education persons from all facets of educational levels to continue to try to answer several questions; "What should be taught to the preservice and inservice science teacher?" "What competencies are of value to the secondary science teacher and can they be identified?" Additional support should be given for further studies that will attempt to delineate those skills that are viewed as basic competencies for the secondary science teacher. We have not yet developed an adequate ranking of skills or competencies needed by secondary science teachers, but it's a beginning.

REFERENCES

Chiappetta, Eugene and Alfred T. Collette. "Science Education Researchers' Perceptions of Skills Necessary for Secondary School Science Teachers." Journal of Research in Science Teaching, 15(3): 233-237, May 1978.

Moore, Kenneth D. "An Assessment of Secondary School Science Teacher Needs." Science Education, 62(3): 339-348, 1978.

Santiesteban, A. Joseph and John J. Koran, Jr. "Acquisition of Science Teaching Skills Through Psychological Modeling and Concomitant Student Learning." Journal of Research in Science Teaching, 14(3): 199-207, May 1977.

Strawitz, Barbara M. "Open-Mindedness and Attitudes About Teaching Science." Journal of Research in Science Teaching, 14(6): 545-549, November 1977.

Chiappetta, Eugene L.; Jay H. Shores; and Alfred T. Collette. "Science Education Researchers' Perceptions of Skills Necessary for Secondary School Science Teachers." Journal of Research in Science Teaching, 15(3):233-237, 1978.

Descriptors--College Teachers; Educational Research; Higher Education; Science Education; Secondary Education; *Secondary School Science; *Secondary School Teachers; *Surveys; *Skills; *Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Robert E. Yager, The University of Iowa.

Purpose

The primary purpose of the study was to improve upon previous attempts to identify specific skills necessary for teaching science in the secondary school. The investigators note the lack of research which clearly identifies skills essential for science teachers -- at least skills where there is general agreement among science educators. They set out to establish such a listing by exploring a reliable method and a valid population for identification of such a set of competencies. The investigators suggest that such a set of competencies or skills will provide "direction for experimental research and for science teacher training." Hence it can be stated that a purpose of the investigation was the establishment of a research base which might influence collegiate programs designed to prepare science teachers.

Rationale

Specification of knowledge and activities "necessary to teaching" is purported to be central to teacher training. The authors report that statements identifying goals, objectives, competencies, and skills have been developed by teacher educators. Few question the value of such statements; however, there is often controversy concerning their validity. The authors state that the controversy arises concerning the studies from which such statements of goals, objectives, competencies, or skills are concluded.

Several references are cited (including Gage, Rosenshine, Balzer, Evans, and Blosser) which suggest a lack of direction to teacher educators for their preparatory programs, the lack of systematic and generalizable studies, and the poor quality of many of the research reports concerned with science teacher education.

Research Design and Procedures

Nearly half of the membership of the National Association for Research in Science Teaching (NARST) was randomly selected. One-third (or 100) of the sample was randomly assigned to each of three groups for involvement in one of three phases of the study.

A modified Delphi technique was used for analyzing the data; different samples of the same population (three groups of 100 NARST members) provided feedback separately at different times and stages in the study. Such a modification purportedly increased the generalizability of the results to science educators, prevented the loss of respondents over time, and eliminated the possibility of manipulating consensus.

In the first round of the investigation competency statements were generated by the first sample of NARST members; in the second round the second sample of NARST members was asked to select levels at which teachers should be able to demonstrate a given skill; in the third round a third sample of NARST members was asked to rank-order the competencies in terms of importance. The perceptions of the three samples of NARST members were considered the independent variables with the products generated at each round the dependent variables.

During round one the randomly selected sample was asked to identify competencies (skills, knowledge, and attitudes) that all science teachers at the secondary school level should possess. Sixty-one persons responded with 432 statements. Five judges were selected for sorting the responses into common categories; cognitive, affective, and personality were the categories selected. Thirty-nine of the statements were assigned to the cognitive category and later grouped into 17 cognitive competencies. Each of these 17 titles was used to organize the competency statements or indicators which specified knowledge and skills perceived as necessary for science teaching. These cognitive competencies were used as the focus for the entire study since "they can be developed through teacher training" and since they comprised the "bulk of usable responses."

During round two the second randomly selected sample was asked to identify a level using Bloom's Taxonomy of Cognitive Objectives for each competency statement. The statements were "stripped of terms" which would link them to one of the taxonomic levels, namely knowledge, comprehension, application, analysis, synthesis, and evaluation. A total of 45 of the 100 persons sampled responded. Statements which were not consistently assigned a single taxonomic level were deleted from the study because such statements would be subject to varied interpretation by the total audience. After round two, 15 cognitive competencies remained.

During round three the third randomly selected sample was provided the list of competencies which remained. Each consisted of a title and one or more competency statements. Fifty-two persons responded and ranked the competencies in order of their perceived importance.

Findings

The major finding of the study was the ranking of the 15 competencies after round three of the study. The rank order was determined by the sums of the ranks assigned to each competency. Two separations in the ranking occurred. The third and fourth competencies are separated by 61 points, and 12 and 13 are separated by 78 points. Hence it can be stated that the first three competencies are definitely the most important with competencies 4 through 12 the next important; competencies 13 through 15 are least important.

The rank order of the competencies is as follows:

1) Provides for a humanly supportive environment;
2) Plans and organizes instruction;
3) Possesses a sound science subject matter background;
4) Relates psychological development with the learning of science subject matter;
5) Demonstrates the ability to communicate effectively;
6) Incorporates effective laboratory activities into instruction;
7) Teaches science using the inquiry, process, and discovery approaches;
8) Uses a variety of instructional strategies and techniques;
9) Provides for individual differences;
10) Employs evaluation skills and procedures;
11) Employs effective classroom management techniques;
12) Relates science to society;
13) Possesses a knowledge of the history and philosophy of science and of its social implications;
14) Possesses a background in mathematics;
15) Knowledgeable about various science curricula.

The Kendall coefficient of concordance was computed to determine the degree of consistency to the ranking of the 15 competencies. A rather low concordance, namely 0.23 ($p < 0.001$) was found. This suggests a low degree of consistency in the ranking by the third sample of respondents.

Interpretations

Since the 15 competencies are similar to those reported using a consensus model, the authors report that "conventional wisdom" has identified many skills that need "validation in the classroom." The authors call for research that will relate teacher competencies to student achievement and student attitude toward science. They call for the development of research instruments which will permit researchers to measure the degree which science teachers can demonstrate the skills.

The investigators report that many of the 15 cognitive competencies can be developed in in-service or graduate programs. Further, they urge that several of the competencies should become the focus for secondary science teacher education. Science teachers competent in the 15 areas will encourage experimental research on teacher effectiveness.

ABSTRACTOR'S ANALYSIS

This study represents a new dimension in assessing major competencies with input from a significant percentage of a major research association in science education. The identification and ranking of 15 necessary competencies for teachers of secondary science is a worthwhile contribution. The results can be compared with other reports — most resulting from creative endeavors, panels of experts, or other less objective means. The fact that the results of the study compare favorably with other attempts at identifying needed competencies for science teachers is reassuring. It also means that the study has not resulted in any really new conceptual contributions for science education.

The use of a modified Delphi technique is of interest and a valuable methodological contribution. Using a significant number of science educators who are members of a research society such as NARST is desirable and affords some unique possibilities for design and analysis. Some opportunities for correcting concerns for reliability and validity in such studies of perceptions were not examined, however.

Several questions arise from the manuscript, the design, the results, the conclusions, and recommendations. Certainly there can be some questions raised concerning the assertion that NARST members are primarily concerned with teacher education and that significant numbers conduct research on teacher preparation. Their depth of understanding of teachers and instruction could be questioned as well. When only slightly over half of 300 randomly selected members chose to respond, other questions concerning their interest in teacher education, depth of understanding, and "unique" suitability for this kind of study arise.

The authors failed to capitalize fully on the modified Delphi technique. They did not show how the use of three separate groups for each phase of the study made the results more generalizable. When nearly 50 percent of the sample was lost in each phase, the claimed "prevention of loss of respondents over rounds" seems like a questionable advantage. Would not a group selected because of their specific involvement in teacher education have been a better choice for reactant sample? A teacher education organization (AETS), a group associated with funded projects for science teacher education (UPSTEP), or a group identified with promising practices could perhaps have provided a better sample group for responding.

