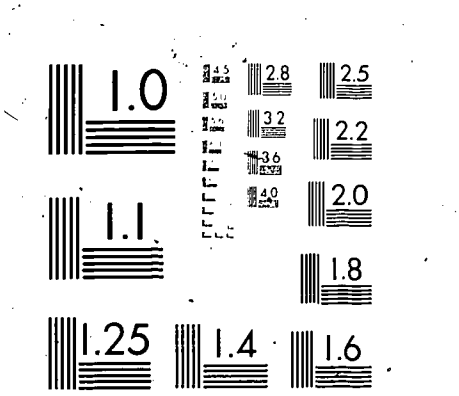

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

DOCUMENT RESUME

ED 168 889 SE 027 521

AUTHOR Blosser, Patricia E., Ed.; Steiner, Robert L., Ed.
TITLE Investigations in Science Education, Vol. 4, No. 3. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 78
NOTE 71p.
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 *Abstracts; *Educational Research; *Instruction; *Problem Solving; Process Education; Research; Research Methodology; Science Education; *Teacher Education

ABSTRACT

This issue of "Investigations In Science Education" (ISE) contains abstracts and critical analyses, prepared by science educatcrs, of research reports published in professional journals. The 11 reports discussed involve instruction, problem solving, process skill development, teacher education and behavior, and attitudes. Each abstract includes bibliographical data, research design and procedure, purpose, research rationale, and an abstractor's analysis of the research. (PB)

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC® Clearinghouse for Science, Mathematics, and Environmental Education

Published Quarterly by

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

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

NOTES . . .
from the Editors

Volume 4 No. 3 of INVESTIGATIONS IN SCIENCE EDUCATION contains abstracts and critiques of eleven journal articles. No attempt has been made to cluster these articles although three of them refer to achievement, in some aspect, and several others focus on teachers--their verbal behaviors, methods for educating, or characteristics to be considered when evaluating teacher performance. Readers may identify other aspects of interest that could serve to relate one article to another.

As Stan Helgeson has written in past "Notes," we continue to invite the readers of I.S.E. to make use of this publication as a vehicle for dialogue. If you have suggestions for methods for accomplishing this objective, please share them with us.

Patricia E. Blosser
Editor

Robert L. Steiner
Associate Editor

Volume 4, Number 3, 1978

Anderson, Elaine J.; Hermes T. Demelo; Michael Szabo; and George Toth. "Behavioral Objectives, Science Processes, and Learning from Inquiry-oriented Instructional Materials." Science Education, 59(2):263-271, 1975.

Descriptors--*Academic Achievement; *Behavioral Objectives; Biology; Cognitive Processes; *Educational Research; Inquiry Training; *Instruction; Learning Processes; Science Education; *Secondary School Science

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Claudia Douglass, Central Michigan University.

Purpose

The primary purpose of the research was to investigate the effect of behavioral objectives on immediate learning from an inquiry-oriented instructional program. The achievement on knowledge level and higher than knowledge level items of a posttest was compared for students who scored high and students who scored low on a test of ability to use scientific processes. The interaction of a student's ability to use process skills and the presence of objectives was also investigated.

The specific hypotheses were as follows:

1. there is no significant difference between learners who receive behavioral objectives and those who receive a placebo using the criterion of specific subject matter achievement,
2. there is no significant difference between learners who score low or high on a test of ability to use scientific processes using the criterion of specific subject matter achievement, and
3. there is no significant interaction between the treatment conditions (behavioral objectives or placebo) and learner facility with scientific processes using the criterion of specific subject matter achievement.

Rationale

Research into the effects of behaviorally stated objectives on learning has produced mixed results. Many investigations have shown favorable outcomes in terms of achievement and retention while others have shown negative results with the addition of objectives. Gathering information from many studies cited in this paper, Anderson et al. claimed that a confounding factor in the literature reports was that neither the cognitive level of the objectives nor the criterion measures had been held constant.

When knowledge, comprehension, application, and analysis level objectives were identified and employed with programmed instruction, significant differences in achievement were noted between the levels of comprehension and application. However, no difference was demonstrated between knowledge and comprehension levels (Stedman, 1970). Anderson and her colleagues designed a controlled study which in part demonstrated the role of the cognitive level of the objectives. Their intent was to carry the work of Stedman (1970) a step farther.

Research Design and Procedure

A posttest-only control group design was employed (Campbell and Stanley, 1963). The independent variables were the treatment conditions (receiving objectives or receiving a placebo) and facility with scientific processes (high or low). The dependent variable was a subject matter achievement test.

The subjects were 40 students enrolled in an elementary science methods course in a major Eastern university.

The evaluation instruments were the Process of Science Test (POST) and a subject matter achievement test designed to contain knowledge level and higher than knowledge level questions. The overall reliability of this test was $KR_{21} = .77$, while the reliability for those 16 items

measuring knowledge level cognition was KR_{20} = .65, and those 16 items measuring higher than knowledge level cognition was KR_{20} = .60.

The instructional material was a short unit based upon the BSCS Inquiry Slide Set Structure and Function: Control of Blood Sugar: A Homeostatic Mechanism. A set of 24 behavioral objectives (16 knowledge and 8 higher than knowledge level) was developed and validated by a panel of judges. These were incorporated into an introductory passage which was given to the students just prior to the instruction. The control group received a passage containing an unrelated discussion of a newly developed science curriculum.

Two weeks prior to the treatment the subjects completed the POST. It was scored and the subjects were divided into two groups based on the median score. Then the students were randomly assigned to treatment groups and administered the treatment passage. Following no introduction, they participated in a 75-minute class on homeostasis conducted by an experienced biology teacher.

Findings

The study showed that there was a significant difference in criterion subject matter achievement between students who received behavioral objectives and those who received a placebo ($F = 4.7$, $df = 1,37$, $p < .05$). Learners who scored high on the POST did significantly better on the criterion subject matter achievement measure than students who scored low on the POST ($F = 12.4$, $df = 1,37$, $p < .01$). Finally, the interaction between the treatment and the ability to use scientific processes was not significant (F value not reported). A table was presented which showed without statistical proof that students did better on the knowledge level subtest of the criterion measure than on the higher than knowledge level subtest, maintaining the same patterns with regard to ability to use science processes and treatment group performance (Table III).

Interpretation

The results of the study suggested that inquiry-oriented biology materials that incorporate behaviorally stated objectives enhanced immediate learning and that facility with scientific process skills increased student performance on the criterion achievement measure. The researchers concluded that the study reinforced the assumption that objectives enabled learners to retain essential material and to organize complex cognitive processes and that the degree of facilitative effect of the objectives was probably related to the quality and quantity of the instructional material.

ABSTRACTOR'S ANALYSIS

It has been stated several times that research dealing with behavioral objectives has been inconclusive (Duchastel and Merrill, 1973; Rowe and DeTure, 1975). Rowe and DeTure (1975) recommended that if one were to pursue research dealing with objectives, he/she should refine the topic or conceptualize it in a different way. Anderson et al. did just that by incorporating the notion of cognitive level and the investigation of an interaction. The rationale for the study would have been strengthened if, rather than dwelling on the current state of confusion, literature dealing with the reasons for expecting an interaction and the inclusion of cognitive levels and science process abilities had been incorporated.

The research methodology was the weakest aspect of the study. Improvements could have been made in the choice of the sample, the statement of validity of the POST, and the reporting of the results. A more appropriate sample would have been one which had no previous knowledge of behavioral objectives. Properly trained preservice elementary teachers should be able to recognize the objectives and their purpose. A sample of individuals not familiar with the purpose of objectives would have been more generalizable.

The POST was used to establish independent variable grouping, however, validity and reliability data were lacking from the report. A review

of this instrument indicated that the only norms for it were established for tenth grade students although later it was deemed suitable for students in grades 10-12. It was noted in that same review that several of the items tested general intelligence and reasoning ability rather than the actual processes of science (Buros, 1972). It would be interesting to know the correlation between intelligence and the POST. If the correlation was between .4 and .6, the results may be confounded and an analysis of covariance would be more appropriate (Cochran, 1957).

A minor point dealing with the reported reliabilities of the achievement tests was that a KR_{21} was used as the reliability figure for the entire test while a KR_{20} was used for the subscore portions. To improve continuity, it would have been more appropriate to use the same measure in both cases. The author of this review was unable to find justification for this procedure and none was given by Anderson et al. (Richardson and Kuder, 1939; Thorndike, 1971). The results were analyzed and presented clearly. A two-way analysis of variance was an appropriate statistical tool to use; however, a complete ANOVA table including the interaction values would have been more useful. A technical point regarding the results is that the probability values reported at the bottom of Table I were reversed and should have read $p < .05$ and $p < .01$. The reporting of results such as those shown in Table III of Anderson's paper is to be complemented. So often supportive data which do not meet the strictest research criteria are omitted from the literature and other investigators in the area never gain access to the information. The choice of design, specifically one which omitted a pretest, was most appropriate in that a pretest x treatment interaction likely would have appeared.

The results were generalized beyond the sample. It was implied that they supported the work of others regarding immediate learning and retention. This study did not deal with retention in any way and this discussion was out of place. The results and implications of the research conducted in this study will help clarify the current state of the literature in the area of behavioral objectives and their meaning. The recommendations for further research regarding

this particular study are that it be repeated with a larger and more appropriate sample, and that the instructional period be extended. Carefully controlled and conducted studies are always of value in science education and especially in an area such as that addressed by Anderson and her colleagues.

REFERENCES

Buros, O. K. The Seventh Mental Measurements Yearbook. Highland Park, NJ: Gryphon Press, 1972.

Campbell, D. T. and J. C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963.

Cochran, W. G. "Analysis of Covariance: Its Nature and Use." Biometrics, 13:261-281, 1957.

Duchastel, P. C. and P. F. Merrill. "The Effects of Behavioral Objectives on Learning: A Review of Empirical Studies." Review of Educational Research, 43(1):53-70, 1973.

Richardson, M. W. and G. F. Kuder. "The Calculation of Test Reliability Coefficients Based on the Method of Rational Equivalence." Journal of Educational Psychology, 30:681-687, 1939.

