DOCUMENT RESUME

ED 259 903 SE 045 872

AUTHOR Blosser, Patricia E., Ed.; Helgeson, Stanley L., Ed.
TITLE Investigations in Science Education. Volume 11, Number 2.
INSTITUTION ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.; Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 85
NOTE 73p.
AVAILABLE FROM SMEAC Information Reference Center, The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (subscription $8.00, $2.75 single copy).
PUB TYPE Collected Works - Serials (022) -- Reports - Research/Technical (143) -- Information Analyses - ERIC Information Analysis Products (071)
JOURNAL CIT Investigations in Science Education; v11 n2 1985

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; Career Education; Classroom Environment; Cognitive Style; Computer Assisted Instruction; Elementary Secondary Education; Environmental Education; Higher Education; Science Curriculum; *Science Education; *Science Instruction; *Student Attitudes; *Teacher Attitudes; *Teacher Education
IDENTIFIERS *Science Education Research

ABSTRACT

Abstracts and abstractors' critiques are presented for 11 articles related to instruction, attitudes, and teacher education. The research studies examined: (1) an instructional strategy incorporating process and conceptual knowledge to promote transfer of learning; (2) effects of instruction using a Visual Response System on achievement of male juvenile delinquents; (3) student's general reasoning ability and usefulness of pictures in science textual material; (4) classroom climates (degree of freedom provided by the teacher) and content achievement in college-level science; (5) chemistry students' cognitive preference and patterns of achievement; (6) student perceptions of instructional effectiveness of small group discussion, role-playing activities, and computer-based simulations on student competencies in dealing with classroom transactions; (7) effects of an environmental science education program on inservice teachers' attitudes; (8) a survey of K-12 science educators' attitudes and practices on career education; (9) hierarchical structures of environmental concerns; (10) formal/informal aspects of the supervisor's role and supervision effectiveness; and (11) workshops designed to improve teacher knowledge about energy. (DH)

SCIENCE EDUCATION INFORMATION REPORT

INVESTIGATIONS IN
SCIENCE EDUCATION

Volume 11, Number 2, 1985

THE ERIC SCIENCE, MATHEMATICS AND ENVIRONMENTAL EDUCATION CLEARINGHOUSE
in cooperation with
Center for Science and Mathematics Education
The Ohio State University

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

Clearinghouse for Science, Mathematics, and
Environmental Education

INVESTIGATIONS IN
SCIENCE EDUCATION

Volume 11, Number 2, 1985

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: $8.00 per year. Single Copy Price: $2.75.
Add $1.00 for Canadian mailings and $3.00 for foreign mailings.

Volume 11, Number 2, 1985

NOTES FROM THE EDITOR:

This issue contains critiques of articles related to instruction, to attitudes, and to teacher education. Within the section of articles of research focused on instruction, Finley and Smith reported on an instructional strategy incorporating the interrelation of process and conceptual knowledge, Test and Heward discussed the success of a visual response system used with juvenile delinquents, Koran and Koran studied the usefulness of adjunct pictures in science textual material, and Haukoos and Penick examined the effects of different classroom climates on college students' science learning. Within the attitude section, McNaught studied the relation of college students' cognitive preferences and chemistry achievement, Reynolds and Simpson examined student reactions to computer-based simulations, Bethel et al. reported on the effects of an inservice program on teachers, Smith surveyed science educators' attitudes about career education, and Richmond and Baumgart described a heirarchical analysis of environmental attitudes. Two critiques of research involving teacher education conclude this issue: Ritz et al. looked at perceptions of supervisor effectiveness, and Van Koevering and Sell evaluated an energy education workshop for teachers.

Patricia E. Blosser
Editor

Stanley L. Helgeson
Associate Editor

INSTRUCTION

Finley, Fred N. and Edward L. Smith. "Student Performance Resulting From Strategy-Based Instruction in a Sequence of Conceptually Related Tasks." Journal of Research in Science Teaching, 17 (6): 583-598, 1980.

Descriptors--*Earth Science; Elementary Education; Elem School Science; Geology; Grade 4; Science Education; *Science Instruction; *Transfer of Training

Expanded abstract and analysis prepared especially for I.S.E. by Linda R. DeTure, Rollins College.

Purpose

The purpose of the study was to design an instructional strategy that incorporated the interrelatedness of process and conceptual knowledge. The intent was to improve transfer of knowledge to more complex tasks. The authors' aim was to examine the effects of the approach via a sequence of tasks for classifying igneous rock samples. Specifically, two questions were asked: a) Does instruction on prior parallel task enhance learning and transfer to subsequent tasks? b) Does explicit instruction on task strategies enhance learning and transfer in a sequence of task within a conceptual network?

Rationale

The research project sought to extend and further test the research of Bessemer and Smith who proposed that transfer of learning could be improved by designing instruction that incorporated conceptual networks, tasks and strategies. Smith et. al. demonstrated the probability of vertical transfer, thus providing a mechanism for the transfer of learning from one task to another. Padilla and Smith found that transfer did result for a sequence of parallel conceptual network and that explicit teaching of generalizable strategies did enhance learning. These studies provided the framework for developing the specific instructional strategies of this study. To answer the research questions, the authors manipulated the sequence of instruction on parallel tasks and the type of instruction, i.e., whether or not it

dealt with specific task strategies. The dependent variables examined included pretest scores for accuracy, trials-to-criterion and posttest scores for accuracy.