The three tasks asked the three sample groups seem very divergent. What is the advantage of having a group generate 432 competencies? Why another group for classifying the cognitive levels? How are the generation activity and the classification activity related to the ranking activity? How do we know that another sample of NARST members would have generated the same results? And yet the findings of the study represent a list of 15 competencies in rank order—by a group of 52 NARST members. What would the 48 NARST members who did not rank the competencies have done? This seems like a legitimate question when the degree of consistency from the 52 respondents is so low. Is it really fair to comment, "One might not expect a higher association to be found among 55 sets of rankings for 15 items"? Perhaps this lack of consistency deserves more attention. Where were the differences? Who were the people? How do their views of teachers and instruction compare? Since it is assumed NARST members have a depth of understanding of essential skills for teachers of secondary school science, is it necessarily so that this "understanding" is the same or even similar?

Is this work of the five-person panel of judges a desirable part of the research design? Would there not be more validity to using another group of NARST members? Why use a science coordinator, a teacher, and a student teacher supervisor as a panel of judges? What was the purpose of the three-category division that the judges proposed? Since all was ignored except for the cognitive competencies, could not the study have started at this point? What is meant by the statement that the cognitive competencies "comprised the bulk of the usable responses"? What is the evidence that cognitive competencies represent "skills which can be developed through teacher training" as opposed to other competencies?

What made the authors feel that the responses in each rank were representative of the perceptions of the target population? Could not a check be made? What about other NARST members? What about non-respondents? Was other information collected on those choosing to respond?

The authors contend that all competencies identified in the study "should be considered by science educators in their research efforts on science teacher competence and science teacher effectiveness."

What is the nature of the other competencies that were proposed? It is fine to state that all should be considered but what is the rationale? This seems especially strange when the authors admit that the 15 competencies are not all that different from a number of other lists that have been generated in other ways. Perhaps a question concerning the value of such global and general lists could be raised. How does one do research knowing that 52 NARST members rated "provides for a humanly supportive environment" as the most important cognitive competency for secondary science teachers?

After choosing to classify and to rank cognitive competencies, the authors conclude that affective and personality characteristics are important variables. Perhaps this whole dimension should have been omitted by definition or it should have been developed more carefully. For example, what was the nature of some competencies in these categories which were identified in round one of the study?

To contend that the 15 cognitive competencies can be developed in extensive in-service or graduate programs seems immaterial. And this is followed with a sentence that eight of the 15 could become a focus for a pre-service program. Why eight? Why the focus at all? Because 52 NARST members agreed to rank the 15 statements?

The closing sentence offers some question. Why will the production of new teachers who have the 15 competencies increase the probability of experimental research on secondary science teacher effectiveness? The rationale needs to be expanded, explained, or illustrated.

Research concerning perceptions of professionals regarding teacher characteristics remains interesting and valuable to a degree. Like so much educational research, more precision and more specificity is needed. We need to know the backgrounds, the positions and the philosophies of a given group of respondents. Wouldn't there be a vast difference in the perceptions of a respondent who had been involved with human relations and one who had not? Such differences could exist -- and probably do -- for each of the 15 competency areas.

We need more rationale -- a better framework from which to conduct our research. Perhaps agreement as to 15 or 25 general competencies is fine. However, to be meaningful and useful these must be better defined and exemplified. To affect teacher education, to influence science teaching, we need to do more than to classify statements and to rank general competencies as to importance.

Butzow, John W. and Zahir Qureshi. "Science Teachers' Competencies: A Practical Approach." Science Education, 62(1): 59-66, 1978.
Descriptors--*Educational Research; Performance Based Education; *Performance Based Teacher Education; Science Education; *Science Teachers; Secondary Education; *Secondary School Science; *Skills; Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by John P. Smith, University of Washington.

Purpose

The purpose of this study by Butzow and Qureshi was to identify and validate observable teaching competencies for use in the preparation of prospective high school science teachers.

Rationale

In the past decade, the growing demand by the public for greater teacher accountability has given rise to attempts by educators at all levels to identify teacher competencies that are both demonstrable and valid. The investigators have stepped into this milieu in an effort to provide a substantial research-based footing for the development of teaching competencies appropriate for use in the education of secondary school science teachers. Most of the effort to develop competencies is being carried on within state departments of education, teacher education institutions, and individual school districts. While this effort is national in scope, it is highly fragmented with little agreement on what competencies to include and what to exclude at the various stages of a teacher's development. In an attempt to go beyond the extent lists of competencies, the investigators have based the development of their proposed competencies on the statements, rankings, and observed behavior of science teachers identified as highly competent.

Research Design and Procedure

The subjects were 21 high school science teachers selected by a stratified random sampling of science teachers identified as being highly competent by a statewide (Maine) random sample of 201 teachers. Stratification was by county with each stratum representing one of the 16 counties in Maine.

The study was conducted in two parts. In part 1, each of the 21 subjects was questioned in an interview about their ideas of a "competent high school science teacher." Competency, as used in the study, was defined for the subject by the interviewer. Following the interview, each subject was requested to rank order the list of competencies he/she had developed during the interview. The audiotape recordings of the interviews were reviewed for additional competencies mentioned but not listed during the interview. In cases where competencies were added to the lists by the investigators after a review of the audiotapes, the amended lists were returned to the respective teachers for re-ranking of all items.

Finally, a list of 12 generic competencies was developed by the investigators from an analysis of transcripts of each of the 21 interviews. The subjects in the study were then asked to rank the 12 generic competencies. Kendall's Coefficient of Concordance (w) was used to calculate the inter-rater agreement of the rankings. The w value of 0.40 was significant at the 0.01 level with a standard deviation of 48.12. A X^2 test was used to determine whether there was a significant difference between the competency rankings. Compared against each other, no significant differences were found between the rankings of adjacent competencies. The competencies are listed below in order from highest to lowest rank.

1. A teacher must know his/her subject and keep striving to update his/her knowledge.
2. A teacher must have a good rapport with his/her students.
3. A teacher must recognize the individual academic abilities of his/her students and try to encourage each of these students.

4. A teacher must make his/her lessons interesting without wasting time.

5. A teacher must be able to control discipline problems in order to protect the learning experience of his/her students.

6. A teacher must plan class lessons in advance with the idea to present scientific concepts and ideas in an organized and clear manner.

7. A teacher must foster unbiased, independent and critical thinking in his/her students.

8. A teacher, especially the one who is teaching low or average ability students, must relate the scientific ideas he/she is teaching to the daily life experiences and the needs of the students.

9. A teacher must evaluate, in a traditional or any other practical sense, the academic progress of his/her students and make the results available to them as soon as possible.

10. A teacher must, appropriately, respond to the sudden diversions of students' thoughts.

11. A teacher must take appropriate actions and instruct students about laboratory safety practices.

12. A teacher must fulfill his/her professional responsibilities.

Part 2 of the study was designed to identify "the extent to which these teachers used these competencies..." Each of the 21 subjects was videotaped for 10-20 minutes while teaching. Most of the videotapes were made on the same day as the interview. Following the videotaping, the investigators divided each of the 21 videotapes into five-minute segments. One five-minute segment was randomly selected from each videotape for viewing by judges. A panel of 14 judges was asked to view each of the 21 five-minute videotape segments and to rate the competencies using a four-point scale of 1 (superior), 2 (satisfactory), 3 (poor demonstration), and 4 (no demonstration). The panel of judges were all members of the New England Region of the Association for the Education of Teachers in Science (AETS). All ratings were made independently by the judg[illegible] The judges' ratings were used to validate the competencies and as a basis for identifying a final list of competencies.

Findings

Inter-rater agreement among the judges was calculated for each of the 12 competencies using Kendall's Coefficient of Concordance: w values ranged from 0.24 to 0.77 and all were significant at the 0.01 level. Standard deviations ranged from 8.94 to 15.82.

Chi-square was used to test the difference between the number of times a competency was demonstrated versus the number of times not demonstrated. Competencies 1 (knowledge of subject), 2 (good rapport), 4 (making lessons interesting), 5 (discipline), and 6 (lesson planning) were all demonstrated at the 0.01 significance level.