Rowe, M. B. and L. DeTure. A Summary of Research in Science Education-1973. New York: John Wiley and Sons, 1975.

Thorndike, R. L. (ed.). Educational Measurement. Washington, DC: American Council on Education, 1971.

Brown, L. M.; P. W. Tweeten; and C. Pacheco. "Attitudinal Differences Among Junior High School Students, Teachers, and Parents on Topics of Current Interest." Science Education, 59(4):467-473, 1975.
Descriptors--*Attitudes; *Biological Sciences; *Educational Research; Parent Attitudes; Secondary Education; *Secondary School Science; *Social Factors; Student Attitudes; Teacher Attitudes

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Dorothy Gabel, Indiana University.

Purpose

This study examines the differences between adolescent, teacher, and parent attitudes on topics of current interest. Differences according to sex are also studied.

Rationale

Brown (1973) has shown that science teachers tend to de-emphasize concepts that are of social nature which are relevant to adolescents. If students' values differ from those of adults, schools may wish to provide students the opportunity to clarify their values.

No references to other research in this area are given by the authors.

Research Design and Procedure

This study consists of a survey of junior high school students, teachers, and parents on their attitudes toward socially relevant ideas. The survey was conducted by students in two seventh grade life science classes at an upper-middle class, predominantly white junior high school in Albuquerque, New Mexico. The 60 junior high school students were paired and each pair interviewed three groups of people: (1) three male and three female junior high school students, ages 11 to 14; (2) two teachers, male or female; and (3) four parents, two mothers and two

fathers. Although this should have resulted in 360 interviews, only 125 interviews were used in the analysis.

The interviews consisted of students showing those being interviewed a picture card (10 in all) depicting an item of societal importance. Topics included marijuana, police, school, baby, atom bomb, smog, whiskey, teacher, army, and cigarettes. The person being interviewed responded by rating the idea presented as (1) very good, (2) somewhat good, (3) so-so, (4) somewhat bad, or (5) very bad.

Data were analyzed using a one-way analysis of variance on the scores for each topic according to the three classifications of those interviewed: student, teacher, or parent.

Findings

The investigators reported the following findings:

1. No statistically significant difference in scores between males and females, using analyses of variance, were reported.

2. The scores of teachers and parents were significantly different from those of students on marijuana, police, school, baby, teacher, and cigarettes as analyzed by Tukey's Honestly Significant Difference Test (HSD). The only significant difference between teachers' and parents' scores was on their view of the teacher.

Interpretations

Adolescents perceive certain segments of current social issues differently from the adults with whom they associate. Teachers should be cognizant of this fact and provide experiences in the curriculum to help students develop their value system.

ABSTRACTOR'S ANALYSIS

This study should be classified as a report on a school project conducted by students rather than a research report. The students who participated in this worthwhile project probably benefitted more than those persons interested in research in this area.

There are several flaws in the research design that makes one question both the internal and external validity of the survey.

1. The sample of this study consists of persons interviewed by students that fall into categories of adolescents, teachers, and parents. Little is known about the sample. No attempt was made to interview a random sample. Each pair of students was to interview 12 persons. Only 125 out of a possible 360 were interviewed. This would certainly bias the sample as some students may be interviewing different numbers of persons than others.

2. The techniques used for interviews were not uniform. There is little evidence that all students used the same technique or the person interviewed would respond in the same way to different interviewers (low reliability).

3. Students interviewed teachers and parents who were known to them and with whom they closely associated. Responses of adults to children could very easily be biased in that the adults gave responses they wished children to emulate.

In addition to the above, the authors give very little rationale for the study and do not place it in a theoretical framework. In fact, only one reference is made to work done by others.

The report is quite straight-forward and easy to understand. Because of the unequal numbers of persons in each category, Table VII becomes a little difficult to comprehend. Perhaps these numbers would be best

represented as percentages so the reader could compare the proportion of responses in each category.

The one value that this report has is that it brings to the attention of the reader that children do not share the same values as adults and that this should be a consideration in their education. Most teachers probably would not consider this new information, however, even though many may not include societal issues in their teaching.

REFERENCES

Brown, L. M. "Relationships Between Progressivism, Traditionalism, Dogmatism, and Philosophical Consistency in Selected Urban Secondary and Elementary School Teachers." Unpublished Ph.D. dissertation, University of New Mexico, 1973.

Deiter, Donn L. and Paul B. Hounshell. "Biology Teacher Evaluation: Factors Significant in the Process." Journal of Research in Science Teaching, 12(2):139-146, 1975.

Descriptors--*Bias; *Biology; Educational Research; Evaluation; *Evaluation Criteria; Science Education; *Science Teachers; *Teacher Characteristics

Expanded Abstract and Analysis Prepared Especially for I.S.E. by John E. Lutz, National Technical Institute for the Deaf and Rochester Institute of Technology.

Purpose

According to the authors, the intent of this study was to gain some understanding about the evaluation of biology teachers.

Rationale

The major assumptions identified by the authors for this study were:

1. Criteria exist which characterize outstanding biology teachers.
2. OBTA (Outstanding Biology Teacher Award) judges are appointed for their ability to recognize outstanding biology teaching and these persons are competent to judge biology teachers.
3. Judges employ these criteria when making decisions about the suitability of candidates for the award.
4. A "value hierarchy" of criteria exists, with some criteria being of more importance than others in evaluating candidates for the award.
5. Assessment of candidates for the award is subjective and individual judges value specific criteria differently.

Research Design and Procedure

This study was a descriptive evaluation of the traits and characteristics of outstanding biology teachers as determined by OBTA program judges. It was planned to determine the importance of various evaluation criteria used to define competent biology teachers.

The study involved 220 judges from the 47-state selection committees active in the 1970 OBTA program. A questionnaire was developed that included 111 items which might have been used as criteria in the teacher evaluation process. Each of the judges was asked to rate the various criteria according to the weight given them when evaluating candidates for the award. Completed questionnaires were received from 179 judges.

Chi-square analyses assessed the variance between the observed and expected response frequencies. Comparisons of ratings were also made between different judge groups.

Findings

Twenty-one factors were identified as significantly important from the study and could be grouped into three major areas related to teachers' intrinsic personal traits, teacher-student interrelationships, and concerns for skills and proficiencies as a science teacher. Eight factors were rated significantly different between different judge-groups.

Interpretations

The authors stated that they did not want their conclusions to be interpreted as judgments of the OBTA program, although they saw it as a model for teacher evaluation by a professional group. They concluded that specific criteria exist which are significant to competent judges and that various judges valued certain criteria differently, depending upon the judge's occupational role.

ABSTRACTOR'S ANALYSIS

The differences between research reports that command confidence and those that suggest reservations can be explained with these criteria:

1) clearly defined purpose, 2) detailed design and procedures, 3) estimation of flaws, 4) adequate data analysis, and 5) justifiable conclusions. This report will be analyzed for these criteria.

1. The stated purpose of any study should include a brief description of where the study fits into the general structure of knowledge; that is, a theoretical framework gleaned from the literature should show its need and importance. This report was adequate in this regard and included brief statements of concerns, literature review, and assumptions for the study.

2. Although a section of the report was labelled "Study Design," a specific pre- or non-experimental design was not identified. Apparently, the descriptive survey discussed by Novak (1963) could be assumed as the model for this study. Furthermore, a precise and clearly worded description of procedures was lacking. The reader had to piece together the procedures from several different sections and then still lacked a clear understanding of the instructions given to the OBTA judges and how evaluator bias was really determined. A test of the adequacy of design and procedural description lies in the replicability of the study from the directions given in the original report; a replication from this report probably would not be possible.

3. No mention was made of any constraints or limitations of the study. An estimation of the effects of probable limiting factors would have been useful to help define the generalizability of the findings. Although the questionnaire designed for the study was based upon several apparently valid sources, no confirmation of content validity was given and no discussion of other measurement concerns was included. Although the return rate of the questionnaire was 81 percent (a fine return), no discussion of the potential selective bias of the returns was included.

4. Since this study investigated possible differences between distributions and not between parameters, the appropriate statistics were nonparametric. The chi-square test for independence

permitted a relatively easy method of data collection and did not require assumptions about the population. Apparently, the 21 items identified as significant were sorted according to some logical grouping—the inclusion of a formal factor analysis might have helped partition these significant criteria into sets for interpretation.

A major presentation of data was accomplished through Table III; however, this presentation might have been more useful if it had been organized differently. For example, the significant criteria could have been listed in the first column, and the results from the various judge groups could have followed in succeeding columns with the response categories collapsed into "Important" and "Not Important." Comparisons of response frequencies for a given criterion could then be made more easily between the various groups.

5. The broadly stated intent of this study was "to discover something about the evaluation of biology teachers." This intent was fulfilled. Within the framework of the listed assumptions and the analyzed data, the conclusions were justifiable.

General Comments

Perhaps the greatest frustration for the reader was the inconsistency of the report's logic. Although research questions were identified and assumptions presented, the relationships between the two were not clearly specified. In fact, there was inconsistency that required reader interpretation among the study's five assumptions, three research questions, four expectations, and three null hypotheses.

Although 21 items were found to have been rated significantly high, further work is needed to validate these items as desirable competencies of biology teachers. They do represent a very useful foundation for additional research.

Finally, it is often easier to criticize another's work than to generate one's own. This study represents a fine beginning to what can be a very useful research program to validate biology teacher competencies and then to investigate teacher appraisal in relation to those competencies. It is hoped that this analysis may contribute to and encourage the continuation and expansion of this line of research.

REFERENCE

Novak, Joseph D. "A Preliminary Statement on Research in Science Education." Journal of Research in Science Teaching, 1:3-9, 1963.

Graybill, L. "Sex Differences in Problem Solving Ability." Journal of Research in Science Teaching, 12:341-346, 1975.

Descriptors--Developmental Tasks; *Educational Research; Logical Thinking; *Problem Solving; Science Education; *Secondary Education; Secondary School Science; *Sex Differences.

Expanded Abstract and Analysis Prepared Especially for I.S.E. by David P. Butts, University of Georgia.

Purpose

Because of society's expectations, girls appear to have had fewer hands-on experiences related to science learning and cognitive development. Because of these gender-related differences in their experience base, it was hypothesized that girls would be functioning at a lower cognitive developmental level and thus less successful in solving science-related problems.