Research Design and Procedure

A sample of 120 fourth graders from heterogenous backgrounds was randomly assigned to five treatment groups of 24 students each. For each of the independent variables, instruction and sequence, experimental and control conditions were defined. The experimental strategy (S) instruction involved explicit instruction on the three tasks. In the control instruction, Feedback (F) instruction, students were not told how to carry out the task, but received feedback as to whether the rock samples were correctly placed. The experimental sequence treatment, Cumulative (C), required students to be instructed on all three tasks. In the control, Isolated (I), students received instruction on a single task.

The three tasks are described on two levels, analytical and context specific concepts. The nature of the tasks is hierarchial in that the first task is a subroutine of the second which is, in turn, utilized in the third task. The Comparison to Standard task requires students to compare an unknown with a standard and verbally report the results. The second Single Variable Classification task requires students to classify samples by grain size utilizing a three cell scheme. The Double Classification task involved classifying igneous rock samples in a 3X3, 9 cell classification scheme based on two variables, grain size and a qualitative description of the amount of light colored grain.

The hypotheses were tested with a posttest only control group design. Three days prior to the treatment all students received preliminary instruction in the concepts needed to perform the task. The pretest-treatment-posttest cycle followed. All three instructional task series were completed in three weeks. The t-statistic was

utilized to compare Cumulative Strategy and Cumulative Feedback group means for the Comparison to Standard task. A planned comparison ANOVA was used for the Single Variable task and a two-way ANOVA was utilized for group performance on the Double Variable Classification task.

Findings

The results are described according to the three dependent variables; pretest accuracy, representing transfer of learning from previous tasks; trials-to-criterion indicating the facilitation of learning during instruction; and posttest accuracy to determine competence of students to perform the tasks.

Pretest accuracy was assessed for both the Single and Double Classification tasks. When examining Cumulative Strategy and Cumulative Feedback versus Isolated S and F for the Single Classification task, no significant differences were found for either sequence or instruction. However, for the Double Classification task, a significant difference occurred for cumulative instruction ($p = 00$) but no significant difference was noted for type of instruction, nor did significant interactions occur.

For comparisons of groups on trials-to-criterion no significant differences were found for either the Comparison to Standard task or for the Single Classification task in regard to type of instruction. No significance was found for sequence in the second task. For the Double Classification task the results were similar to the pretest accuracy findings. A significant difference occurred for sequence ($p = .00$) but none for type of instruction or interaction.

The students' ability to perform the tasks was reflected in the posttest accuracy scores. Again for Comparison to Standard and Single Classification the differences were not significant. There was improvement of approximately one sample over pretest performances and the researchers felt a possible ceiling effect was taking place.

For the Double Variable Classification task a significant main effects difference occurred from instruction within the sequence, but the type of instruction did not differ significantly. A significant interaction ($p = 0.02$) was found between sequence and instruction, with Strategy instruction within the task sequence being most effective.

In summary, only the Double Variable Classification task resulted in significant differences. Transfer of Learning was enhanced by instruction on the two previous tasks but not by the strategy based instruction. For Facilitation of Learning, participants of the entire task sequence performed the Double Variable Classification task more easily than students who did not. Again, strategy based instruction did not make a difference. For task performance, students receiving instruction within the task sequence and instruction strategy within the sequence had a greater advantage.

Interpretations

The authors conclude that the results support their approach to instructional design based on designing conceptually related tasks and using strategy based instruction because both learning and transfer were heightened during the final completion of the geologic task. When teaching complex science tasks, it would appear to be useful to teach simple but conceptually related tasks first. Also it may be effective to teach strategies for performing simpler tasks if the strategies can be utilized for a similar higher level task. The results indicated that students receiving instruction within the sequence were better able to transfer what they had learned to more complex tasks. Also, strategy instruction within the task sequence resulted in the best performance of the four experimental groups.

While the above results support the hypotheses regarding planning, a note of caution is encouraged. The benefits of the treatments did not emerge in early tasks and the overall magnitude of difference between groups for the final performance was moderate.

ABSTRACTOR'S ANALYSIS

An underlying assumption of education today is the expectation that the transfer of learning to more complex tasks readily and/or automatically takes place. An observable example of the fallacy of this assumption is often noted in mathematics with children who understand the numerical mechanics of problem but cannot decipher the same problem when it is presented in word form. Finley and Smith are to be commended for addressing the transfer of learning as a problem to be solved in itself and for going beyond the unidirectional approach of process versus content. The sparsity of the related studies in the literature review is conceivably indicative of the assumption of generalizability and certainly a lack of the problem being regarded as a multifaceted issue of instructional design.

Conceptually this study is aimed at the heart of the instructional process by addressing a central question: Can instruction be designed in such a way that the transfer of learning can be reasonably assured? While the question will not be answered unequivocally with this study, or even a series of studies, positive support should encourage researchers to continue to ferret out the parameters of the issue. How people go about solving problems remains an intriguing question that makes designing instruction to provide children with strategies for solving problems an interesting approach. However, a potential problem with the approach is the likelihood that a multitude of strategies for any given problem exists. In formulating strategies, some will surely be more effective than others and children with effective strategies will be more successful problem solvers. In this study, the subjects were taught a strategy which may not have been one they would have derived. Learning the strategy could have had either beneficial or deleterious effects depending on the nature of the learner. This thinking is highly speculative, but given the diversity of problem solving approaches, aptitude and treatment somehow ought to be taken into consideration. The results of the study, while positive and encouraging, were moderate enough to leave room for

possille alternate explanations. Perhaps even, the ATI effects partially nullified the significance of the strategy treatment. In the related Padilla-Smith (1979) study only one group profited more from strategy training than from practice and feedback and this advantage was not evident in later months. Still, teaching strategies as a means to learn more about problem solving seems to be an area that warrants additional study.