Competencies 3 (individual abilities), 7 (critical thinking), 8 (relating ideas), and 12 (professional responsibilities) were non-significant at the 0.05 level.

Competencies 9 (evaluation), 10 (responding to sudden diversions), and 11 (laboratory safety) were not demonstrated at the 0.01 significance level.

Interpretations

The investigators found that a set of highly rated science teaching competencies could be generated by experienced science teachers and judged demonstrable during instruction. It is interesting to note that five of the six highest rated competencies were judged as significantly demonstrated, whereas, three of the four lowest rated competencies were judged significantly not demonstrated. The investigators concluded that pre-service and in-service training programs "should be designed to promote such competencies as primary goals." An extended description of each of the five validated competencies was included in the conclusions.

ABSTRACTOR'S ANALYSIS

The investigators are to be commended for their effort to identify an important but manageable generic set of secondary school science teaching competencies. The teacher-generated list reported in this study, however, is not unique in that each of the 12 competencies indentified may be found in other such lists and is applicable, with the possible exception of laboratory safety, to many other areas of instruction. One must decide, then, what this research contributes to the growing list of products of the nationwide effort to identify for both training and accountability purposes valid and demonstrable teaching competencies. I would agree with the investigators that we will probably never develop a set of competencies appropriate for all situations.

Since the identification of teaching competencies does not lend itself readily to a research design, one must examine with great care the investigators' method for generating the list of competencies in order to make some judgment about their usefulness.

First, the terms competency and validity occur frequently in the research report and are central to the study. Competency, however, is never defined for the reader. The investigators do mention that competency was defined for the 21 teachers interviewed but this definition was not included in the report. Validity, as used in this study, was finally defined for the reader halfway through the Results and Analysis section "as its [competency] significant demonstrability by the teachers and its subsequent ability to be observed by a panel of judges."

Second, it is not clear just how the 21 teachers were chosen nor do we know anything about the characteristics of the group, e.g., science subject areas, number of years in teaching, large school or small school, etc. The investigators indicate that the subjects to be interviewed were selected by a stratified random sampling technique but they do not indicate the size of the group from which the 21 were selected if a larger group was nominated by the nominating teachers. I can only infer that such a group was nominated since the 21 teachers used in the study represented a sample. Although each of the 21 teachers interviewed was identified as being highly competent, the criteria upon which the nominations were based, if other than individual criteria, were not reported. The investigators, also, failed to make clear whether the 201 randomly selected teacher/nominators were all science teachers.

Third, the mechanics of the teacher interviews were not clear. The reader might reasonably ask: "Who conducted the interviews?" "Were the interviews structured or unstructured?" "What were the criteria for ranking the competencies?" Since the final list of 12 generic competencies was derived from an analysis of the 21 interviews and submitted to the teachers for ranking, there is a question in my mind as to why the investigators bothered with asking the teachers to rank the self-generated preliminary competencies during the interviews. The criteria for ranking the 12 generic competencies were not reported by the investigators.

Fourth, it is not clear how the judges who rated the videotapes were selected; but more importantly the investigators fail to report whether the judges received any training in identifying and rating the competencies. Consequently, one does not know the extent to which the judges agreed about classroom exemplars of the competencies prior to actually rating the videotapes.

Fifth, with regard to the videotaping, the investigators fail to report whether the taping covered a variety of settings, e.g., lecture, prelaboratory, laboratory, etc., and place the lesson sequence, e.g., beginning, middle, and end. If the 10-20 minute tapes were made only at the beginning of a lesson and only of lectures, then one might reasonably expect that some important science teaching competencies would not be demonstrated.

Sixth, the procedure for rating the competencies is very sketchily reported. The investigators do not indicate whether the judges viewed the videotapes together or separately; if separately, was the order of the tapes changed for each viewer? Furthermore, the unit of analysis, e.g., time, sentence, phrase, etc., was not identified for the raeder.

One must, also, ask about the frequency counting procedure used. The investigators reported that the demonstration of competency 1 (knowledge of the subject) was rated as superior 63 times, satisfactory 162 times, and poor 36 times. The investigators state, "In other words, the competency was judged as demonstrated 261 times (63+162+36 = 261)." One must ask if each of the ratings was a discrete instance of behavior--in which case the observations could be summed, or was one judge's "superior" rating of a competency another judge's "satisfactory"? If the latter is the case, then, many observations of a competency may have been counted more than once which in turn raises concern for the reliability of the raw data. Furthermore, how does one count 33 non-demonstrations of a behavior? How can one count a competency if it is not demonstrated unless it is made clear to the rater that a certain behavior should occur under certain circumstances and counted as a non-demonstration when it does not occur; if this was done, the investigators did not say so. In my opinion, the reliability of the judges' ratings is quite questionable.

Seventh, it is theoretically inappropriate to use a ranking statistic, i.e., Kendall's Coefficient of Concordance, for measuring agreement among judges' ratings. In effect, the use of this statistic builds in a high number of ties where one has 14 judges, 12 competencies, and only 4 possible rankings (ratings). If such a statistic is used, then, the investigators are obligated to describe what happens to the sampling distribution when a high number of ties are built into the matrix.

Finally, in reporting the results, competency 3 (individual abilities) was left out in the initial discussion of the X^2 results but mentioned later in a restatement of the findings.

The investigators have chosen a worthy, albeit complex subject for their research but have confounded whatever might be of value by vague and incomplete reporting. I am especially concerned about the reliability of the judges' ratings due to the possible, and I think probable, lack of discreteness of each rating and the use of Kendall's Coefficient of Concordance to test inter-rater agreement. Therefore, the statistical base for determining the validity of the competencies (as defined by the investigators) on which the investigators have built their case is highly questionable. The competencies do have face validity, however; in fact, the 12 competencies identified in the study have been rather commonly accepted objectives of science teacher education programs for some time.

Finally, replication of this study would be impossible if one uses only this report. Ill-defined terms and vague research procedures will only lead to the production of similarly confounding results.

Lazarowitz, Reuven and Addison E. Lee. "Measuring Inquiry Attitudes of Secondary Science Teachers." Journal of Research in Science Teaching, 13(5):455-460, 1976.

Descriptors--Attitudes; Educational Research; *Evaluation; *Inquiry Training; *Measurement Instruments; Science Education; Secondary Education; *Secondary School Science; *Teacher Attitudes

and

Lazarowtiz, Reuven. "Does Use of Curriculum Change Teachers' Attitudes Toward Inquiry?" Journal of Research in Science Teaching, 13(6):547-552, 1976.

Descriptors--Changing Attitudes; *Curriculum; Educational Innovation; *Educational Research; *Inquiry Training; Science Education; Secondary Education; *Secondary School Science; *Teacher Attitudes

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Roger G. Olstad, University of Washington.

Purpose

The purpose of these studies was to develop an instrument intended to measure inquiry attitudes of secondary school science teachers, and, using this instrument, to investigate the effect of curriculum usage on science teacher attitudes toward inquiry.

It was hypothesized that science teachers who used "new" programs (defined as BSCS, PSSC, HPP, CHEMS, CBA, and ESCP) would hold a more favorable attitude toward an inquiry approach to science teaching than science teachers who did not use these programs. It was further hypothesized that there would be a relationship between the number of years that the "new" programs had been used by these teachers and their attitude toward an inquiry approach to science teaching.

Rationale

These studies assumed that effective implementation of the "new" curricula required science teachers to use inquiry strategies.

Studies were cited which indicated that there were variations in the implementation of these programs' inquiry-approach base (Amos, Dillon, Gagné, Parakh, Watson), that these variations were often due to teachers' personal traits and philosophies (Black, Gallagher, Montague & Ward, Sadler), that mere use of inquiry science curricula did not assure a positive philosophy with regard to the use of inquiry strategies (Barnes, Gallagher, Montague & Ward), that training did positively effect teacher use (Barnes, Kochendorfer, Ost), and that there was a relationship between the amount of teaching experience and the effective use of these programs (Blankenship, Barnes, Kochendorfer). Most of these studies were conducted prior to 1970.