Rationale

As exemplified in the work of Piaget, one's intellectual development is expected to be a function of earlier experiences. Informal descriptors of girls' experiences support the assumption that their experience base is substantially different than that of boys.

Research Design and Procedure

A posttest-only-no-control group design with non-random selection of three pairs of boys and girls at ages 9, 11, 13 and 15 served as the sample. No treatment was involved. Each student was asked to solve four problems. Based on recordings of these individual interviews, the students' level of cognitive development was estimated.

Findings

Boys consistently performed at higher cognitive levels than did girls. Observable differences in the ease with which boys manipulated materials were noted. Boys were said to be more successful in solving problems than were girls.

Interpretations

Cognitive development—or its lack—was not related to IQ or school achievement. Boys did better than girls because of the gender-related social dimensions. Schools value "such traits as agreeability, punctuality, cooperative, neatness, and submissiveness." Girls more than boys display these traits in conforming to schooling expectations—and at the same time in that conformity retard their cognitive development.

ABSTRACTOR'S ANALYSIS

In posing a question concerning a factor that may help teachers improve their approach to science instruction, the author begins with a hypothesis that has grown out of her personal teaching experience. This is very positive. Relating gender differences to the social milieu is useful. Although the author suggests that these differences are essential variables in cognitive development and problem solving, the study deals only with cognitive development and ignores both the literature and the analysis of the problem solving of students.

In this exploratory non-experimental study there is an intuitive search to see if differences exist between boys and girls, although the hypothesis suggested that the study would not "see if" but would demonstrate the transition. As hypothesized, one would have expected to find a pre-post design with a treatment in which a transition was attempted and observations made to document its existence or lack of occurrence. The author vaguely describes the sample which could consist of as few as

three pairs of students (N = 6) or three pairs for each of four age levels (N = 24). Previous experience was identified as a significant variable—but was not assessed in any systematic way. Although problem solving performance was identified as one dependent variable, no measure or analyses were reported. While informative, neither the graphs nor the report provide the reader with any data analysis to support that the differences observed were significant.

In this clearly written report on a topic of keen interest to science teachers, the reader is left with the clear impression that the discussion section reflects the author's opinion since no empirical data analyses are reported to support the conclusions. Teachers need more research that is based on important classroom observations as this study is. But, teachers need to have such questions followed by intelligent design, analysis, and interpretation. This study would be substantially improved by these three additions.

Jaus, Harold H. "The Effects of Integrated Science Process Skill Instruction on Changing Teacher Achievement and Planning Practices." Journal of Research in Science Teaching, 12(4):439-447, 1975.

Descriptors--*Autoinstructional Methods; College Science; *Educational Research; Elementary School Teachers; Higher Education; *Process Education; Science Education; *Scientific Methodology; *Teaching Skills

Expanded Abstract and Analysis Prepared Especially for I.S.E. by William S. LaShier, Jr. and James D. Ellis, University of Kansas.

Purpose

The study was designed to ascertain the effectiveness of integrated science process skill instruction on prospective elementary teachers' (1) achievement of the integrated science process skills, (2) selection of integrated science process skill instructional objectives, (3) writing of both integrated science process skill instructional objectives and learning activities in lesson plans, and (4) attitude toward the use of these skills in the classroom. A secondary purpose of the study was to determine the effects on the teachers of reading a persuasive communication as measured by the dependent variables already listed. The persuasive article advocated the teaching of integrated science processes in the elementary classroom.

Rationale

The intent of many of the new elementary school science curricula has been to develop the science process skills in children. This instructional emphasis is also reflected in the type of science process skill training received by prospective elementary school teachers. Butts and Raun (1969), among others, have determined experimentally that training inservice teachers in the basic science process skills (as defined by the Science Education Commission of AAAS) had positive impact on the teachers' competence in using the skills, attitude toward the skills, and use of the skills in classroom science activities. However, until the present research, no one had investigated

the direct effects of training prospective elementary teachers in the integrated science process skills.

The present study was also designed to further explore Hughes' (1970) finding that the use of a persuasive communication about using integrated science process skills significantly improved the attitudes of prospective elementary teachers toward teaching science as a process.

Research Design and Procedure

Procedures. The experimental design used was a static group comparison using three intact classes of 90 preservice teachers enrolled in elementary science methods classes. Prior to receiving one of the three treatments, each student demonstrated his/her ability to write acceptable behavioral objectives and design adequate science lesson plans at the 95 percent mastery level. Also prior to treatments, the intact classes were found to be statistically equivalent for verbal and mathematics scholastic aptitude test scores and cumulative grade point average.

The control group, Class A (N = 26) received placebo instruction in contingency contrasting and mastery learning. Class B (N = 31) received instruction in the integrated science process skills via the integrated science process skill self-instructional pamphlets. Class C (N = 33) received the same pamphlets plus the three-page persuasive communication. Treatment for all classes lasted four class sessions.

Instruction for groups B and C consisted of ten self-instructional pamphlets developed by Okey and Fiel (1971). Approximately eight hours were required to complete the tasks and self-tests which dealt with identifying variables, graph construction, relationships between variables, data processing, analyzing investigations, constructing hypotheses, operationally defining variables, designing investigations and experimenting. The additional treatment afforded the group C subjects was a three-page handout which contrasted factual, conceptual, and process approaches to teaching science.

Measures. The four major instruments used in the study were administered after the groups had completed the related treatment. The instruments included:

1) Integrated Science Process Skill Achievement Test which measured the acquisition of integrated science process skills by the subjects. Content validity and test reliability were reported for this test.

2) The Selection of Objectives Questionnaire consisted of a randomly ordered list of ten integrated science process objectives and ten science content objectives all relating to a unit of instruction on the topic Plants and Plant Growth. Each subject was told that all 20 objectives were appropriate and they were to select any ten that they felt should be included in a unit on plants. The content validity and reliability of the instrument were reported.

3) The Integrated Science Process Skill Attitude Measure consisted of 30 statements to which the subjects reponded on a five-category, Likert-type scale. For example, a subject might mark strongly disagree to "I am more interested in the number of science facts and concepts a child remembers than in how well he solves science problems." The reliability of the instrument based on a test-retest method was reflected in the Pearson product-moment correlation coefficient of 0.76.

4) Lesson Plan Objectives and Activities consisted of five behavioral objectives with one or more activities per objective. The lesson plans were evaluated by counting the number of integrated science process skill objectives and activities per lesson. No reliability or validity estimates were reported.

Findings

There were significant differences ($p < .01$) between means for Class A (30.46) with Class B (46.19) and Class A (30.46) with Class C (46.48) for the achievement test, with an overall F value significant beyond .001 level. There were significant differences ($p < .01$) between means for Class A (4.19) with Class B (6.22) and Class A (4.19) with Class C (6.75) for selection of objectives questionnaire, with an overall F value significant beyond .005 level.

There was a significant difference ($p < .05$) only between means for Class A (.46) with Class C (1.27) for the number of process skill objectives written in lesson plans, with an overall F value significant beyond the .05 level.

There were significant differences ($p < .05$) between the means of Class A (2.80) with Class B (4.70) and Class A (2.80) with Class C (5.00) for the number of process skill learning activities written in lesson plans, with an overall F value significant beyond the .01 level.

Interpretations

The researcher confirmed that preservice teacher-training is successful in promoting competence in the integrated science process skills. This confirmation was important since many preservice teachers in the control group lacked proficiency in some of the process skills. Many of the control group subjects apparently had little previous exposure to science process skills in their high school or college science courses.

Since preservice teachers who receive training in science process skills wrote significantly more process-oriented instructional objectives and activities than did their untrained peers, it was suggested that having teachers competent in science process skills was a first step in providing for development of process skill acquisition by children.

It was also implied that training teachers in process skills will increase use in the classroom of process oriented elementary science curricula. It was suggested that even though the attitudes toward process skills were not different among groups, training was required to implement this desire to teach process skills. The results of this study did not support previous studies which showed that reading of a persuasive communication could significantly influence the attitudes of its readers. It was suggested that perhaps the 700-word communication used was not powerful enough to affect teacher attitudes.

ABSTRACTOR'S ANALYSIS

The study supported the suggestion that well-planned instructional packages can improve short-term achievement of the selected objectives. There was no attempt to measure long-term retention. While the study demonstrated that achievement of process skills also influenced application of the knowledge in written lesson plans, no evidence was gathered to determine whether the preservice teachers operated differently in the classroom as a result of the instruction. Previous studies found that training inservice teachers in process skills caused teachers to teach these skills in the classroom science activities. Since preservice teachers have different concerns and experiences from inservice teachers, it may not be reasonable to generalize from the previous studies and infer that preservice training of four days will have an impact on actual classroom instruction. The major finding was to produce supporting evidence for planned instruction increasing achievement in preservice teachers. A priori judgment perhaps could have accurately formed this conclusion.

The failure of the study to support previous findings that a persuasive communication can influence attitudes was explained by the lack of power of the communication. The effect of the communication was not measured independently from the instructional treatment. A two-way factoral design with four groups would have included a treatment group receiving the persuasive communication treatment separately from the instructional treatment. Analysis of the relationship of the

additional group to the other groups would indicate the relative effect of the persuasive communication. Without the independent comparison, the study could not completely interpret the effect of the persuasive communication.

Indices of validity and reliability were reported for all dependent measures except the lesson plans. The appropriateness and interpretability of the indices is questionable. The reliabilities were all based upon the test-retest method, where the same items are used in both measures. With only four days between pre- and post-test, it is very likely that spuriously high correlations between the two administrations will result from a memory effect.

The validity data for the achievement test was based upon the degree to which the nine tasks related to the nine process skills. There were 53 items scored for the nine tests. A more meaningful relationship could have been drawn between the individual items within each task and the related process skills. A spuriously high correlation may have resulted from the correlation of the nine tasks with the respective process skills.

The validity of the objective selection questionnaire was established by the ability of experts to select process oriented objectives. The instrument was administered and scored according to directions that ask the respondent to select those objectives which are most important to teach. The instructions for the subjects suggested a more value-oriented measure than did the validation procedure. The validation procedure and the use of the instrument should be directed at measuring the same trait.