When examining the research report there were some areas of concern. The description of the methodology was not altogether clear. The table of task descriptions was most helpful because it briefly identifies what the children were expected to do. The description of instructional treatments was not so clear. In the population and sample selection section, the authors reported that children were randomly assigned to five treatment groups of 24 students each. However, only four treatments were defined; the Strategy versus Feedback instruction and Cumulative versus Isolated sequence. How that fifth group functioned is not specified. In addition, the experimental treatment for sequence was not adequately defined particularly the isolated condition, e.g., which task was not defined until later in the paper. Since sequence was ultimately the variable to be significant in the posttest accuracy groups, a more detailed description or summative table of treatments should have been included.

This study was fairly ambitious in scope and complex in that the tasks, the treatments and the statistical analysis were layered and hierarchial. From the written report, I found it difficult to always follow what the researchers had done. None of the non-significant results were reported in table form which is understandable in view of space requirements for journals. But, a summary table of the various means and standard deviations would certainly be worthwhile for the reader of the report.

In the results section for pretest accuracy the means refer to the number of samples placed correctly. No significant group differences occurred for the single classification task. With the double variable classification Cumulative instruction was more

effective for the correct placement. In the trials-to-criterion, it appears that the means refer to the number of trials rather than the number of samples correctly placed. If not, then the isolated sequence group performed significantly better than the cumulative group. Clarification is needed. Results of the posttest accuracy were easy to follow. Based on the description of the experimental design, the statistical analyses used to test the hypotheses were appropriate.

Overall, this seems to be a well designed and thoroughly executed research study. The few weaknesses of the study are more related to a lack of detail in reporting the results than in faulty design. It behooves us as researchers to be cognizant that the reader can only interpret the report and not the actual research.

A very positive outcome of this study is that it directs the reader's attention to the complex problem of designing instruction to maximize the ease of learning and the transfer of knowledge to new tasks. The potential for further research is great. Identifying cognitive problem solving strategies is a task in itself. When that is superimposed on individual differences and instructional design, a large and futile research area is exposed.

REFERENCE

Padilla, M. J. and E. L. Smith. "Experimental Results of Teaching First Grade Children Strategies for Nonvisual Seriation." Journal of Research in Science Teaching, 16: 339-345, 1979.

Test, David W. and William L. Heward. "Photosynthesis: Teaching a Complex Science Concept to Juvenile Delinquents." Science Education, 64 (2): 129-139, 1980.

Descriptors--Audiovisual Aids; Biology; *Delinquency; *Educational Equipment; *Educational Research; *Instructional Innovation; Overhead Projectors; *Photosynthesis; Science Education; *Science Instruction; Secondary Education; Secondary School Science

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 determine the effects of instruction using a Visual Response System (VRS) and photosynthesis program on the achievement of male juvenile delinquents enrolled in a science class at a residential correctional facility.

Rationale

The investigation was conducted within the conceptual framework of applied behavioral analysis. It was related to a variety of previous research utilizing behavioral techniques to increase student performance in academic skill areas, including McCarty et al. (1977), Van Houten and Thompson (1976), Ballard and Glynn (1975), Fawcett and Fletcher (1977), Neef et al. (1977), Rapport and Bostow (1976), Knapczyk and Livingston (1973) (1974), Lahey et al. (1973), and Tyler and Brown (1968). Use of the VRS as an instructional strategy for special education was related to previous research by Barrette (1971), Eachus (1971), Heward and Eachus (in press), Shadding (1979), Joynes et al. (1978) and Marshall (1979).

Research Design and Procedure

A multiple baseline across behaviors desig.. was utilized to determine the effectiveness of an instructional program designed to teach selected science concepts to juvenile delinquents in a correctional institute. In this type of design, behavioral data are gathered simultaneously on two or more independent but functionally related behaviors. These target behaviors serve as dependent variables. Once it is determined that these dependent variables maintain an acceptable level of stability, an intervention or treatment is applied to one of the dependent variables while monitoring the others. When the dependent variable reaches a previously determined criterion level, intervention is applied to the second dependent variable while monitoring the others. The procedure is continued until intervention is applied to each of the dependent variables.

Seven male juvenile delinquents, ages 13 to 18, who were enrolled in a science class at a correctional facility, were selected to participate in the study based on their performance on a pretest covering photosynthesis. The test was given to identify students who would benefit most from the instructional program.

A four part unit on photosynthesis was developed and taught to the students in a classroom containing an eight-station VRS (For a detailed description of a VRS see Heward, 1978). Each student desk contained an overhead projector and feedback unit, including a red light, green light, teacher call button and electrically operated counter. The teacher's desk contained a sound system and master switches for the overhead projector and counter. Instructional materials included a master test, student response slides, teacher overhead transparencies, atom transparencies and student response transparencies. Teaching each part of the unit served as the four interventions or treatments.

The target behaviors, one set corresponding to each part of the unit, were measured by administering the master test. The behaviors included: (1) writing the three major parts of a green plant and

stating two functions of each part; (2) writing the chemical symbols for the atoms and molecules in the photosynthesis equation, and, given the symbol for a compound, calculating and writing the number of molecules and atoms in the compound; (3) writing the photosynthesis cycle; and (4) writing a definition of photosynthesis that includes seven distinct components. The master test served as a pretest and was administered six additional times, once after each intervention and as a three and six-week posttest. Only three students were available for the six-week posttest. Data from the testing were used in the multiple baseline analysis.

Each intervention was approximately 40 minutes in length. Five daily 40 minute instruction sessions were required for the four interventions because of the extra time needed for administering the test. The general teaching procedure followed a stimulus-student response-feedback model. Students were required to complete each step with 90 percent accuracy before going on to the next step. Students were provided with points for correct responses. Points were exchanged for cigarettes or lottery tickets for soft drinks and post s. A questionnaire about the program was administered to the students after the three-week follow-up test.