Research Design and Procedure (Instrument Development)

The instrument, known as Inquiry Science Teaching Strategies (ISTS), was purported to measure inquiry attitudes of secondary school science teachers. Eighty-three items were initially developed with reference to three areas: classroom teacher-student interaction, laboratory investigations, and textbooks used. The items were selected according to their relationship to inquiry teaching which was defined as including student formulation of problems and hypotheses, student designed studies involving data collection and interpretation, and student acceptance of different results for discussion and interpretation in a learning environment where the student is actively participating and the teacher is student-oriented rather than subject-oriented.

Half the statements were positively worded regarding inquiry, the others negatively worded, and they were set up using a typical five-point Likert scale ranging from strongly agree to strongly disagree. Procedures involving a panel of judges (seven professors of science education) and a group of 30 secondary school science teachers reduced the instrument to a total of 40 items and established its content validity.

Construct validity was established by the technique of known groups by obtaining mean scores from five groups to whom the instrument was administered.

Group 1:	7	Professors of science education
Group 2:	16	Graduate students in science education
Group 3:	47	Inservice secondary science teachers having had a 4½ day workshop in inquiry methods
Group 4:	44	Preservice secondary science teachers in their first day of class in the methods course
Group 5:	735	Secondary science teachers from various cities in Texas

Instrument reliability was computed using Kuder-Richardson Formula 20, yielding alpha coefficients of internal consistency for each of the groups.

Findings (Instrument Development)

The mean scores on the instrument by group were as follows: 1) 181.57, 2) 160.68, 3) 146.82, 4) 145.40, 5) 143.27. The authors stated that, "as is to be expected when using this technique (of known groups), the mean score of each group differed according to the level of involvement in science education—i.e., the more the involvement, the higher the mean score."

Tests of significance between pairs of group means showed that Group 1 differed significantly from Groups 2, 3, 4 and 5; Group 2 differed significantly from Groups 3, 4 and 5; and Groups 3, 4 and 5 did not differ significantly from each other.

Reliability (in alpha coefficients) was .54 and .48 for the first two groups, and ranged from .69 to .85 for the latter three groups.

Interpretations (Instrument Development)

The authors concluded that the results support the use of ISTS in determining the inquiry attitudes of secondary school science teachers. They further suggested that ISTS made possible a comparison of the degree of acceptance of the inquiry approach by the individual teacher in relation to known groups and could be used by supervisors or methods course instructors to recognize areas in which teachers need help. This recommendation was based on the construct validity of ISTS and the high alpha coefficient (.85) for Group 5, the most heterogeneous population studied.

Research Design and Procedure (Attitude Change)

The ISTS and a personal data form were distributed (by district science supervisors) to an unstated number of Texas secondary school science teachers. A total of 507 responses were returned to the supervisors for analysis. These teachers represented 30 Texas communities, and most secondary school science subject and grade levels.

Findings (Attitude Change)

It was found that biology, chemistry, and earth science teachers who used the "new" programs had significantly more favorable attitudes toward inquiry strategies than did teachers of those subjects not using these programs. For physics and life science teachers there was no significant difference between users of "new" programs and those not using them.

When compared by years of teaching experience, experienced biology, chemistry, and earth science teachers using "new" programs had significantly more favorable attitudes toward inquiry strategies than had less experienced teachers of these subjects. No significant relationship was found for physics teachers, and there were not enough life science teachers using "new" programs to analyze by years of teaching experience.

The author also analyzed the results on 10 specific items (no rationale provided for the selection) of the inventory, comparing teachers using "new" programs with non-users. Half of the selected users and non-users were significant. It was also found that the degree of rejection of negative items was lower than the degree of acceptance of positive items.

Coefficients of internal consistency of ISTS ranged from .71 to .81 for users and from .66 to .86 for non-users.

Interpretations (Attitude Change)

The author stated that "the results of this study show that secondary science teachers who use new programs in their teaching activities have more favorable attitudes toward inquiry strategies than non-users, and that years of experience in the use of new programs is related to more favorable attitudes toward inquiry strategies. However, it seems that this is not a universal picture. There is no assurance that for each particular element of an inquiry approach teachers will hold a favorable attitude to the same degree. There are no inquiry or noninquiry teachers. Neither the use of new programs nor the length of time they are used assure proper interpretation of the inquiry approach advocated by the programs. The results suggest that teachers need more training in specific competencies of elements of inquiry activities...."

ABSTRACTOR'S ANALYSIS

The attempt to develop an instrument designed to measure science teachers' attitudes toward inquiry is commendable. Whether this instrument measures this quality depends upon what is meant by "attitude." The authors provide no definition of theirs and the reader is left with his/her own understanding of that word. For this reviewer, a definition which has been found useful is "the difference between knowing what to do and doing it." For example (using one of the items from the ISTS), knowing that "students are often capable of designing valid experiments" is different than being skillful in structuring lessons involving student designed experiments, and both of these are different than using lessons involving student designed experiments. It is in the latter case that one gains an insight into the teacher's attitude regarding student designed experiments.

Thus, when the authors write that "ISTS makes possible a comparison of the degree of acceptance of the inquiry approach," it should be clear that what is meant is intellectual acceptance and not necessarily acceptance in teaching practice. Further, when the author states that the results suggest that teachers should have more training in "specific competencies," one should note that training in a competency is not the same as changing an attitude, although the one might lead to the other. (The research results which support the call for more training in specific competencies were not reported in the two papers reviewed.)

The development of the ISTS instrument and the interpretation of its use would have been greatly enhanced had the author(s) examined the relationship of teacher scores to teacher practice. While the authors were careful to develop the content and construct validity of their instrument, the most important kind of validity for such an instrument is its predictive validity. With a high degree of predictive validity, the user (and we readers) would have some confidence in what the instrument was really measuring, e.g. knowledge about inquiry or practice involving the use of inquiry. What the obtained scores predict concerning practice remains an unanswered but vital question.

The title of the second article ("Does Use of Curriculum Change Teachers' Attitudes Toward Inquiry?") poses a question which is never answered, and the research design used does not allow for its answer. The support for the hypothesis that teachers (of only <u>some</u> science subjects) who use "new" programs have more favorable attitudes toward inquiry strategies does not necessarily mean that the use of the "new" programs <u>caused</u> any change in these attitudes. It is equally plausible that these teachers already had more favorable attitudes, and that these more favorable attitudes toward inquiry influenced their selection of curricula, resulting in their use of the "new" programs.

As for the relationship between the number of years of teaching experience of users of "new" programs and their more favorable attitudes toward inquiry strategies, this could be explained in a number of ways in addition to the author's assumption that "teachers who teach the new programs will develop more favorable attitudes toward inquiry...." For example, it could be that longer users of "new" programs had to have more favorable inquiry attitudes to stick with the programs that long. Again, the author's findings do not necessarily support his conclusions.

The author offers no explanation for his finding that biology teachers' attitudes were significantly different between users and non-users, while for life science teachers there was no significant difference, or the finding of no difference in attitude for physics teachers, whether users or not. These "discrepant" findings must be accounted for if one intends to build a generalizable model regarding science teachers' attitudes toward inquiry strategies. (Perhaps the non-significant findings for physics teachers could be accounted for by the fact that the initial selection of the item pool was largely from biological sources.)

The assumption that 44 student teachers in their first day of class in a science methods course have a higher level of involvement in science education than 735 experienced Texas science teachers is questionable. Yet, the authors state that the higher mean score for the student teachers vis-a-vis experienced science teachers is to be expected because of their greater involvement, and upon this assumption rests the determination of the instrument's construct validity.

The recommendation that ISTS be used as a diagnostic instrument in which a science teacher supervisor looks at an individual teacher's responses to individual items on the inventory and proposes assistance in areas where attitudes are lower is an approach which few, if any, science supervisors would find workable unless an extraordinary supervisor-teacher relationship had been previously established. Indeed, in these studies it might be asked to what extent were the responses (were they anonymous?) affected by the use of science supervisors to distribute and collect the instruments?

Finally, for the author's stated conclusions in the second paper (quoted previously), it is left to the reader to interpret the meaning of what is written, and to compare this with the findings as cited.

The relationship of science teacher attitude to practice is important. Future research should take into account research already done regarding teacher attitude and its measurement, much of it generic in nature and not pertaining to science teachers per se. Further, the predictive validity of attitude measures must be attended to by serious researchers in this area.