No reliability or validity was reported for the rating of the lesson plans or the number of process objectives and process oriented activities. It was reported that the lesson plan was not an instrument per se. Interrater reliability and concurrent validity methods might have been used to establish these indices.

The overall group comparisons were reported at levels of significance varying from $p < .001$ to $p < .10$. The Neuman Keuls test of the pairwise comparison of means was performed for those overall analyses significant at $p < .05$. It was suggested for attitudes that evidence of a trend existed for treatment groups having higher means when the level of significance was $p < .10$. If the level of significance for rejection of the null hypothesis was set at $p < .05$, then it wasn't reported as such, and any overall analysis which failed to meet the pre-established criterion indicates no systematic differences between the groups.

A univariate anova was used when five dependent variables were analyzed. When a series of univariate analyses are performed in one analysis, the probability that one of the analyses will be significant by chance alone is increased. This possibility is a limitation of the analysis procedure selected.

The analysis of the lesson plan was based upon the number of written process oriented objectives. Although these results were found to be significant, it was difficult to interpret mean score data. When inspecting the number of objectives, the treatment group with the highest score was 1.27 out of a possible score of 5. Does 25.4 percent process oriented objectives in a written lesson plan suggest competency?

REFERENCES

Butts, D. P. and C. E. Raun. "A Study of Teacher Change." Science Education, 53(1):3-8, 1969.

Hughes, E. F. "Role Playing as a Technique for Developing a Scientific Attitude in Elementary Teacher Trainees." Paper presented at the annual meeting of NARST, Minneapolis, March, 1970.

Okey, J. R. and R. L. Fiel. Science Process Skills Program. Bloomington, IN: Laboratory for Educational Development, Indiana University, 1971.

Kaplan, Eugene H. "A Model Biology Curriculum for Heterogeneous Seventh Grade Biology Classes Containing Culturally Deprived Students (in Israel), II. The Construction of the Model." Science Education, 59(3):321-332, 1975.

Descriptors--*Biology; *Culturally Disadvantaged; *Curriculum; *Secondary School Science; Educational Research; Science Education; Secondary Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Ellen Stephanie Simmons, Teachers College, Columbia University.

Purpose

The purpose of the author's experiment was to develop and employ a model for the teaching of seventh grade biology. This model included an emphasis on individualized instruction using a multi-media approach in order to provide for a maximum amount of learning regardless of the student's cultural background. The author hypothesizes that a program consisting of less teacher-centered instruction with a "mastery learning" concept effects greater cognitive and affective gains on the part of the student.

Rationale

The author states that due to the nature of settlement throughout Israel, the Ministry of Education in 1964 proposed educational reform based upon the needs of students from all socio-economic levels. The Ministry of Education recognized that the educational needs of that segment of society deemed culturally deprived were not being met. The emphasis of the author's model was an attempt to effect a better learning milieu for these deprived students. This model was, therefore, an attempt to provide a pragmatic interpretation to learning and teaching theory with the production and implementation of curriculum and educational materials. Twelve teaching strategies and 12 pedagogical techniques were drawn from literature on the teaching of culturally deprived students; these strategies and techniques then served as the foundation for the author's model.

Research Design and Procedure

In order to implement the 12 strategies and techniques for teaching culturally deprived students, a pilot study and three subsequent revisions were conducted in several seventh grade biology classes as summarized in Table 1.

Table 1

A Summary of Kaplan's Treatments

Group	Date	Classes	Schools	Number of Students	Units
Pilot	Not stated	Small groups	2	Not stated	Not stated (Five five-lesson units)
1	1969	6	4	Not stated	Two units of a projected 8-unit sequence were constructed and employed-- A. Water as a Habitat B. The Organism and Its Environment
2	1970-71	6	1	Not stated	Not stated
3	1971-72	5 Experimental	1	N = 154	Not stated
		4 Control	4	N = 251	

Because one of the major concerns of this study was to devise a curriculum which would meet the needs of a diverse Israeli society, the schools involved in the program were selected to reflect this diversity in the ethnic and economic backgrounds of their students. In a preceding article appearing in the same journal which outlines the theoretical foundation for his model, Kaplan also categorized students into the following intelligence groups:

1. Intellectually gifted (more than one S.D. above the mean IQ score)
2. Average
3. Below average (within one S.D. below the mean)
4. Low performance (beyond one S.D. below the mean)

A curriculum, therefore, was designed to help all students (with the exception of the lowest group), but special emphasis was given to the below average group designated as "gifted culturally deprived." The author recognized deficiencies in this particular group including a low reading comprehension, a poor learning and reasoning pattern, a rigidity in the face of change, and a lack of motivation and self-discipline in relation to academic goals and study habits. The curriculum, therefore, was aimed at this low average group with enrichment activities available for the faster students. The author hypothesized that statistically significant gain scores would be attained by this target group as a result of his proposed curriculum innovations. The course of study followed in this experiment was part of a sequential junior high program which included aquatic zooecology on the seventh grade level. A model curriculum was developed which consisted of five lessons within each unit:

1. A teacher-directed didactic introduction or discussion period
2. An open-ended laboratory experience
3. An analysis—a "competency measure" and a programmed lesson of core concepts
4. An analysis—an additional time period for slower students to complete the core material or an enrichment period for faster students
5. A unit test—a test with maximum score possible being 80 percent for those completing only the core program, and additional points being an incentive for those students involved and tested on the enrichment lessons.

The unit test was aimed at placing the students under a minimum amount of pressure and at allowing a maximum amount of encouragement. Students were told that "mastery learning" could be achieved through completion of only the core lessons of each unit.

As a result of the model being aimed at the gifted culturally deprived, the lecture method was deemed inappropriate. Since this was the predominant method employed by Israeli teachers, child-centered teaching methods had to be taught to those teachers involved in the experiment. A Teacher's Guide was also provided as an additional resource.

Findings

The author reported achievement results for only his 1971-72 study. These results dealt primarily with the difference in gain scores between pre- and post-achievement tests. These tests covered knowledge of basic biological concepts appropriate for the seventh grade curriculum throughout Israeli schools. In the 1971-72 study, Kaplan had an experimental group of 154 students and a control group of 251 students. In comparing the results of the increase in gain scores between the pre- and post-tests of these two groups, Kaplan reported that the experimental groups attained a higher percentage of increment of improvement than did the control group. This increment was true in three subtests: basic knowledge, comprehension, and higher functions.

In addition to the achievement test, the Attitudes Towards Biology test, a modified version of the NALSMA Report test of attitudes towards mathematics, was employed. This test consisted of twelve areas. Using an analysis of variance between the control and experimental scores on this test, three areas were found significant at the 0.05 level in favor of the experimental group.

Interpretation

Kaplan infers that the gifted culturally deprived students can achieve "mastery learning" if teachers are trained to utilize a more student-oriented approach to teaching. He also infers that the gifted culturally deprived students achieve a level of mastery of subject matter only if they are allowed a flexible amount of time for learning.

ABSTRACTOR'S ANALYSIS

While the author reports significant gains achieved, especially for his 1971-72 research, several shortcomings are inherent in the reporting of his study. This deficiency in reporting can be seen in the following.

1. The reader is not given a full summary of each year's study; comparisons among the various investigations are lacking.

2. The article is not presented in a scientifically-oriented format; e.g., the purpose of the study is found under the results category. Therefore, only through the reader's reorganization of the material can a full understanding of the project and its results be gained and understood.

3. The author makes several generalizations without citing specific supporting data; e.g., he not only fails to mention the twelve sub-topics of the Attitudes Toward Biology test, but he also neglects to specifically state which three categories he found significant.

4. Some of Kaplan's tables do not contain sufficient information; e.g., he defines four groups of students based upon IQ in his first article, but he fails to statistically analyze any data in relation to this method of grouping.

5. The author does not describe in any great detail the length or duration of each lesson and/or the content of each unit; e.g., Kaplan neither specifically states the units employed for his 1970-71 or his 1971-72 research nor tells how long the units were studied.

6. The author assumes the reader's understanding that the control group involved the traditional lecture method and the traditional curriculum techniques; no information whatsoever is given.

Therefore, the author's lack of reporting his investigation in a scientifically-oriented manner detracts from the merits of his endeavor.

While Kaplan states in his article that his main objective is to document the amount of significant gain among the gifted culturally

deprived students, his research does not show this factor adequately through his statistical design. Instead, the findings reported for 1971-72 are those for all intellectual levels and not for just the gifted culturally deprived. Furthermore, his documentation does not indicate what percentage of the 154 students in the experimental group were considered in this category; therefore, the gains he reports (Table 6) would be for all intellectual groups and not for just this select group.

In contrast, a more adequate statistical design could have involved a comparison of the experimental and the control treatments for each of Kaplan's four intelligence groups. Using pre- and post-test gain scores, analysis of variance could have been performed between the groups (1 and 2, 1 and 3, 1 and 4, 2 and 3, 2 and 4, and 3 and 4). Further analysis could have been performed based upon cultural patterns; e.g., students from Western European and American backgrounds could have been compared with those of Asian-African backgrounds for each ability group. Figure 1 summarizes this suggested research design.

In addition to the deficiencies inherent in his research design, Kaplan's testing procedures are questionable since he does not justify what he does. This point is illustrated in his 1969 study when he reports administering an achievement test after only one week under his program (see Table 5). A more logical choice for administering the test would have been after the entire treatment had been given. His questionable testing procedures are also illustrated in his use of a standardized test for his 1971-72 research. It would have been more appropriate for him to have constructed his own test or to have located one which could have assessed the significance of his method. The need for a new type of testing which would be an integral part of any given curriculum and the implementation of this type of testing concept is discussed in relation to Israeli schools by Tamir (1978).

Another area of concern is the author's failure to show the relationship of his investigation to other research in this field. Although other studies dealing with culturally deprived students are mentioned

in general, no specific references are employed for comparison. This deficiency is paramount because of the need for similar programs in the United States.

While obvious deficiencies are prevalent in Kaplan's reporting and research design, his model is valid. This model is an objective treatment of subject matter administered within the biological sciences. It would have been appropriate for Kaplan to have made recommendations for the implementation of this model in other subject areas. Since the education of the culturally deprived is a difficult problem at all levels of education, Kaplan's model could be employed as a teaching method to help every student achieve "mastery learning."