Findings

The investigators reported that the students averaged 6.3 questions correct (range 0-11) on the overall pretest. The students averaged 29.5 questions correct (range 23-31) on all four parts of the criterion measure on the probe following intervention. They averaged 25.4 questions correct (range 13-31) three weeks following treatment. On the six-week follow-up, the students averaged 27.3 questions correct (range 21-31) on all four parts of the criterion instrument. Student responses on the questionnaire revealed that they were positive toward the VRS and photosynthesis unit.

Interpretations

The investigators concluded that the instructional program was effective in teaching selected concepts in a photosynthesis unit to juvenile delinquents. Students retained the knowledge and were able to generalize what they had learned across academic settings. Students were positive toward the VRS and photosynthesis unit. It was further concluded that the instructional program enabled the teacher to incorporate into the program positive features such as higher rates of student response, more immediate feedback, an increase in variety of responses, and greater number of student-student interactions.

ABSTRACTOR'S ANALYSIS

The following analysis includes an evaluation of whether or not the investigation met several criteria stated by Baer, Wolf and Risley (1968) as being essential for any applied behavior analysis research. Such research should be applied, behavioral, analytic, technological, conceptually systematic, effective and display generality. The intent of the evaluation is to provide greater insight into the research methodology used by Test and Heward and provide guidelines for science educators who wish to use applied behavior anlaysis in their future research efforts.

Knowledge of science concepts is very important to individuals if they are to function adequately in a scientific and technological society. It is even more important to such a society that its individual members are scientifically literate. Knowledge of basic science concepts is an essential characteristic of a scientifically literate person. Baer, Wolf and Risley (1968) state that if behaviors

are selected for study because of their importance to the individual and society rather than theory, they may be classified as applied. Since juvenile delinquents are often low in academic achievement, the selection of their knowledge of science concepts for study by Test and Heward met the criterion of being applied.

A measure of the behavioral attributes of a study is whether or not the subject's behavior is significantly altered. The investigation was revealed as being behavioral when the subjects answered questions correctly, following intervention, that they had previously missed on the pretest.

In applied behavioral analysis, the criterion of being analytic refers primarily to whether or not the investigators demonstrate in a believable fashion, to the audience for whom the research is intended, that the intervention is responsible for the occurrence or non-occurence of the target behavior. Test and Heward met this portion of the criterion of being analytic. They provided convincing evidence that reliable control was exercised over the subjects' identified behaviors; i.e., the investigation exhibited a high level of internal validity.

Further explanation is needed concerning the criterion of being analytic because demonstrating enough control in applied research is often not as clear as it is when investigators use an experimental design. The two most common methods of demonstrating reliable control in applied research is the use of withdrawal and reversal designs or multiple baseline designs. (For a comprehensive description of these designs see Tawney and Gast, 1984.) Test and Heward chose the multiple baseline across behaviors design. In this design, internal validity is determined in a large part by the extent to which the investigators accomplished the following steps: (1) select three or more independent but functionally related behaviors, (2) measure the behaviors until a stable baseline is established, (3) intervene on one behavior while monitoring the others until a predetermined criterion level is reached, and (4) intervene on next behavior while monitoring the others (Tawney and Gast, 1984). Test and Heward completed these steps, and their data

show that the target behaviors were reliably altered by the intervention. In other words, in applied research it is necessary to examine the results as well as the design to determine if the investigation has internal validity. In experimental research, internal validity is determined primarily by the design itself.

Another aspect of being analytic refers to the simplification and separation of component processes (Baer, Wolf and Risley, 1968). Is the intervention a single component, or are the component parts of the intervention separated and studied so that the effectiveness of each component can be identified? Test and Heward did not meet this criterion. Their intervention was successful, but it was not a single component. It cannot be determined, for example, if the VRS, teaching unit, feedback, active participation, or reward system were the most effective components of the intervention. It was in fact a multiple treatment. The next step should be additional research to determine the degree of effectiveness of each of these components.

If only the published report were considered, the investigation would not meet the criterion of being technological. This criterion requires that an investigation could be replicated well enough to obtain similar results by a typically trained person who is given only a description of the procedure. The intervention by Test and Heward was complex and would require a fairly lengthy and detailed description to replicate. Published reports are by nature brief. The investigation may meet the criterion of being technological, but this cannot be determined without a more complete description of the procedure.

The investigation met the conceptual systems criterion. The procedure followed a stimulus-student response-feedback model and included various general principles of learning, rather than being simply a bag of tricks. For example, the intervention included active student participation, immediate feedback and rewards for desirable behavior.

The criteria of being effective and having generality were met by the investigation. The intervention was effective in altering the target behaviors enough to be socially important, and facts learned by the students were generalized across academic settings.

The investigation by Test and Heward should be extremely interesting to researchers in science education. They have used a research design that is conspicuous by its absence from science education research literature. However, this is not surprising because multiple baseline designs are a relatively new approach to research. They were introduced by Baer, Wolf and Risley in 1968 into the literature of applied behavior analysis (Tawney and Gast, 1984) and are only recently gaining acceptance by basic researchers as an objective approach for conducting research. The multiple baseline across behaviors is but one of the multiple baseline designs with which science educators will want to become familiar, especially those who are involved in special education and/or analytic behavioral application. Test and Heward provide science educators with an excellent introduction to a research design that has a great deal of potential. The investigation was well conceived, executed, and reported. In many respects, it can serve as a model implementation of a multiple baseline across behaviors design for science education researchers who are interested in single subject research.