REFERENCES

Amos, R. "Teachers' Opinions About the Importance of Scientific Method in English Ordinary Level Biology Courses." Journal of Research in Science Teaching, 7:303-313, 1970.

Barnes, L. W. "The Nature and Extent of Laboratory Instruction in Selected High School Biology Classes." Unpublished doctoral dissertation, The University of Texas-Austin, 1966.

Black, H. T. "PSSC Physics in Indiana." The Teachers College Journal, 33:127-129, 1962.

Blankenship, J. W. "An Analysis of Certain Characteristics of Biology Teachers in Relation to Their Reactions to the BSCS Biology Program." Unpublished doctoral dissertation, The University of Texas-Austin, 1969.

Dillon, M. A. "The Status of New Science Curricula in Texas Public Schools." Unpublished Doctoral dissertation, The University of Texas-Austin, 1969.

Gagné, R. M. The Conditions of Learning, 2nd ed. New York: Holt, Rinehart and Winston, 1970.

Gallagher, J. T. "Teacher Variation in Concept Presentation in BSCS Curriculum Programs." BSCS Newsletter, 30:8-19, 1967.

Kochendorfer, L. H. "A Comparative Study of the Classroom Practices and Teaching Rationale of High School Biology Teachers Using Different Curriculum Materials." Unpublished doctoral dissertation, The University of Texas-Austin, 1966.

Montague, E. J. and R. M. Ward. "The Development of Problem-Solving Abilities in Secondary School Chemistry." Journal of Research in Science Teaching, 5:354-357, 1968.

Ost, D. H. "An Evaluation of An Institute for Teachers of Secondary School Biology." The American Biology Teacher, 33:546-548, 1971.

Parakh, J. S. "A Study of Teacher-Pupil Interaction in High School Biology Classes." Journal of Research in Science Teaching, 5:183-192, 1968.

Sadler, P. M. "Teacher Personality Characteristics and Attitudes Concerning PSSC Physics." Journal of Research in Science Teaching, 4:28-29,40, 1967.

Barufaldi, J. P.; J. P. Huntsburger; and R. Lazarowitz. "Changes in Attitude of Preservice Elementary Education Majors Toward Inquiry Teaching Strategies." School Science and Mathematics, 76(5): 420-424, 1976.

Descriptors--*Attitudes; Elementary Education; *Elementary School Science; Educational Research; *Inquiry Training; *Preservice Education; Science Education; Teacher Education; *Teaching Methods

Expanded Abstract and Analysis Prepared Especially for I.S.E. by David R. Stevenson, Chiganois Elementary School, Debert, Nova Scotia.

Purpose

Barufaldi et al. investigated changes in attitudes held by preservice elementary education majors toward inquiry techniques that are used as part of teaching strategies for elementary school science programs.

Rationale

The investigators gave examples of inquiry skill development activities for elementary school science programs and emphasize the role of teacher as stimulator. The research of Allen (1970), Bruce (1971) and Smeraglio and Honigman (1966) was cited in support of behavioral changes in children exposed to stimulating teachers. Although not stated, the inference was made that the direction of the teacher as initiator of inquiry activities plays a significant role in causing desired student behavior and therefore the attitude of the teacher toward the mode of learning should be known.

Research Design and Procedure

Barufaldi et al. conducted a pre- posttest investigation of attitude change, using an elementary science methods course as intervening treatment. The course was stated to be an exploration of facets of modern elementary science programs. The variable being treated was the set of attitudes preservice teachers held toward inquiry techniques.

The researchers used a previously modified form of the Inquiry Science Teaching Strategies (ISTS) as the test instrument. The ISTS is described as containing 40 items, half stated positively (inquiry supporting) and half negatively (non-inquiry supporting). Subjects responded on a Likert-type scale to each statement. A possible score of 200 is indicated. Previous validation yielded reliability scores of 0.41 and 0.83 by Program TESTAT (an item analysis) and a Cronbach α coefficient of 0.96.

From an elective elementary science methods class of 146 students (fall of 1973), 74 students were given the ISTS on the first day of classes (pretest) and the last day of classes (posttest).

Mean scores for subjects and for seven test items were compared using one-way analysis of variance. The item analysis was used to check consistency in attitude change.

Findings

The investigators found the pretest mean score to be 146.14 and the posttest mean score to be 155.41, a difference significant beyond the 0.01 level (F ratio 77.54).

Six of the seven items also showed changes significant beyond 0.01 and the seventh was significantly different beyond the 0.02 level. The researchers noted that change was evident in 13 of 20 items (65 percent) in each of the positively and negatively expressed groups of items.

Interpretations

The authors concluded that the subjects had more favorable attitudes toward inquiry strategies following completion of the course.

ABSTRACTOR'S ANALYSIS

Many educational investigators seem bound by unfortunate research limits. Unyielding parameters also face the physicist who wants to be free of earthly gravity. Nonetheless, all investigators should strive for a purity of setting. The research should be true to the cause it addresses, and bias must be absent or recognized.

Educational investigators would seem to have more difficulty in overcoming biases connected with the research setting than is desirable. The shortage of funds for ordinary research forces the choice between ideal and possible. What can be done is done and sour grapes have to be eaten. The results, however, may be open to too much criticism and yield too little that is fresh and useful. To grasp possible research topics may mean using subjects at hand, either those who can be persuaded to accept a benign treatment or those under the "control" of the investigator. If the latter, they likely come with certain mental sets, prepared if not necessarily positive. They may be aware that part of the routine is completion of an answer sheet, and the instructor happens to teach certain concepts that are within the question-answer context of opening day.

The foregoing is meant to be more realistic than cynical. For it is held to be realistic that a response pattern is stimulated, is likely to be expected, and that self-selecting subjects will cooperate so that concomitant testing will reveal the mental set. A negative or neutral result, on the other hand, would be an unwelcome outcome for a teacher-investigator and an indication of the worth of his methodology.

The analysis need be little more than shallow. Since the turn of the century it has been shown that investigations of classroom techniques will yield desired results regardless of the investigator, his method of testing, and the methodology within his program of instruction. Barufaldi et al. may be added to a long list.

None of the parameters of the study indicate alterations in concept, methodology, or research design from previously reported studies.

Comments made above about the setting of educational research deserve more specific exploration here. Texts on research design caution investigators about sample size and interpretations that flow from research. The cautions should be heeded, and possibly journals should tighten their editorial policies to exclude the obvious and the limited findings. There seems room for criticism with respect to the research under review. At best, it seems that 74 selected students (out of 146—no explanation given) were successful in receiving the message that inquiry teaching strategies are "good" and learnable, or that the attitude set is.

Most educators should not be surprised at the outcome and should wonder what further research will follow and why. It may be of more use to identify the non-learners or recalcitrant students and weed them out of a program that requires inquiry skills and allow their assignment to other roles in life.

Too many articles (this may not be one) flow from doctoral study, bearing the names of several persons, reporting the obvious. It may be time to stem the flow of reports. Even more dramatic would be the upgrading of research in this area of investigation. Perhaps attention could be diverted to the environmental factors that influence the success of graduates in implementing their positive attitudes at the classroom level.

REFERENCES

Allen, Leslie R. "An Evaluation of Certain Cognitive Aspects of the Material Objects Unit of the SCIS Elementary Science Program." Journal of Research in Science Teaching, 7(4):227-281, 1970.

Bruce, Larry R. "A Study of the Relationship Between the SCIS Teachers' Attitude Toward the Teacher-Student Relationship and Question Types." Journal of Research in Science Teaching, 8(2):157-164, 1971.

Smeraglio, Alfred and Fred K. Honigman. "Examining Behavioral Differences Between 'New' and 'Traditional' Methods of Elementary Science Instruction." A Working Guide to the ESS Program. Newton, Massachusetts: Educational Development Corporation, 1966.

Cotten, D. R.; J. J. Evans; and M. S. Tseng. "Relating Skill Acquisition to Science Classroom Teaching Behavior." Journal of Research in Science Teaching, 15(3):187-195, 1978.
Descriptors--Behavior Change; Educational Research; Elementary Education; *Elementary School Science; *Elementary School Teachers; *Inquiry Training; Science Education; *Skill Development; *Teacher Education; *Verbal Communication

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Judy Egelston-Dodd, National Technical Institute for the Deaf, Rochester.