REFERENCES

Tamir, P. "The Potential Role of Examinations in Innovative Curricula." The American Biology Teacher, 40(7):426-428, 1978.

GROUPS	TREATMENTS											
	Control						Experimental					
	Pre-Test		Post-Test		Gains		Pre-Test		Post-Test		Gains	
	European and American	Asian and African	European and American	Asian and African	European and American	Asian and African	European and American	Asian and African	European and American	Asian and African	European and American	Asian and African
1. Intellectually gifted												
2. Average												
3. Below average												
4. Low performance												

Figure 1.--Suggested Research Design

Nelson, Miles A. "Evaluation of a Cooperating Teacher Training Program." Science Education, 59(4):505-515, 1975.

Descriptors--*Cooperating Teachers; Educational Research; *Instruction; Science Education; Secondary Education; *Secondary School Science; *Student Teachers; *Teacher Supervision; Teaching Methods

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Hans O. Andersen, Indiana University.

Purpose

The writer's purpose was to describe a training program for cooperating teachers which provided the participants with specific supervisory skills and to explore the program's effects on a novice's practicum experience. Specifically, the investigator sought an answer to the following question: Is there a correlation between the behaviors a cooperating teacher has acquired in a training program and the classroom performance of his assigned student teacher?

Rationale

The author noted that numerous articles have been written that describe how to train a cooperating teacher but that few investigators have attempted to relate the training provided cooperating teachers to changes in the behavior of student teachers. More specifically, articles describing the influence of the cooperating teacher in the attitude and teaching style of the student teacher are available but very few of these articles relate what the cooperating teacher does consciously to modify the behavior of a practice teacher to changes that are later evidenced.

Research Design and Procedure

Six two-hour workshops were provided for participating cooperating teachers. The objectives for these workshops were:

a. To identify beginning science teachers' strengths and weaknesses in the area of planning, questioning techniques and student involvement.

b. To learn how to communicate, to the beginning teacher, identified strengths and weaknesses.

c. To plan experiences aimed at eliminating observed weaknesses in student teacher behavior.

The purposes for each workshop are here delineated.

Session 1

Purpose-- Demonstrate the effect of focusing on a single aspect of a lesson and critiquing it.

Action--The cooperating teachers were shown a 10-minute micro-teaching tape of a novice, a post lesson critique focused on a single teaching skill, and the reteach tape. In the discussion that followed viewing two cycles of teach, commentary and reteach, the cooperating teachers concluded that a dramatic change in novice teaching behavior could result from this type of effort.

Sessions 2 and 3

Purpose--The cooperating teachers viewed lessons, practiced identifying weaknesses in a student teacher's planning, and devised experiences which would provide the practice needed to eliminate the undesired behaviors.

Action--The cooperating teachers were shown examples of beginning teachers who a) planned and did not plan for student involvement; b) practiced and did not practice a demonstration before using it in class; and c) did and did not use performance objectives in their planning. The discussion following each segment focused on the differences in performance of students who did and did not plan. Then participants were led to a discussion of tactics they could use to eliminate these weaknesses.

Session 4

Purpose--To detect weaknesses in a student teacher's questioning techniques and plan tactics to eliminate questioning weaknesses.

Action--The investigator began the session with an examination of questioning types and a discussion of the appropriate use of questions. A list of questions classified by type with a statement regarding an appropriate use of the questions was distributed to facilitate discussion. Video-taped lessons of beginning teachers were shown and analyzed.

Session 5

Purpose--Examining techniques used to get student involvement.

Action--Video-tapes illustrating successful and unsuccessful attempts to stimulate student involvement in lessons were shown to stimulate a discussion of novice errors and tactics cooperating teachers could use to help the novice eliminate the errors.

Sessions 5 and 6

Purpose--Examine the purpose for and handling of the critique session.

Action--The investigator provided the cooperating teachers a model of the critiquing session and research data supporting his claim that the purpose for the critique was to get the novice involved in discussing one or two positive or negative aspects of a lesson he/she had just taught and planning how to improve his/her next presentation. Video-tapes of critique sessions were also presented and analyzed.

Thirty-nine (39) cooperating teachers, 19 of whom had completed the training sessions, and 52 beginning teachers participated in this study. The novices were randomly assigned to cooperating teachers in their respective disciplines.

To determine if cooperating teachers used the methods presented in the workshop, a 26-item questionnaire was sent to each novice. The questionnaire directed the novice to rate the frequency of the cooperating teacher's comments (0-4 scale) relative to planning, questioning techniques, student involvement and the cooperating teacher's supervisory style. The questionnaire contained items classified as follows:

6 questions on planning comments
6 questions on questioning techniques
7 questions on student involvement comments
7 questions on critiquing style

Ratings were totaled within each area and used as the independent variable. A 2 x 2 x 2 factorial design which had as factors; Replication, first or second semester; type of experience, internship or student teaching; and Cooperating teacher training, yes or no. Scheffé's approximation was used to correct for unequal cells.

To determine the effects the supervision had on the novice, trained judges who had never seen the novice teach were used to rate prepracticum and post practicum video-tapes. The judges' reliability (.90) in using a low-inference observation instrument of the checklist variety was established. The stepwise multiple regression technique used to analyze the data is well described by the author.

Findings

1. Irrespective of replication or type of experience, the trained cooperating teachers were rated by the novice as making more comments about the novices, questioning techniques, and student involvement. The differences were significant at the .05 level. The differences for planning comments were not significant but the trained cooperating teachers were favored. The analysis of the data on critique style also significantly favored the trained group. In general, an interpretation of the data supported the conclusion that the trained teachers performed the supervisory job better as supervision was defined by the model.

2. The data seem to support the claim that the frequency of comments made by the cooperating teacher which are related to the skills developed in the training program are correlated with the beginning teacher's performance. The magnitude and direction of the changes were purposely not discussed by the investigator.

Interpretations

The investigator indicated that the short six-hour training program had an effect on the teacher's style of supervision. While avoiding statements on magnitude and direction, the investigator reported correlations between supervisory behavior and novice post practicum behavior.

ABSTRACTOR'S ANALYSIS

That we should be accountable is a concept that we all should embrace. Taxpayers and students have placed faith (and money) in us as persons who are capable of preparing teachers. Yet, as Nelson points out, there is very little hard evidence available supporting our claim that we are able to prepare teachers. Yes, most of us have testimonials and many of us have observed our students doing what we and many others think is a credible job. But, as Nelson undoubtedly had to conclude, hard data support is scarce and hard to obtain.

Nelson's effort to train cooperating teachers is most noteworthy and absolutely essential. However, one must wonder if a six-hour session can produce the results he had hoped to obtain. Two questions are unanswered. They are: 1) To what extent did the trained cooperating teachers become committed to the model and conscientiously implement it in their critiques? How much enthusiasm, and how much time did they devote to the task? 2) To what extent did untrained cooperating teachers employ elements, or all, of Nelson's model? The statistics reveal a difference between the trained and untrained but the magnitude of the difference cannot be practically estimated, and the

untrained could well have been partially trained or, because of other factors, oriented toward behaviors described in the model.

I have used similar techniques as those described by Nelson. As did Nelson, I pieced together the techniques from a variety of research sources and philosophical statements. I believe the model is sound but to test its soundness I would have to develop testing procedures that most of us avoid like the plague. That is, I would have to teach Nelson's model to one group, the antithesis of Nelson's model to another group of cooperating teachers. Then I would have to determine that the cooperating teachers were handling their critiques in a manner consistent with the models. Then I could compare changes made by the student teachers that could be attributed to the critique process. (This design could be improved if we could identify those students whose behaviors are immune to either treatment.)

I doubt if Nelson needs any suggestions on how his research might have been improved. His list of desired and desirable changes probably exceeds mine. Hence, I would prefer to compliment Nelson on an excellent study performed in a naturalistic setting in which the investigator attempted to control as many variables as possible. It is obvious that the data were carefully collected and analyzed. And it is obvious that Nelson was very careful and conservative in stating conclusions and implications of his effort. He had to be!

Nelson's descriptions of the workshop sessions were particularly excellent. I believe I could replicate his effort in spite of the fact that his descriptions are exceedingly brief. This brings me to another point. I assume I could obtain a more elaborate description of the workshops from Nelson, but that is not mentioned anywhere. Also, are the video-tapes available for rent or lease, or to copy? Does the investigator, himself, feel that a replication or a replication with modification would be desirable? If so, under what conditions?

In way of criticism, would a description of the trained and untrained cooperating teachers be of use to persons reading the report? How

were the cooperating teachers obtained? Were there differences between the trained and untrained cooperating teachers that could account for noted variances? Was Nelson free to make any such comparisons and report them in the literature?

In conclusion, permit me to offer my opinion that Nelson's effort was timely, it was apparently well implemented, and it was well reported. I also suspect that it was underfunded and that Nelson was unable to control aspects of the study he would have liked to have controlled.

Seymour, Lowell A. and Lawrence F. Padberg. "The Relative Effectiveness of Group and Individual Settings in a Simulated Problem-Solving Game." Science Education, 59(3):297-304, 1975.

Descriptors--*Biology; Educational Research; *Group Activities; Individual Activities; *Instruction; *Problem Solving; Science Education; Secondary Education; *Secondary School Science

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Lynn W. Glass, Iowa State University.

Purpose

This research was designed to investigate two aspects of group work. A modification of a simulated problem-solving game devised by Jay Hall was used to compare the relative effectiveness of individual work versus group work and group work alone versus group work following individual work.

Stated as directional hypotheses, the investigators sought to determine (1) whether students working in groups will score better (lower) as a group than the average score of students working individually, and (2) whether students working in groups, with the previous experience of individually attacking the problem, will score better (lower) than students working in groups without such prior individual work.