Reviewing the investigation brought to mind several questions in need of further consideration and discussion by science education researchers. Are multiple baseline designs appropriate only for applied behavioral analysis in special education? Are the designs objective enough to be useful in experimental research involving any type of single subjects? Could it be possible that the designs are more appropriate than traditional experimental designs for a variety of situations found in science education research?

REFERENCES

Baer, D. M., M. M. Wolf, and T. R. Risley. "Some Current Dimensions of Applied Behavior Analysis." Journal of Applied Behavior Analysis, 1: 91-97, 1968.

Ballard, K. D. and T. Glynn. "Behavioral Self-Management in Story Writing with Elementary School Children." Journal of Applied Behavior Analysis, 5: 387-398, 1975.

Barrette, P. P. "The Effects of Contingency Management on the Rate, Accuracy, and Lenght of Sentence Composition of Deaf Children." Unpublished doctoral dissertation, University of Massachusetts, 1971.

Eachus, H. T. "Modification of Sentence Writing." American Annals of Deaf, 116: 25-43, 1971.

Fawcett, S. B. and R. K. Fletcher. "Community Applications of Instructional Technology: Training Writers of Instructional Packages." Journal of Applied Behavior Analysis, 10: 739-746, 1977.

Heward, W. L. "Visual Response System." Exceptional Children, 44: 466-468, 1978.

Heward, W. L. and H. T. Eachus. "Acquisition of Adjectives and Adverbs to Sentences Written by Hearing Impaired and Aphasic Children." Journal of Applied Behavior Analysis, in press.

Joynes, Y., S. H. McCormick, and W. L. Heward. "Teaching Mildly Retarded, Juvenile Delinquents to Complete Employment Applications: Evaluation of a Mediated Instructional Program." Manuscript submitted for publication, 1979.

Knapczyk, D. R. and G. Livingston. "Self-Reading and Student-Teacher Supervision: Variables with a Token Economy Structure." Journal of Applied Behavior Analysis, 6: 481-486, 1973.

Knapczyk, D. R. and G. Livingston. "The Effects of Prompting Question-Asking upon On-Task Behavior and Reading Comprehension." Journal of Applied Behavior Analysis, 7: 115-121, 1974.

Lahey, B. B., M. P. McNees, and C. C. Brown. "Modifications of Deficits in Reading for Comprehension." Journal of Applied Behavior Analysis, 6: 475-480, 19.3.

Marshall, A. E. "Teaching Self-Management Skills to Juvenile Delinquents in the Visual Response System." Unpublished masters thesis. The Ohio State University, 1979.

McCarty, T., S. Griffin, T. Apolloni, and R. E. Shores. "Increased Peer-Teaching with Group-Oriented Contingencies for Arithmetic Performance in Behavior-Disordered Adolescents." Journal of Applied Behavior Analysis, 10: 313, 1977.

Neef, N. A., B. A. Iwata, and T. J. Page. "The Effects of Known Item Interspersal on Acquisition and Retention of Spelling and Sight-Reading Words." Journal of Applied Behavior Analysis, 10: 738, 1977.

Rapport, M. D. and D. E. Bostow. "The Effects of Access to Special Activities on the Performance in Four Categories of Academic Tasks with Third Grade Students." Journal of Applied Behavior Analysis, 9: 372, 1976.

Shadding, A. L. "Evaluation of a Mediated Instructional Program to Teach Computation of Math Fractions." Unpublished master's thesis, The Ohio State University, 1979.

Tawney, J. W. and D. L. Gast. Single Subject Research in Special Education. Columbus, OH: Charles E. Merrill Publishing Co., 1984.

Tyler, V. O. and G. D. Brown. "Token Reinforcement of Academic Performance with Institutional Delinquent Boys." Journal of Educational Psychology, 14: 413-423, 1968.

Van Houton, R. and C. Thompson. "The Effects of Explicit Timing on Math Performance." Journal of Applied Behavior Analysis, 9: 227-230, 1976.

Koran, Mary Lou and John Koran. "Interaction of Learner Characteristics With the Pictorial Adjuncts in Learning From Science Text." Journal of Research in Science Teaching, 17 (5): 477-483, 1980.

Descriptors--*Academic Aptitude; Aptitude Treatment Interaction; General Science; Junior High Schools; *Pictorial Stimuli; *Reading Research; Science Education; *Science Materials; Secondary Education; Secondary School Science; *Student Characteristics

Expanded abstract and analysis prepared especially for I.S.E. by David P. Butts, University of Georgia.

Purpose and Rationale

In this report the investigators proposed to study the relationship between a student's general reasoning ability and the usefulness of pictures in science textual material. Because pictures can serve a variety of functions for a reader--an advance organizer, a reinforcement, or a review--and since much science text material has a large number of complex relationships, it is reasonable to expect that pictures would be a valuable aid in helping the student learn these relationships. Furthermore, this help might indeed be related to the student's ability in that pictures before would help focus the attention of the student to more concrete expressions of the concept. Pictures during or after might be expected to be a source of reinforcement or review for the high ability student. Thus it could be expected that the use of pictures in a text could have contrasting effects with students of different aptitudes--a classic example of aptitude treatment interactions in the science classroom.

Research Design

Using a posttest only control group design, three treatments were used.

R X_a O_1

R X_b O_2

R X_c O_3

where: X_a = picture before text

X_b = picture after text

X_c = no picture

$O_{1,2,3}$ = Test of General Intelligence
Criterion test on water cycle

Findings

Conclusions were based on results for 87 seventh and eighth grade students who had access to the text materials and then were given a criterion test. Since the treatment consisted of a four-page text and then a test, the time of the experiment was assumed to be less than one hour.