Purpose

The first two authors of this article developed a written format package to teach preservice and inservice elementary teachers how to use inquiry and process-oriented behaviors during science instruction. This study was subsequently evaluating the effectiveness of the model by: 1) comparing the trained teachers' ability to use process skills as measured by paper/pencil test items with an untrained group of teachers, and 2) demonstrating an increase in the use of indirect teaching behaviors by the experimental (trained) teachers.

Rationale

Cutbacks in funding for preservice and inservice training in science education for elementary teachers raised the need for a print format package of instruction to accomplish the same goals. Few elementary teachers are adequately trained in science content, and fewer still receive any special methods instruction specific to the unique affect and processes required for effective science teaching. Thus the end of the NSF institutes and money for consultants needed for revamping science instruction and enhancing the content expertise of elementary teachers represented an appropriate rationale for undertaking instructional development of the inquiry/process skills package.

The study is obviously an evaluation effort being reported under a research label. If the results were used for revision of the package, then the rationale for the study would be strengthened as one of formative research.

Research Design and Procedures

A pre-post test quasi-experimental design was employed for the first question. The equivalence between the two groups of teachers was determined only by pretest scores for criterion variables. Subjects in the experimental group (N = 70) were enrolled in either preservice or inservice science methods training while control subjects (N = 32) were enrolled in a general curriculum course.

The instrument used to determine treatment effects for inquiry/process skill attainment was designed by the developers using a modification of the AAAS-Science Process Measure for Teachers, Form A. The modified instrument has a Cronbach alpha coefficient of 0.86 (N = 23) and a test-retest reliability coefficient (with a seven-week interval) of 0.48 (N = 32). The hand-scored process skills tests based on a rater protocol yielded an interrater reliability of 0.98 (mistakenly reported as interpreter reliability).

The second question was addressed with a change score design using the Instrument for the Analysis of Science Teaching (IAST) Version 2 (Hall, 1972) to collect interaction data during the first twenty-minutes of each class. The first two authors served as observers and had inter-observer reliability coefficients of 0.82, 0.79 and 0.84 (Scott Coefficient Pi) prior to coding the experimental teachers' pre-treatment behaviors.

Teachers were observed and then exposed to the packaged instruction, either singly or in small groups, and performed all the activities in each unit. Upon completion of each unit, teachers were to modify the activities and present them to their elementary students. No mention was made of grade level assignment for these teachers, and it is assumed that all grades were represented.

Teachers were observed and their interactive behavior again coded for each unit. The developers provided some unspecified form of "help" in getting the teachers to use the inquiry process during the first unit but this assistance was "faded out" during subsequent units.

Outside of classroom teaching time, teachers spent from 18-24 hours interacting with the instructional package. At no time did the control teachers read or interact with the materials, and, it is assumed although not specified, that the control group did no teaching of science or any other content area before being posttested.

Before analyzing between group differences on the skill achievement question, the pretest scores were analyzed using a one-way ANOVA to show equivalence on the nine process test scores. Since two tests significantly differentiated the groups, the pre-posttest analysis included a one-way ANOVA for the seven equivalent process tests, and an analysis of covariance was used to compute the F value for the other two.

Analysis of the data collected by classroom observation yielded a series of eight indices (five teacher behaviors and three student behaviors). Correlated t statistics were used to compare pre-and post treatment behaviors.

Findings

The trained teachers outperformed the control group ($p < .05$) on five process measures (observing, classifying, predicting, controlling variables and the composite score). Measuring, communicating, inferring and identifying variables were not different between groups although inferring and identifying variables had been different on the pretest. When the means were adjusted using an analysis of covariance with the pretests as covariates, the experimental group was also higher on the process score for identifying variables. Gain scores were also significantly greater for the experimental group ($p < .05$) on seven process skills (all <u>except</u> measuring and communicating).

Teacher behaviors which were significantly modified by treatment with the inquiry package including lower ratio of closed to open questions and lower percentage of continuous lectures. The other behaviors were unchanged.

A *post hoc* investigation of the questioning behavior using Gallagher's questioning categories (Nelson, 1969) yielded data which were analyzed with a correlated *t* statistic. This analysis revealed that, after interacting with the instructional package, the teachers reduced their use of cognitive memory questions and significantly increased their use of divergent questions.

Student behavior also changed significantly after their teachers had studied the units. Students exhibited a significantly more positive attitude toward science instruction, were involved in a proportionally greater amount of non-verbal activities and more peer interactions, as measured by the interaction analysis instrument (IAST v. 2). Additionally, correlated *t* statistics applied to pre and post treatment data revealed significant increases in the frequency of student group activities and peer interaction during both small group and total class discussion.

Interpretations

The authors suggest there were incongruities between the package strategies and the evaluation instrument. For example, the package emphasized creativity in devising measuring techniques while precision and accuracy were primary criteria on the test which resulted in no significant gains in the treatment group's measuring skills.

The instructional package was judged effective and successful in changing teacher behavior from lecture to inquiry mode.

If teachers are to develop skill in teaching scientific processes using an inquiry approach, they must receive specific instruction toward those goals. There is a real need for instructional packages in print format such as this one.

ABSTRACTOR'S ANALYSIS

The failure of the authors to see the problem of inadequate science instruction as not only one of inappropriate preparation in teaching effectiveness skills but also one of inappropriate materials and instructional strategies is evident in their design. With inquiry oriented instructional materials and strategies, teachers will use an inquiry/process oriented approach and will use more indirect behaviors, ask more questions and have students who behave more independently and achieve higher over a long range of instruction (Egelston, 1971). In their design control group subjects were never allowed to teach the content. It is therefore unknown if interaction with the instructional package (treatment) accounts for the between group differences, or if teaching the content itself would affect criterion scores.

Missing information was the nature and extent of the assistance in using an inquiry approach provided in the first few units and then "faded out." Indeed, this technical assistance could account for the gain scores, behavior changes and between group differences. Documentation of the necessary assistance would be the best formative evaluation for this instructional package.

The report was well written and the study was thoroughly done. In particular the post hoc questions revealed additional detail which helped clarify the results.

It would have been helpful to know the authors' perceptions of the limitations of the category system. Other systems of interaction analysis have been developed and used in science classrooms, and the choice of the IAST should have some rationale.

Finally, the implication that evaluation items for the criterion variables did not match the strategies in the units leads this reviewer to advocate revisions. There was no indication that revisions were planned.

REFERENCES

American Association for the Advancement of Science. Science Process Measure for Teachers: Form A. Lexington, Mass.: Ginn and Company, 1973.

Egelston, J. C. "An Observational Investigation of Classroom Behavior Under the Inductive and Traditional Methods of Teaching High School Biology Laboratories." Unpublished doctoral dissertation, SUNY at Buffalo. Ann Arbor, Michigan: University Microfilms, No. 7121700, 1971.

Hall, G. E. A Manual for Users of the IAST, Version 2: A System of Interaction Analysis. The University of Texas at Austin: The Research and Development Center for Teacher Education, 1972.

Nelson, L. N. The Nature of Teaching: A Collection of Readings. Waltham, Mass.: Ginn and Company, 1969.

Tamir, Pinchas and Hava Zoor. "The Teacher's Image as Reflected by Classroom Experiences." Journal of Biological Education, 11(2): 109-112, 1977.

Descriptors--*Biology; Classroom Communication; *Educational Research; Evaluation; Secondary Education; *Secondary School Science; Statistical Analysis; *Student Teacher Relationship; *Teacher Behavior

Expanded Abstract and Analysis Prepared Especially for I.S.E. by James R. Campbell, St. John's University.

Purpose

The purpose of this study was to determine the instructional priorities which have emerged in the implementation of the Israeli-BSCS Adaptation Biology Program. The authors distributed a questionnaire to a random sample of teachers and another sample of their students to determine the priorities and frequency of occurrence of nine selected instructional practices. Table I lists the instructional practices and Table II outlines the study and the various comparisons which were utilized. The first comparison involved the contrast between how important the teachers viewed the instructional practice (priority) with their perception of just how frequently the practice was utilized. The second comparison involved the student's rating of each practice with his or her perception of how frequently it occurred. A third comparison contrasted the priority ratings of the teachers and those of the students. The final comparison involved the contrast of the reported relative frequency of use of each practice by the teachers and their students. The contrast between what is important and how often it occurred was the central theme of the study.