Rationale

The value of groups for developing attitudinal qualities, motivation, social interaction skills, in-depth understanding, greater retention of information gained, etc. has been accepted either intuitively or empirically by large numbers of persons. The authors cite several studies that provide support for various values of group work. The authors of this study cite a concern of theirs that many teachers are misinterpreting the results of the "Lost on the Moon" exercise as evidence for the superiority of group work over individual work. Hence, the rationale for this study was to design an investigation that would study the relative effectiveness of individual work versus group work and individual work versus an individual-group sequence when applied to the successful completion of

a problem-solving task. This study does not depend on any single prior piece of research for a basis; it does, however, contribute to and build upon a large cluster of research that deals with values of group work.

Research Design and Procedure

Thirty-two heterogeneously grouped by ability, coeducational, tenth grade biology classes taught by 15 teachers were used to obtain data. Each class was divided randomly into two teams of 14 students each. On the fourth class period of the school year team No. 1 completed the "Lost on the Moon" exercise first as individuals and then in groups of three or four students. Students in team No. 2 completed the exercise first in groups and then individually. Data are reported by class and by team as mean individual scores and mean group scores. This counterbalanced design can be represented diagrammetically as thus:

$$R \quad X_i O \quad X_g O$$

$$R \quad X_g O \quad X_i O$$

when i represents completing the exercise as individuals and g represents completing the exercise in groups.

In the "Lost on the Moon" exercise, students are directed to rank from most important to least important fifteen items available to a crew of astronauts stranded on the moon. Scores are calculated by adding the absolute differences between proposed rankings by the students and rankings provided by NASA. Low scores, indicating least difference between proposed rankings and NASA rankings, are best; scores can range from 0 to 112.

Findings

A summary of the data quickly demonstrates an apparent advantage of group work over individual work.

Team No. 1 (Individual-Group Sequence)		Team No. 2 (Group-Individual Sequence)	
Individual Scores		Group Scores	
Mean	= 39.9	Mean	= 33.4
Median	= 39.0	Median	= 34.0
Group Scores		Individual Scores	
Mean	= 31.2	Mean	= 34.4
Median	= 30.5	Median	= 34.0

The first hypothesis compares the team No. 2 group scores to the team No. 1 individual scores. The results of the median test indicate that these two sets of scores differ in central tendencies. The first hypothesis is supported, students working in groups perform better than students working alone.

The second hypothesis compares the team No. 1 group scores to the team No. 2 group scores. The median test was used to test the difference in central tendencies of these two sets of scores. The second hypothesis, that students working in groups with prior individual experience will perform better than students working in groups without prior individual experience, was not supported.

Interpretations

The authors interpret the results of their study to support the use of small groups rather than individuals in simulated problem-solving activities. If the use of additional class time is not to be considered detrimental to some other aspect of instruction, then the individual-groups sequence can be interpreted as superior to the use of groups alone. Both of these findings were anticipated by the authors since students working in a group have the advantage of several viewpoints and sets of background knowledge and skills; further, students in the individual-group sequence had more time to consider the problem and had an opportunity to formalize their own ideas regarding the problem before being subjected to the influence of other students' ideas.

ABSTRACTOR'S ANALYSIS

This reviewer finds the article to be an important contribution to the literature. Undoubtedly it will find more value among practitioners seeking to find ways to teach their learners more effectively than it will among researchers trying to build a model of human learning. The article can be read and interpreted easily and uses statistical techniques that should be understood by a very wide spectrum of readers.

The authors chose to use the median test to determine if the two independent populations differed in central tendencies. A much more powerful test would have been the Mann-Whitney *U* test. More of the information in the data is used since the Mann-Whitney *U* considers the rank value of each score rather than simply its location with respect to the combined median. Using the information presented by the authors a probability of error equal to 0.06 can be calculated for the second hypothesis rather than 0.15 value arrived at with the median test. This test would have provided the authors with more support when they interpreted the individual-group sequence to have some advantage over group work alone.

Although data collection procedures appear to have been sound, it is unknown whether classroom directions were standardized with respect to administering the "Lost on the Moon" exercise. For example, were all teams of students given the same amount of encouragement when asked to do the exercise the second time? How were "Why do we have to do the exercise again?" questions handled?

A change in the article which would permit future researchers to replicate the study would be a short explanation of how the original "Lost on the Moon" exercise was modified for this research.

Discussion of possible data comparisons that do not contribute to the purpose of the study and are not made within the study tend to distract the reader from the main thesis of the authors.

Shymansky, James A.; John Penick; Ronald G. Good; and Charles C. Matthews. "Using Macroanalytic Techniques to Study Teacher Behavior Patterns." Journal of Research in Science Teaching, 12(3):221-227, 1975.

Descriptors--*Behavior Patterns; *Classroom Observation Techniques; *Data Analysis; Educational Research; Interaction Process Analysis; Measurement Techniques; Science Education; *Teacher Behavior

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Thomas P. Evans, Oregon State University.

Purpose

The purpose of the investigation was to develop a modified version of Campbell's macroanalytic technique and to determine the effectiveness of the modified technique in identifying new long-term patterns of teacher behavior.

Rationale

The investigation was an extension and application of ideas presented by Campbell (1973) in a study in which he developed and utilized a macroanalysis technique for analyzing teacher classroom behavioral data obtained through systematic observation of classroom interaction.

The following are assumptions that seem to be inherent in the study but were not stated as assumptions by the researchers:

1. Macroanalytic techniques for analyzing classroom observational data are superior to matrix analysis, because they preserve longer sequences of behavior, utilize greater amounts of coded observational data, and provide greater potential for identifying and improving instructional strategies.
2. Predominant classroom behavior patterns are best identified by analytic techniques that utilize the maximal amount

of coded observational data and collapse all repetitive behaviors within a given time duration into single units, reflecting sequences of different behavior.

Research Design and Procedure

Observational data were obtained by coding the classroom behavior of a fifth grade science teacher for approximately 32 minutes utilizing the SCAS Classroom Interaction Categories-Teacher Behaviors (1971). The SCAS system contained 18 categories of teacher behavior. Observations were coded every three seconds unless there was a change in behavior within the three-second interval. When a change occurred, it was coded. The resulting 770 observational tallies were analyzed using Campbell's five-tally macroanalytic technique (1973) and a modified version of the technique (hereafter referred to as MACROanalysis).

The macroanalytic technique developed by Campbell involved the organization of observational data into units, each consisting of five sequential tallies of behavior. If, for example, a 30-second segment of teacher behavior was coded as BBBADDAABB, a macroanalysis of the behaviors using the Campbell five-tally technique would result in the following patterns of behavior: BBBAD, BBADD, BADDA, ADDAA, DDAAB, and DAABB. Since the observation data were coded in three-second intervals, each of the resulting patterns described 15 consecutive seconds of behavior.

MACROanalysis was developed by modifying the Campbell macroanalytic technique. The modified version organized coded observational data into five-tally patterns, but it collapsed all uninterrupted chains of repetitive behaviors into single units; i.e., a series of coded behaviors such as ABAAACD would be reduced to ABACD. Applying MACROanalysis to the same 30 seconds of coded behavior previously analyzed using the Campbell technique (BBBADDAABB) would result in one pattern, BADAB, rather than six. The resulting pattern was broader in nature in that it reflected a sequence of different behaviors over a longer interval of time.

Each set of behavior patterns obtained by subjecting the coded observational data to both the macroanalytic and MACROanalytic techniques was ranked based on the number of times each sequence of behavior occurred in the coded observational data. The percentage represented by each sequence of behaviors based on the total number of identified behavior patterns was calculated for each analytic technique. Visual comparisons were made between the rankings, and the nonparametric Test for Significance of Differences Between Two Proportions was used to determine if significant differences existed between percentages of patterns ranked 1,2,3 and 4 in the MACROanalysis and the percentage of similar patterns obtained through macroanalysis.

Findings

The major findings reported by the researchers follow:

1. analysis utilizing Campbell's macroanalytic technique revealed a total of 767 behavior patterns, 208 of which were different;

2. analysis utilizing the MACROanalytic technique revealed a total of 441 behavior patterns, 111 of which were different;

3. only four of the 25 most frequent behavior patterns identified by each type of macroanalytic technique were common to both analyses;

4. the four common patterns were ranked 1,2,3 and 4 in MACROanalysis and 2,7,15 and 18 in Campbell's macroanalysis;

5. these four patterns accounted for approximately 33 percent of the total number of patterns identified using MACROanalysis and 10 percent of the total number of patterns identified using Campbell's technique;

6. the percentage of time that each of these four patterns appeared in MACROanalysis was shown to be significantly

greater than the percentage of time they appeared utilizing Campbell's macroanalysis;

7. nineteen of the remaining 21 behavior patterns in the top 25 identified utilizing MACROanalysis did not appear in the 50 most frequent patterns identified by Campbell's macroanalytic technique;

8. six of the remaining 21 behavior patterns in the top 25 using the MACROanalytic technique did not appear in any of the 208 different patterns identified using Campbell's macroanalytic technique;

9. MACROanalysis revealed 51 behavior patterns not found using Campbell's macroanalysis; and

10. a summary of the data contained in the individual SCAS categories revealed that categories S2 (observes student behavior but does not respond), S7 (asks questions which do not tell students what to do), and S3 (accepts student behavior without evaluation) were the predominant teacher behaviors with frequencies of 466, 161, and 123, respectively.

Interpretations

The following statements summarize the conclusions, inferences, and implications that the researchers drew from the findings:

1. MACROanalysis had advantages over macroanalysis because of its greater efficiency and manageability and its increased potential for identifying broader behavior patterns over longer intervals of time;

2. failure of the MACROanalytic technique to include series of repetitive behaviors was not critical because of the nature of the lost information and the ease by which the frequency

of each individual category of behavior and its percentage of the total number of observed behaviors could be retrieved;

3. MACROanalysis has a great deal of potential for having an impact on research and the training aspects of education; and

4. teacher behavior patterns which characterize and best facilitate the attainment of individual science program goals could be identified utilizing MACROanalysis: the technique appeared to have a great deal of potential for defining teacher effectiveness at all levels.