The findings of the study were essentially

--low ability students benefit more from the pictures regardless of their placement, before or after text material

--high ability students performed best with no pictures.

Interpretations

The authors conclude that pictures benefit low-ability students. They extended the conclusion in an explanation that this occurs because a picture provides a means that helps students deal with abstract words in the text--and in using pictures, it helps students to store and retrieve new ideas.

ABSTRACTOR'S ANALYSIS

In this study the researchers have made important first steps on a significant issue. The widespread use of textbooks in science instruction is well documented. The enormous investment in these texts assumes that we have a solid empirical basis for their design. It also assumes that one design is usable with all students. Koran and Koran's study here provides a significant clue that this assumption needs much further study. In a simple and clearly designed study, they provide evidence that pictures in a text are not equally useful for all students. The word "useful" here, however, represents a challenge. The conclusion of this study would be strengthened if evidence were more explicit on what was measured in the criterion measure. Further, the conclusions would be strengthened if they were supported by several instances, rather than being limited to a 60 minute involvement with a single concept. A somewhat side issue is seen in the authors' explanation of their findings. These explanations include how students use pictures which really represent hypotheses in need of further study. This study represents excellent "first words" for continuing research.

Haukoos, Gerry D. and John E. Penick. "The Influence of Classroom Climate on Science Process and Content Achievement of Community College Students." Journal of Research in Science Teaching, 20 (7): 629-637, 1983.

Descriptors--*Academic Achievement; *Biology; *Classroom Environment; *College Science; *Discovery Learning; Higher Education; *Process Education; Science Education; Science Instruction; Two-Year College Students

Expanded abstract and analysis prepared especially for I.S.E. by David R. Stronck, California State University, Hayward.

Purpose

The purpose of this study was to measure the effect of two specific classroom climates on learning of science process skills and content achievement in college level science classes. The two classroom climates were described as discovery classroom climate (DCC) and nondiscovery classroom climate (NDCC). In this context, the term discovery means the degree of freedom that the teacher provides in classroom interactions.

Rationale

A search of the literature shows that classroom climate studies have been limited mostly to science teaching in elementary and secondary classrooms or teacher education programs. This study extended the research on classroom climate into the postsecondary situation. The variable of the degree of freedom in classroom interactions was considered in terms of its influence on the science process skills and content achievements of community college students.

Flanders (1965) demonstrated that students achieved more in an indirect classroom climate. Indirect verbal behavior of the teacher, e.g., eliciting student statements, listening and accepting student ideas, increases the freedom of the students and encourages them to show a variety of responses. Direct verbal behavior by lecturing, giving directions, reciting facts, criticizing, or praising, tends to reduce the variety of possible student responses.

Several studies demonstrate increases in the learning of process skills when the teacher's behavior is less directive. For example, Raghubir (1979) showed that students in the discovery classroom climate made significant gains in formulating hypotheses, making assumptions, designing and executing investigations, understanding variables, observing carefully, recording data, analyzing and interpreting results, and synthesizing new knowledge, in comparison with students in a nondiscovery classroom climate.

Research Design and Procedure

The 78 subjects of this study were students at the College of DuPage, Glen Ellyn, Illinois, a two-year, comprehensive, community college. The subjects were enrolled in four intact sections of Principles of Biological Science, an introductory college biology course. Two randomly selected sections received a treatment designated Discovery Classroom Climate (DCC). Two others received a Nondiscovery Classroom Climate (NDCC). Each treatment was defined by the classroom behavior of the teacher and the nature of the activities performed by the students. Throughout the study, all subjects received exactly the same content although the classroom climates were of two types. Although the laboratory equipment and supplies were identical for the DCC and the NDCC, students in the DCC treatment were encouraged to select and explore an activity from an assortment of questions. The teacher of the DCC laboratory offered little direction or evaluation. On the other hand, students in the NDCC laboratory received exact directions for what to do with the laboratory materials and how to manipulate them. The NDCC teacher clearly praised or rejected their results. In the lecture/discussion portion of the course, the DCC instructor encouraged the students to discuss and interpret the projection of 35-mm photographic slides. During the NDCC sessions, the same slides were used but were presented with the impression that science was complete and final beyond the need for students' questioning.

One teacher, the primary investigator, did all of the instruction in this study. Daily audio taping of the teacher-student interactions allowed a systematic monitoring of the teacher's classroom behavior throughout the study. These interactions were then coded into the established categories of the Science Laboratory Interactions Categories (Shymansky and Penick, 1979). Then the interactions were compared with the criterion levels of teacher behavior for DCC and for NDCC. These data helped the teacher to maintain the desired treatment during the study according to the preestablished criteria for DCC and NDCC.

The research design was the pretest-posttest control group design. To measure the influence of classroom climate on students' achievement of science processes, the Science Inventory (SPI), Form D, was used. According to Welch and Pella (1967), this instrument of 135 items measures the understanding of methods and processes by which scientific knowledge evolves. They also describe the validity and reliability of this standardized instrument.

The Biology Achievement Test (BAT) was used as a posttest only for measuring the content achievement of the course. Neither formal validity of the test questions nor reliability of the test was measured. However, a committee of biology faculty at this community college designed this instrument to provide their Office of Testing a proficiency instrument for students seeking credit without taking the course. This committee established the face validity of the BAT.

Findings

An analysis of covariance (ANCOVA) was used to determine if significant differences existed between the DCC and the NDCC treatments. When significant results were obtained, Duncan's Multiple Range Test for variability was used to find where the differences were for the science process achievement data. For analysis of content achievement data, since only a posttest was given, an analysis of variance (ANOVA) was used to find possible differences among the four treatment sections.