Rationale

This study is one of several interconnected studies which examined the priorities which emerged for teachers and their students. Previous studies examined the expectations and the objectives of random samples of teachers and students. In a sense it is the emerging methodologies which are examined and analyzed.

Research Design and Procedure

The study used a 22-item two-part questionnaire to obtain the data. The first part of this instrument utilized a nine-item, five-point Likert scale which asked the participants to indicate how important they considered each of the instructional practices.

The second part of this instrument used a five-point scale which required the participant to indicate how often the various practices occurred in their classrooms. For example both groups were asked how important they felt field trips were in this course and then asked how often did they go on such trips or utilize such trips for homework assignments.

The authors reported that the instrument had an alpha (Cronbach) reliability which varied between r = .72 to .87 with the different samples. A stratified random sample was used which included city academic schools, rural kibbutz schools and agricultural schools. Questionnaires were sent to a random sample of 80 teachers and 66 replied (82 percent response rate). The authors also asked the teachers if they would administer the questionnaire to their students. Twenty-four teachers did comply with this request (36 percent response rate) and a total of 624 tenth grade students (age 16 years) completed the questionnaire. One assumption underlying this procedure is that the data from the 24 teachers can be generalized to the 66-teacher sample. The authors did compare the results of the questionnaire of these 24 teachers with the larger group and found no differences.

The statistical method utilized in this study was a series of t-tests between each sample's priority and frequency of occurrence for each of the nine instructional practices.

Findings

1. Students saw the text and its questions and exercises as more important than did the teachers. The students also reported higher

levels of use of these items. In terms of the t-tests, the students rated the importance of the text significantly higher than the frequency of use. (Although both ratings were in the high range.)

2. For interviews and securing information from other sources, teachers saw these activities as more important and occurring more frequently. The disparity between the rated importance and occurrence of other sources was significantly higher for both groups. Both groups prioritized this activity but both reported much lower levels of occurrence. The same finding was found for students with interviews. They valued this activity highly but experienced such interviews much less frequently (.001 level).

3. Teachers found radio, television, lab and field trips more important and had relatively higher levels of use than students. Both groups indicated significantly higher priority levels for these activities than actual occurrence.

4. Teachers rated original assignments and individual investigations as more important and also indicated higher levels of use. Both groups had significantly higher levels of priority for individual investigations than for its frequency of occurrence.

Interpretations

1. The authors concluded that the disparity between teachers and students about classroom events was expected and was useful to the implementation team.

2. The authors reported a high level of use for many conventional instructional practices despite the innovative nature of the curriculum.

3. Lab was perceived as valuable and used frequently by both groups.

4. The two practices involving originality and doing individual assignments and did seem to have the level of occurrence that the team expected. In fact, the low level of occurrence of these items should have implications for the developers of this curriculum.

ABSTRACTOR'S ANALYSIS

This study has several important strengths. First, it is a serious attempt to utilize formative evaluation in the large-scale implementation of an innovative curriculum project. The study is one of several done in comparing the perceived priorities of teachers and students with their perceptions of the frequency of occurrence of related practices. Thus, the priorities which emerge are examined both by the students and their teachers and both viewpoints are contrasted. By doing several related studies it is possible to use each one as a point of reference to the other and arrive at more useful conclusions.

In terms of design, the study utilized a stratified random method in selecting the sample which has to be a strength. The authors rejected the use of the Flanders observation system because of the difficulty in using it with a large random sample. This procedure would seem justifiable particularly when the researcher is not concerned with the affective climate of the dialogue.

One weakness in design which did emerge was the failure to utilize a control group or stratified series of control groups. Since the authors were interested in the implementation at three different settings, then it would have been profitable to have control groups in each setting. Without such baseline data it is difficult to judge the effects the new curriculum is having. These effects are one of the primary reasons for conducting evaluation studies in the first place. By using a control group design the authors would have been able to compare the priorities of their samples with samples of the traditional tenth grade biology classes. What if both groups' priorities were the same? What if many of the practices turned out to occur just as infrequently for both the traditional and the new implemented curriculum? Without a baseline for comparison, none of these questions can be answered and it is not possible to isolate the effects of the implementation process.

Another design problem was the non-random nature of the student sample. By asking teachers to volunteer to test their students the authors could not call this sample a randomly selected one. In fact, the 3 percent response rate caused them to compare the teacher scores of these 24 teachers with the larger sample. Regardless of the negative results of this comparison we can still not generalize the student results to the larger sample with any degree of certitude. The authors should have conducted a detailed analysis of why so few teachers cooperated. Did any one of the three sublevels have higher percentages of respondents? The better way would have been to randomly select students and send the questionnaire directly to them. This procedure would have resulted in two random samples.

Another area closely related to design concerns the instrument used in this study. This two-part instrument was developed from an informal observation of the "transactions" the team saw during the implementation of the program and also from the types of recommendations they made to the participating teachers. This procedure would assure that the instrument had some degree of content validity (although the authors made no such claim). The instrument could have been strengthened by establishing other types of validity. It does seem possible that the frequency of the nine practices could have been observed in a small sample of classrooms and actual tabulations made of just how frequently each occurred. This data could then be used with the data produced by the questionnaire for both teachers and students. This procedure would isolate just how accurate each group's perceptions were of these classroom events.

What if students or teachers were inaccurate by 20 or 30 percent? Is it possible that the students were more accurate observers than the teachers? Such data are important if we are to make conclusions in relation to perceived priorities.

In terms of reliability, the authors reported using an alpha reliability for each of their samples. This method of determining reliability usually is performed on just one trial administration for each sample. The authors don't mention any other mechanisms of determining the reliability of the questionnaire. One problem with using only the one administration for the subjects is that the individual's day-to-day variations are not considered. Such variations do occur for many types of tests and are important to control for in the construction of any instrument. A test-retest method is frequently used to isolate this factor. It could have been used to augment the alpha reliability which was used. By using the two procedures the authors would strengthen their claim for developing an instrument which produced consistent results.

Another, more serious problem concerns the selection of items for the test. The authors don't mention any trial administration or any item analysis in establishing the ability of each item to discriminate.

In Part II most of the frequency data were based on just one item. Only three of the practices used two items to determine the frequency of occurrence. Perhaps the frequency of each practice could have been determined by several homogeneous items. It is interesting to note that in one of the three practices that contained two items there were no significant differences for one item and significant differences for the other. If both items were supposed to isolate the same practice, how can such difference occur?

Another problem with the instrument was no mention in the research report of the directions utilized by the samples. These data are always useful for a reviewer to analyze. This reviewer had a final question about the difference between the affective part of the test and the observational part. In all the results the two parts are scaled on the same point value but the question arises: Are the two value systems really equivalent? Can one really equate neutral importance to some degree of frequency of one or more of the practices? Perhaps the analyses should have remained separated for both parts of the test.

Despite these weaknesses the instrument is a blend of high inference items (affective portion) and low inference items (observational part). Such a blending was first recommended by Rosenshine et al. (1971) and must be considered a strength.

In summary, this report of a formative evaluation study shows the state of the art of researchers in this area. The emerging priorities of any curriculum are important items to isolate and measure. Hopefully more and more school systems will utilize evaluation procedures to isolate the effects of each new implemented curricula.

REFERENCES

Rosenshine, Barak and Norma Furst. "Research on Teacher Performance Criteria," Chapter 3. Research in Teacher Education—A Symposium, B. O. Smith (ed.). Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1971.

I N D I V I D U A L   S T U D I E S

Lowery, Lawrence, and William H. Leonard. "A Comparison of Questioning Styles Among Four Widely Used High School Biology Textbooks." Journal of Research in Science Teaching, 15(1):1-10, 1978.
Descriptors--*Biology; Educational Research; Instructional Materials; *Questioning Techniques; Science Course Improvement Project; Secondary Education; *Secondary School Science; Teaching Methods; *Textbook Evaluation; *Textbooks

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Linda R. DeTure, University of Florida.