ABSTRACTOR'S ANALYSIS

The investigators in this study are to be commended on their simple but imaginative ideas for analyzing classroom observational data. The assertion that they made regarding the lack of advances in analytic techniques for observational data other than matrix analysis is unfortunately true. The first AETS Yearbook, A Review of Research on Teacher Behavior (Balzer, Evans, and Blosser, 1973) revealed that research involving systematic observation of science classroom behavior increased dramatically in the late 1960s and early 70s, but most if not all of the attention was directed toward development, modification and application of observation instruments. Certainly, none was devoted to new analytical techniques for analyzing observational data. In recent years the number of studies involving systematic observation of science classroom behavior appears to be declining. Many reasons could be given for this decline, but a portion of the problem is directly related to the lack of further advances in analytic techniques. MACROanalysis represents an attempt to develop and apply a new technique for analyzing coded observational data. The real contribution of the technique, despite the overly optimistic views of the researchers concerning its potential for identifying teacher effectiveness, may be its potential for generating a renewed interest in systematic observation of science

classroom behavior. Further, it builds on the existing ideas presented by Campbell and could, along with Campbell's macroanalysis, serve as a springboard for engendering additional analytic techniques for examining classroom observational data.

Both the research and research report would have been greatly improved if the researchers had avoided what appear to be discrepancies in the data, if they had not ended the report with conclusions that present the new technique as a potential analytic panacea for identifying effective classroom practices and science teacher behavior, if they had not based the investigation on a questionable assumption, and if they had been more precise and consistent in their definition and comparison of the nature of the behavior patterns identified through MACROanalysis and macroanalysis. Each of these points is discussed separately and more completely in the following paragraphs.

Several minor discrepancies in the data were identified. First, the researchers reported that 770 observational tallies were analyzed with Campbell's macroanalytic technique, resulting in 767 five-tally patterns of behavior. The analysis should have resulted in 766 patterns of behavior. Secondly, a table summarizing the frequencies of behaviors revealed 782, and not 770, tallies of observational data when they were totaled by the reviewer. A total was not presented in the table. Third, the summary table reported that the S3 category (observes student behavior but does not respond) had a frequency of 466, representing 58 percent of the total behaviors. The percentage should have been 60 percent if the total number tallies was 782 or 61 percent if the total tallies was 770. Such discrepancies were not serious in that they did not significantly affect the conclusions. They were, nonetheless, confusing and did not ensure confidence in the accuracy of the remaining data and analyses.

Some of the conclusions concerning the potential of MACROanalysis went well beyond the findings. It was true that the application of the new analytic technique resulted in the identification of new and broader patterns of behavior. The analysis was based, however, on only 32 minutes of classroom observation, and no evidence was

presented showing that the new and broader patterns had determining influence on independent variables such as criteria of effectiveness. The findings would certainly not support the conclusion stated by the researchers that "MACROanalysis can be used to identify teacher behavior patterns which characterize and best facilitate the accomplishment of the goals of individual science programs" (p. 226). This is not to say that the technique does not have potential for identifying what constitutes effective science teacher behavior. Additional research may reveal some of its potential, but the findings reported in the study alone would not support such optimism.

The researchers' optimism concerning the potential of MACROanalysis may have been directly related to their assumption that the technique would result in the identification of predominant classroom behavior patterns. (A more exact statement of the assumption appears in the rationale.) If by predominant they meant those behavior patterns having determining or prevailing influence on selected dependent variables, the conclusions would seem more realistic. But the assumption was questionable unless predominant behavior pattern had some different meaning that was not revealed by the research report. The statement, rather than being an assumption, was an excellent research problem in need of further investigation.

In their definition of the problem, the researchers introduced MACROanalysis and assumed that the application of the technique would result in predominant patterns of behavior that reflected sequences of different behaviors. Throughout the remainder of the report, they referred to patterns of behavior when discussing and comparing the results of MACROanalysis and macroanalysis. The phrase, "sequences of different behaviors," was not used again in the same context. If the researchers meant by the phrase that MACROanalysis would simply result in different patterns of behavior, then an important aspect of their technique was possibly overlooked. If they actually meant that MACROanalysis would result in sequences of different behaviors or changes in behaviors, then they were inconsistent throughout the remainder of the report. If the latter were true,

they should not have used new patterns of behavior as the criterion for determining the effectiveness of MACROanalysis. The technique by definition would result in new and different patterns of classroom behavior. Regardless, the point to be made is that the researchers failed to capitalize on a very important contribution of MACROanalysis: application of the technique resulted in the identification of patterns of changes in behavior.

Several questions concerning the investigation and research report have been raised, but they do not take away from the fact that MACROanalysis is a very interesting technique for analyzing coded observational data. The technique should stimulate further research in the area of systematic observation of science classroom behavior. The following are several suggestions for future research.

1. MACROanalysis could be used to re-examine the coded observational data in existing studies of science classroom behavior.

2. The relationships between patterns of behavior identified by MACROanalysis and a variety of possible independent variables, such as student gains and teacher characteristics, should be investigated.

3. The potential of the technique for identifying characteristic teaching styles for individual teachers, various teaching strategies, and specific science programs should be investigated.

4. Since there is nothing sacred about a five-tally sequence of changes in behavior, the technique could be used to study sequences of behavior preceding or following a particular behavior or sets of behaviors.

5. Subscripts could be used along with MACROanalysis to indicate the number of repetitive behaviors in a series. For example, the coded data AABBBCDDEEEE would appear as $A_2B_3C_1D_2E_4$.

Similar patterns of changes of behavior could then be analyzed for a variety of factors, including the amount of repetition with a given pattern. A profile for each pattern could be identified by placing the mean length or number of repetitions for each change of behavior in the subscript of the pattern. Such profiles could be investigated to determine whether or not they more accurately described the predominant pattern than simply listing the sequences of changes without subscripts.

REFERENCES

Balzer, A. LeVon; Thomas P. Evans; and Patricia E. Blosser. A Review of Research on Teacher Behavior. Columbus, OH: Association for the Education of Teachers in Science and ERIC Science, Mathematics and Environmental Education Clearinghouse, 1973.

Campbell, J. R. "Pattern Analysis—A Macroscopic Development for Interaction Analysis." Paper delivered at the annual meeting of the National Association for Research in Science Teaching, Detroit, March 27-29, 1973.

Mathews, C. C.; D. G. Phillips; and R. G. Good. Student Structured Learning in Science: A Program for the Elementary School Teacher. Dubuque, IA: Wm. C. Brown Company, 1971.

Tamir, P. and E. Jungwirth. "Students' Growth as a Result of Studying BSCS Biology for Several Years." Journal of Research in Science Teaching, 12(3):263-279, 1975.

Descriptors--*Biological Sciences; *Curriculum Evaluation; *Educational Research; Science Education; *Secondary School Science; Secondary Education; *Science Course Improvement Project.

Expanded abstract and analysis prepared especially for I.S.E. by Chris A. Pouler, Hyattsville, Maryland.

Purpose

This study had a threefold purpose:

1. To provide an overview of the students' growth during four years of studying the BSCS biology.

2. To identify, in the development of the project from its implementation in 1965, trends in various aspects of students' achievement, which may have at least some generalizable attributes.

3. To assess the feasibility of the BSCS adaptation in Israel by comparing BSCS and non-BSCS students regarding their achievement in biology, their acquisition of inquiry skills and their attitudes toward science and nature.

Specific questions involved:

1. The growth of BSCS students regarding biological knowledge, inquiry skills, attitudes and understanding of science process.

2. The best predictors of student growth.

3. The effect of a BSCS course by comparing BSCS students with non-BSCS students. In addition, statistical comparisons assessed differences due to sex, type of school, and ethnic background of students.

Rationale

In 1965 an adaptation of BSCS Yellow Version was introduced to Israli schools. By 1971 approximately one-half of the high school biology classes utilized this curriculum. This study sought to describe the impact of the BSCS Adaptation.

Research Design and Procedure

Biology was taught for three periods a week to all ninth and tenth grade students. Those students choosing to specialize in biology continued their study during the eleventh and twelfth grades for four to five periods per week. All students completed units dealing with Unity and Diversity while biology majors continued with Heredity, Evolution and Ecology. At the end of the twelfth grade, all majors took a matriculation examination.

This study compared the achievement, attitude and process knowledge of the BSCS students and non-BSCS students. Further comparisons were made concerning ethnic background (European versus non-European) and among the BSCS students only regarding sex differences and type of school (city, rural and agricultural).

The sample included three class populations of students who entered the ninth grade in 1965, 1966, and 1967. These classes graduated in 1969, 1970, and 1971. The comparison groups were selected as closely as possible for types of students, teachers and schools.

Since the instruments of evaluation had to test the dependent variables of achievement in content, understanding of process, and attitude, the tests were either (a) locally designed, (b) locally adapted, or (c) translated. To insure worthwhile data, the instruments were administered at the most appropriate time. Listed below are the dependent variables, test of choice and time of administration.

1. IQ:

   Milta test (locally designed), beginning of the ninth grade.

2. Biological Knowledge:

   General Biology Information Test (locally designed), beginning of ninth and end of tenth grades.

   BSCS Quarterly Tests (locally adapted), as units were completed.

3. Science Process:

   TOUS Form W (locally adapted), beginning of ninth grade and end of tenth and twelfth grades.

   SPI (translated), beginning of ninth grade and end of tenth, eleventh and twelfth grades.

4. Attitude:

   Biology Attitude Inventory (locally constructed), beginning of ninth and end of tenth grades.

5. Inquiry Achievement:

   Biology Process Test (locally designed), end of tenth grade.

   Practical Matriculation Test (locally designed), end of twelfth grade.

   Written Matriculation Test (locally designed), end of twelfth grade.

Findings

Comparisons Among BSCS Students

1. BSCS appeared to be adequate preparation for biology majors.

2. Comparing the achievement of tenth grade boys and girls, no significant difference occurred. However, with biology majors, twelfth grade girls outperformed boys in understanding and dealing with inquiry processes.

3. As students progressed from tenth grade (where deficiencies in mastery or inquiry skills existed) to twelfth grade, the mastery level of inquiry skills improved.

4. Regarding the students of city, rural (kibbutz) and agricultural (boarding) schools, only slight differences occurred between the city and rural schools. For all three classes, there was a conspicuous difference between the students from agricultural schools and those from city and rural schools. As for the twelfth grade majors, some of the differences disappeared.