In this study there were no significant differences between the groups on content achievement. Studies with pre-college students show a trend toward higher levels of achievement with indirect methods of teaching. The investigator observes that both groups of community college students had the same primary motivation: to complete the course with a passing grade. Probably children in the pre-college grades lack this adult attitude of responsibility and benefit from increased motivation through indirectness of instruction. Their increased motivation then leads to higher levels of content achievement.

The total number of students involved in this study was only 78. The number of students in each of the four sections ranged from 11 to 25. Usually students enrolled in community college courses are very heterogenous and are selecting specific sections on the basis of scheduling convenience. Although the statistical analysis of the study is correct, the small sample sizes raise suspicions, especially about conclusions where no significant differences were found. Perhaps additional studies using larger samples would find significant differences in content achievement between CCC students. Similarly, additional studies might find that five-week intensive classes produce the same results as ten-week classes in terms of learning science processes. The investigator concludes that sufficient time must be given to learn the science processes. However, he does not clearly emphasize that he is talking about length of time for the course, not time in the classroom. Certainly time of instruction within the intensive short course must be approximately the same as that for the ten-week course because both sections are for the same number of credits. There seems to be no logical reason why students in the intensive five-week version of the course failed to achieve significantly higher scores on the SPI, as did the students in the ten-week course. This difference may result from the small sample sizes. Intensive courses have many educational advantages; the Colorado College in Colorado Springs, Colorado, has a system of using only intensive courses throughout the academic year to provide these advantages to their students. Research studies favor intensive courses although this study seems to make the opposite conclusion.

The literature related to this study is large and presents a series of overlapping terms. Classroom climate may be defined as the dynamic interaction of all elements operating in the classroom. One of the major factors determining the classroom climate is the amount of intellectual freedom granted to students by the teacher. Flanders (1965) measured an aspect of this freedom by observing directness or indirectness in the teaching. This study used a different instrument to measure the teacher's behavior: Science Laboratory Interaction SLIC-Teacher (Shymansky and Penick, 1979). This instrument considers more behavior categories than does Flander's analysis. The investigator introduces the word "discovery" to describe the degree of freedom that the teacher provides in classroom interactions. Other researchers have used the term discovery to define the application of scientific processes. Perhaps the title of this study and the discussion within the study would have been clearer if the investigator had avoided increasing the use of terms and simply provided the title: "The Influence of Less-Directive Teaching Behavior on Science Process and Content Achievement of Community College Students." The use of the term "classroom climate" misleads the reader into anticipating a broader investigation of factors, e.g., socio-economic data.

This study does encourage attention to less-directive teaching to improve the students' understanding of science processes. Many other researchers have advocated the same goal. The chief value of this study is its use of students in the community college environment and therefore the extension of the goal into additional grades through this supporting data.

REFERENCES

Flanders, N. A. Teacher Influence, Pupil Attitudes, and Achievement. Washington, DC: Department of Health, Education, and Welfare, Cooperative Research Monograph No. 12, 1965. (O.E. 25040).

Raghubir, K. P. "The Laboratory-Investigative Approach to Science Instruction." Journal of Research in Science Teaching, 16 (1): 13-17, 1979.

Shymansky, J. A. and J. E. Penick. "The Use of Systematic Observations to Improve College Science Laboratory Instruction." Science Education, 63 (2): 195-203, 1979.

Welch, W. W. and M. O. Pella. "The Development of an Instrument for Inventorying Knowledge of the Process of Science." Journal of Research in Science Teaching, 5 (1): 64-68, 1967.

ATTITUDES

McNaught, Carmel. "Relationship Between Cognitive Preferences and Achievement in Chemistry." Journal of Research in Science Teaching, 19 (2): 177-186, 1982.

Descriptors--*Academic Achievement; *Chemistry; *Cognitive Style; Foreign Countries; Grade 12; High School Students; Science Education; Secondary Education; *Secondary School Science; *Student Characteristics

Expanded abstract and analysis prepared especially for I.S.E. by Glen S. Aikenhead, University of Saskatchewan.

Purpose

Are chemistry students' cognitive preferences related to their patterns of achievement and, if so, in what ways?

Rationale

The author used the Heath (1964) postulate that students attend to subject matter in four different modes, showing preferences for:

1. memorizing (M)
2. identifying a practical application (A)
3. questioning critically and challengingly (Q)
4. identifying the underlying principles (P).

Four previous studies employed the Heath system in conjunction with chemistry courses, all of which were the "new" courses of the early 1960's. The present study was a logical extension and refinement of this past research. The validity of the Heath system was assumed and not rationalized.

Research Design and Procedures

The Campbell and Stanley X O_1 O_2 design was used. The treatment X was two years of chemistry which had been preceded by a four-year integrated science curriculum. O_1 refers to a cognitive preference test while O_2 represents the Australian High School Certificate final test. Students wrote both within a month of each other. The sample comprised all grade 12 chemistry students (N = 304) in 18 schools randomly picked from within the Melbourne vicinity.

The author took advantage of the ipsative nature of the cognitive preference test and constructed scores that compared two preferences; for example, instead of using a P score and an M score, a PM score was calculated by: $P_{score} - M_{score}$. The Cronback reliability coefficients ranges from 0.7 to 0.25 for the individual preferences. The author mentioned that the low reliabilities would not adversely affect the analysis because they "were used only to allocate students to cognitive preference groups" and not used in the main covariance analysis. Statistical results implied that the students' cognitive preferences could be characterized in terms of either QM and PA scores, or QA and PM scores. Each of these four scores was correlated with O_2, the chemistry achievement test.