Purpose

The purpose of this study was to examine the questioning styles of the four high school biology texts more frequently used for college-bound students. Types of questions, including experiential and nonexperiential, frequency and placement of questions within the texts were compared. For the experimental questions the kind of inquiry processes elicited by the questions was examined. The hypothesis was that there would be no quantitative differences in the questioning styles of the four texts.

Rationale

A basic assumption is that questioning by either the teacher or the text plays an important role in eliciting inquiry behaviors in students. Research indicates that the kinds of questions asked during instruction determine the types of operations students perform. Because many teachers rely heavily on the textbook as a program guide, it is appropriate to examine the ability of in-text questions to prompt student inquiry. The authors state that little or no research has been focused on relating textbook questions to the process of inquiry or to science teaching in general. Studies examining questions in written materials have not been applied to commonly used biology textbooks.

Research Design and Procedure

The texts selected for the study were 1973 editions of the four most widely used high school biology texts for college preparatory biology classes. The texts served as the independent variables. The dependent variables were the ratio of questions to sentences per page and the frequencies of types of questions tallied according to the Textbook Questioning Strategies Assessment Instrument (TOSAI) developed by the Cooperative Teacher Preparation Project. A 10-percent random sample of pages from each text was studied. From each page a ratio of questions to sentences and the mean number of questions were computed.

The questions were classified as experimental or nonexperiential, which required the students to focus on phenomena not previously experienced. The experiential questions were classified as rhetorical, direct information, focusing, open-ended and valuing. The nonexperimental questions were identified according to the following inquiring processes: observing, communicating, comparing, organizing, experimenting, inferring and applying. Additionally, the questions were examined by placement in the text: initiatory, contextual, terminal and captional.

For analysis the questioning styles and frequencies were converted to percentages for easy comparison among the textbooks. Only the overall frequency of questions per page was statistically analyzed for significance and this was done by t-test. All pairwise contrasts for the four texts were tested. The particular questioning styles of a given text were compared but were not analyzed by statistical tests.

The variables evaluated were frequency of experiential questions, placement of questions, types of experiential questions and patterns related to the science/learning process questions. Other differences noted by the authors were placement of the laboratory investigations and the general distribution of questions within the text. For example, the Modern Biology text asked the most questions in the first few chapters. Interestingly, it was noted that the level of sentence complexity and reading difficulty were similar for all the texts except for the BSCS Blue Version which was more complex.

Findings

Except for the t-test analysis of the ratio of questions to sentences per page, all the results are observational rather than statistical. The t-test results indicate significant differences between all pair-wise contrasts of the four texts except for the Modern Biology versus the BSCS Blue Version in which the difference was not significant.

With regard to the other variables examined, the following evaluative differences were noted. Comparing the percentages of experiential versus non-experimental questions, the BSCS Green Version (BSCS-GV) has 71 percent experiential question; the Modern Biology (MB) 31 percent; BSCS Yellow Version (BSCS-YV) 52 percent; and the BSCS Blue Version (BSCS-BV) 58 percent. For the placement of question the MB text places the four categories about evenly; BSCS-GV has a high number of captional questions and few terminal questions. BSCS-YV asks few terminal questions. All except the BSCS-GV place questions in context. Of the five types of experiential questions all texts used focusing questions extensively. MB asked the most rhetorical questions and the least valuing. BSCS-BV asked fewer direct information questions than the other texts but more open-ended and indirect information questions. Regarding the process-type questions essentially no observing questions and few organizing and experimenting questions were asked. However, the laboratory investigations were not included in the study. Most of the BSCS-GV process questions involved communicating and applying while the BSCS-BV stressed inferring questions.

Interpretations

Significant differences between the ratio of question to sentences were found for all t-test pairwise comparisons except for Modern Biology versus BSCS Blue Version. The ratio of question to sentence was highest for BSCS Green Version followed by BSCS Yellow Version, BSCS Blue Version and Modern Biology.

The value of the particular questioning patterns of each textbook is dependent upon the assumptions made about questioning. If examined from a Piagetian focus, which indicates that direct and active experience is valuable to learning, the BSCS-GV would be most desirable since it has the most experientially based questions. If asking a high proportion of questions is important, then BSCS Green and Yellow are valuable.

If the position of questions is a critical factor, beginning or end rather than contextual, then none of the texts is desirable. The effect of captional questions (BSCS-GV) has not been researched. All texts ask a rather low number of "higher level questions" and all are relatively deficient in asking higher order inquiry-process questions. If balance of types of process questions is important, then again all texts would rank low.

The authors express a need for more research to clarify the relative values of certain types of questions. However, because questions are considered to be stimuli to inquiry, it is felt that examining in-text questions is a legitimate, but neglected, basis for evaluating textbooks.

ABSTRACTOR'S ANALYSIS

This study provides an interesting and useful addition to the area of research related to questioning. Although it is not truly an experimental study, Lowery and Leonard examine question types and patterns that have been dealt with experimentally in other studies. In particular, they look at in-text questions in regard to question type and to position or placement of questions within the written material as well as the frequency of questions to sentences per page. Research related to the former suggest that it would be appropriate to examine the textbook and its relationship to the processes of inquiry.

Conceptually what this study might do is provide a framework for identifying research questions that would require further experimental examination. When discussing the implications of the study, the authors defined the desirability of the four texts in accordance with the assumptions of related questioning research. For example, the placement of questions research indicates that questions placed at the beginning or end of paragraphs have greater value to readers than those in context and, therefore, none of the texts is desirable since none place their questions at the beginning or the end of paragraphs. These various assumptions may provide a rationale for future research issues, but extreme caution should be used if generalizations are made that would bridge the gap between in-text questions and questioning research that deals with learning outcomes. Essentially the question of textbook questions and effective learning outcomes has not yet been asked experimentally.

The contribution of this study is that it simply looks at an area of questioning in frequent use, i.e., the textbook, and provides some baseline data which is, in itself, necessary. It is time that someone looked at questions, where and how they are as opposed to how they should be.

Because this was not actually an experimental study, in that it had no manipulated variables, it had no experimental design. The categories for classifying questions were established and the questions were subsequently assigned to the various categories and tabulated for the four textbooks. The only statistical test employed was a t-test for the pairwise comparison of the ratio of questions to sentences per page for the four texts. Frequency tables were used to report the results of types of questions asked (rhetorical, focusing, etc.) and kind of process of learning questions (observing, organizing, inferring, etc.).

The authors make a case for not subjecting the differences between categories to statistical analysis because of an assumption that a 10 percent sampling is representative of the whole textbook and, thus, the question of difference is an evaluative judgment rather than a statistical one. There is some room for disagreement with this idea. Any time a sample is used rather than the whole population, there is the chance that it is not representative of the whole population. The reason for using a statistical test to begin with is to give the researcher an idea of how big that chance may be. The authors stated that the number of questions in the textbooks were concentrated in different sections of the books and that the placement of concentration varied among the four textbooks. It then may not be valid to assume that the placement of question types is necessarily uniformly spaced throughout the texts. On the contrary, it would appear that a systematic approach for inserting questions into the textbook was not used at all, thus making it necessary to look at the differences statistically as well as evaluatively.

One other question that could be raised regarding the statistical analysis concerns the use of pairwise t-test rather than a one-way analysis of variance. It would be interesting to know whether or not the differences found to be significant would hold up in cases where the t value was low in a slightly more rigorous test.

The report was well written and was organized in a way to answer the basic questions usually asked of a research paper. The authors did not attempt to generalize beyond the findings which were appropriate in light of the fact that most of the findings were not subjected to statistical analysis. The value of the study is that it seems to be the first to examine the role of the textbook and its relation to the processes of inquiry. Most studies of questioning have used specially prepared materials and did not offer suggestions for textbook.

application. This study examines in-text questions and asks how the research could be applied.

Herein lies the suggestion for future research. The textbook questions could be experimentally manipulated for examining effectiveness on learning outcomes. It would be worthwhile to know how the textbook is used by the teacher to enhance the learning process. If textbooks are infrequently used, it might not matter about the types and number of questions incorporated. However, if the text is central to the instructional process, the question of questioning becomes very important. It would be nice for writers and researchers to get together to design texts that would use questions appropriately and effectively.