5. The correlation between the various measures of achievement ranged from 0.00 to 0.58.

Comparisons of BSCS and non-BSCS Students

1. Although both mean scores and gains in biological knowledge at the [redacted] of the tenth grade were not meeting expected levels, BSCS students did perform significantly better.
2. Eleventh grade BSCS students outperformed non-BSCS students on the SPI.
3. BSCS students scored significantly better than non-BSCS students on most areas of inquiry skills measured by the Biology Process Test.
4. BSCS students did not exhibit the "deficiencies" of non-BSCS students on the practical mode of the matriculation test.
5. There was a significant difference between pupils of European and non-European parentage (in favor of pupils of European descent) in all types of cognitive measures in both experimental runs.

Interpretations

1. The BSCS adaptation curriculum was found as successful as other biology courses in Israeli schools. In fact, BSCS was found superior in a few cases.
2. BSCS was found more adaptable to city and rural schools than to agricultural schools.
3. BSCS adaptation was "less adequate" for lower IQ students and students of non-European descent.
4. To predict achievement, a variety of measures should be utilized.

ABSTRACTOR'S ANALYSIS

This study provided an extensive description of the effect of the BSCS Adaptation Program in Israel during the first six years of implementation. It is important to note that the Adaptation Program referred to a translated version of BSCS Yellow which was modified according to local flora, fauna and culture. Further, to obtain research data, several of the tests were locally constructed. Consequently, the generalizability of results is severely limited but the research methodology is not.

The longitudinal nature of this study provided copious data. The researchers considered many dependent variables—achievement, inquiry skills, understanding of process and attitude. The results were not inconsistent with other research comparing inquiry to traditionally oriented curricula. In some regards BSCS instruction was found superior to the conventional instruction. As for attitude and achievement by biology majors, definitive conclusions are lacking. Likewise a predictor of achievement was not found. The description of inquiry skills did not produce strong evidence favoring either form of biology instruction.

Perhaps the most outstanding feature of the work is the dependence on locally constructed tests. It must be assumed that (a) the locally constructed tests are both valid and reliable, and (b) as they are revised, these tests may become a standard for future research in other countries. Further, the use of criterion referenced tests provided the researchers the means to assess the impact of BSCS. It should be noted that a few normative tests—the SPI and BSCS Standardized—were utilized. With the SPI (and also the TOUS), Israeli students "compared favorably with U.S. and Australian students."

Although the implications were limited in scope, questions for future research have arisen. Examples which can be applied to most educational settings include:

1. Does the BSCS curriculum favor students with particular learning styles?
2. Since students involved in this study experienced knowledge of chemistry and physics, to what extent does concurrent chemistry instruction enhance BSCS achievement?

3. What are valid predictors of BSCS achievement?

4. Are predictors of BSCS achievement identical to predictors of conventional biology achievement?

5. How is the acquisition of process skills related to BSCS achievement?

6. To what extent does the attitude of BSCS students differ from that of conventional students?

7. What teaching characteristics are conducive to more effective BSCS achievement?

This study provided a means to assess the impact of a newly implemented curriculum. The procedure was sound and, coupled with the data, has provided worthwhile questions for future research.

Williams, Russell B. "A Method of Integrating Chemistry Subject Matter into Biological Science Instruction." Science Education, 59(2):167-175, 1975.
Descriptors--Biological Sciences; *Chemistry; *College Science; Educational Research; Higher Education; Instruction; *Learning; *Memory; *Science Education

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

Purpose

There were two main purposes of this study. One objective was to determine the extent of retention of chemistry knowledge among physiology students in a two-year community college by type of chemistry course background and by the length of time since completion of the chemistry course. Also, beginning physiology students were compared with finishing chemistry students to determine the areas of chemistry most frequently forgotten. The second objective was to design a physiology laboratory with a chemistry review integrated into it and to compare the student's performance with a control group taking a traditional physiology laboratory.

Rationale

The basis for the research was to attempt to improve student performance in physiology by using a unified science approach. One goal of unified science is to have pupils achieve a greater understanding of science in both the cognitive and affective domains. Physiology students at Southwest College, which has an open door policy, tend to have poorly developed skills despite the fact that chemistry is a prerequisite for physiology. The author cites evidence that problem solving skills developed in one class carry over to another class. Thus, it is suggested that inadequate preparation in chemistry may be related to the low achievement in physiology. The investigator expected the integrated course to affect the student's appreciation, understanding, and enjoyment of science as well as to improve his cognitive performance.

Research Design and Procedure

Development of Integrated Laboratory Exercises and Achievement Tests

The physiology labs that were developed consisted of a set of objectives, a materials list, a chemistry review and the physiology laboratory procedures. The content of the review was determined by examining the knowledge of chemistry required in the traditional physiology labs. By utilizing an outline format, simplified instructions and background information in paragraph form with a question and answer section, no additional time was required for the experiments.

A multiple choice chemistry pretest was developed to measure eleven areas in chemistry. A physiology achievement test was formulated from three introductory physiology tests to measure student performance of the experimental and control groups at the end of the course. The contents of both tests were validated by a panel of experts and the reliability was measured by split halves and the Spearman Brown formula.

Sampling

The subjects were nursing students enrolled in physiology classes which ranged in size from 8 to 28. The student body of the college is predominantly black. The number of classes was not specifically described but it appears that there were two experimental groups which had the chemistry review and two control groups which did not.

Analysis

The investigator administered the chemistry pretest to measure the first set of objectives which were related to chemistry knowledge retention. The test results were analyzed with (multiple?) t-test according to background and time. The chemistry background was categorized into three groups depending on where the students had taken chemistry. These were: A) in

high school; B) at Southwest College; C) at a school other than A or B. On the basis of time students were also placed in three groups: A) students who had taken chemistry within twelve months; B) students who had taken chemistry within 13 to 48 months; C) students who had taken chemistry more than 48 months ago. The differences were measured by t-test.

The experimental phase of the research was a post test only design. To test the differences in student performance between the control groups who did not get chemistry review and the experimental groups who did, a physiology test was administered and the results were analyzed by t-tests.

Findings

1. For retention of chemistry knowledge according to background no significant differences between groups were found. The means for the groups were: A) high school chemistry, 25; B) Southwest College, 26; C) other college chemistry, 26.6.

2. For retention of chemistry knowledge on the basis of time, the mean for Group A who had taken chemistry within 12 months was significantly higher than for Group B who had taken chemistry within 13 to 48 months or for Group C who had chemistry more than 48 months ago. The respective means for the three groups were 32, 20.8 and 23.6.

3. The investigator found that a loss of chemistry knowledge occurred since the finishing chemistry students scored significantly higher on the chemistry test than did the beginning physiology students. The means were 38.9 and 25, respectively. Of the 11 areas tested, some were forgotten more than others, especially problem solving areas. The author did not report the number of items on the test.

4. The investigator reported that the experimental groups scored significantly higher than did the control groups. The means of the two experimental groups were 65.42 and 61.61 while the control group means were 59.90 and 53.29. The t value was not given but the author stated that it was significant at the 0.05 level.

Interpretations

From this study the investigator drew the following conclusions:

1. the chemistry backgrounds of students from different schools were comparable;

2. students who took physiology soon after completing chemistry had less trouble in physiology;

3. some areas of chemistry were recalled more easily than others, the most difficult being those requiring analytical skills.

Several two-year nursing programs are considering eliminating chemistry as a prerequisite for physiology. This study suggests that knowledge of chemistry is important for a successful level of performance in physiology.

ABSTRACTOR'S ANALYSIS

This study addresses a problem that seems to be rather widespread; that is, the performance problems of students of science in colleges with an open door policy. Frequently the analytical skills the students need for success in courses are not well developed. For this reason, and because there is evidence that skills developed in one science course carry over to another, Williams designed an integrated course. The body of research studies in the development of unified science is rather small and this study could add to that, except that the reporting of the data is too incomplete for the study to be of empirical value.

Although the study potentially has much to offer to the reader interested in unified science and health related curricula, there are several weaknesses which reduce the significance of the study. One relatively minor concern is that the author concentrated on the affective outcomes in his rationale and literature review, but measured only cognitive outcomes. No affective outcomes were evaluated, even informally.

The major deficiencies of the study were in the statistical analysis and the reporting of the data. Sample sizes were not reported specifically and statistical tables were included for only one of the four research objectives. The investigator stated that the reliability of the instruments was determined but failed to report the results. In order to be useful to the reader, the reporting of the results should be complete enough for an independent interpretation. If a mean is reported without either a range or standard deviation, it is meaningless. In this paper, since the author did not report these or the t values, one simply has to take his word that the findings are valid, for there is no way to make an independent judgment.

Another fault of the study is the statistical analysis of the data. It appears that multiple t test were used for analysis since there were three groups compared for the first two objectives and four groups for the last. An analysis of variance would have been more appropriate and might have made a difference in the findings, particularly for the fourth objective. Multiple t-tests sometimes show significance that doesn't hold up under more rigorous analysis. It would have been interesting to see the results if the variables, background and time, had been analyzed simultaneously with an ANOVA. Although no significant differences occurred between background groups, a significant difference was found for time variations. It seems likely that most students in the high school chemistry group would not have had chemistry within 12 months and yet their scores were not significantly different from the other two groups. Without knowing the number of items on the test, it is difficult to determine what those scores mean.

For the analysis of the experimental phase of the study, Williams did not report the sample size, the number of items on the physiology test, or the t values. Because of this, and because he did not describe how he analyzed the results except by t-tests, the validity of the results is questionable. If he had four groups, two experimental and two control, with respective means of 65.42, 61.61, 59.90 and 53.29, then he must have collapsed the experimental and control groups into two groups or he must have done multiple t-tests. If he collapsed the groups, no means were given and the mean would not have been the average of the two since the sample sizes were unequal. If multiple t-tests were used, the difference between the experimental group, 61.61, and the control group, 59.0, does not look like it

would be significant. In either case, without more information it is impossible to make an accurate assessment which places the validity of the whole study in jeopardy.

This study could have made a valuable empirical contribution by providing evidence for the notions that knowledge learned in one science course has carryover value to another, that chemistry is a necessary prerequisite for some of the life sciences, that a unified approach is an appropriate means of increasing student performance, etc. However, the limitations of the reporting and analysis undermine even the basic findings.