In order to refine the study further, the items on this achievement test were classified as:

1. Knowledge (K) or Higher than Knowledge (HK) using Bloom's taxonomy, and
2. nonmathematical completion items ("extended answer" -- EA) or multiple choice and numerical problems ("other" -- O).

In summary, different types of achievement items (K, HK, EA, O) were correlated with four cognitive preferences (QM, PA, QA and PM).

Findings

Correlations between the total chemistry achievement scores and student cognitive preference scores were: QA, 0.11; PM, 0.17; QM, 0.19; and PA, 0.04; with 0.17 and 0.19 being significant at the .05 probability level.

A complex analysis of covariance, looking at differences between K and HK items and between EA and O items, yielded the result that only on the PA scale did students show differences, and this only for K and HK items:

high PA students performed better on HK items than on K items,
low PA students performed better on K items than on HK items.

Next, the study moved to a more finely subdivided analysis by double classifying test items: HK/EA, K/EA, HK/O and HK/EA. Not only were the results of high and low PA students confirmed, but one more finding was added:

high PM students performed better on EA items than on O items
low PM students performed better on O items than on EA items.

Interpretations

The author states that the aims of the Australian High School Certificate chemistry course "imply that preferences for principles and critical questioning are more desirable than preference for memory." The findings of the present study concur with this implication. The following propositions were claimed (p. 184):

1. Students who develop the cognitive preferences consistent with those implied or stressed by the HSC chemistry course tend to perform better on an end-of-course achievement examination than do students whose cognitive preferences are not consistent with these implied preferences.

2. A strong cognitive preference is positively correlated with performance on the sorts of cognitive tasks implied in that preference.

The author cautioned that the development of cognitive preferences (whether as a result of the chemistry course or for other reasons) had not been examined. An implication for chemistry teachers was given: teachers ought to develop P and Q cognitive preferences in students if students are expected to perform on HK items. How this can be achieved is left for further study.

The author states (p. 186): "as with other studies in this area, cognitive preferences have again been demonstrated to be a relevant variable in relation to achievement in secondary school science courses."

ABSTRACTOR'S ANALYSIS

The high quality of the research design and procedures, and the fine clarity of the written report, allow one to discuss the investigation and its interpretations with confidence.

If the Heath cognitive preference system is a Kuhnian paradigm, the present study is a classic example of "normal science." The study articulates the Heath system by trying to increase its precision and reliability, by determining what facts are significant and by determining the scope of these facts (Kuhn, 1970). The author's interpretations carry the implicit assumption that the study did articulate the Heath paradigm and that the results had defined fruitful avenues for further research.

An alternative view is that the Heath paradigm did not fare well. This position moves us into "extraordinary science" and uncovers some interesting issues in science education research.

The statistically significant correlations between chemistry achievement and cognitive preference are admittedly weak (0.17 for PM and 0.19 for QM). In fact, these would account for only about 4% of the variance in test scores. In terms of teaching and motivating

students to achieve on final tests, the significance of other educational variables would seem to overshadow cognitive preference variables. As the analysis of the data moved into the hyperfine structures of test results, QM no longer played a significant role, but PM and PA exhibited weak associations. Other cognitive preferences, and other combinations of cognitive preferences showed no differences or showed inconsistent results. The "extraordinary science" issue emerges: How strongly, consistently, and thoroughly should a cognitive preference construct explain variation in test results in order for the construct to merit the acceptance of the science education community? On the criteria of strength, consistency, and thoroughness, the Heath construct may be reasonably viewed as problematic, thus encouraging one to reexamine the context in which Heath originally postulated the construct. Such a critical analysis/discussion might reveal that the four-category system was philosophically derived as an intriguing hypothesis for legitimate empirical investigation. The construct validity of the Heath construct does not appear to be strengthened by the results of the McNaught study, though this view requires a resolution of the "extraordinary science" issues mentioned above. This view does not question the quality of the study, it simply suggests that there is an additional topic that could have been explored by the author: a critical and explicit look at how well the Heath construct fared, including the author's criteria for making such a judgment.

The statistical significance of a result is clearly different from its educational significance. The former was rigorously investigated by the author, but the latter was simply assumed without a well argued and well documented case. Many science education reports lack rigorous treatment of the educational significance of findings. Many science educators believe that their responsibility as researchers ends at reporting the statistical significance of their results. Other science educators believe that their responsibilities include the improvement of science teaching through communicating the educational significance of their findings. The criteria for educational significance are not

nearly so well developed as the criteria for statistical significance. It is not surprising, then, that the author of the present report did not hazard an academic journey into somewhat uncharted waters. At the same time, however, it is not surprising that science teachers do not generally value our research findings.

REFERENCES

Heath, R. W. "Curriculum, Cognition and Educational Measurement." Educational and Psychological Measurement, 24: 239-253, 1964.

Kuhn, Thomas. The Structure of Scientific Revolutions (2nd ed.). Chicago: University of Chicago Press, 1970.

Reynolds, D. S. and R. D. Simpson. "Pilot Study Using Computer-Based Simulations on Human Transactions and Classroom Management." Science Education, 64 (1): 35-41, 1980.

Descriptors--Affective Behavior; *Classroom Environment; *Computer Assisted Instruction; Higher Education; Instructional Materials; *Methods Courses; Pilot Projects; *Preservice Teacher Education; *Science Education; Student Attitudes

Expanded abstract and analysis prepared especially for I.S.E. by F. Gerald Dillashaw, Bradley University.

Purpose

The purpose of this investigation was to study students' perceptions of the relative instructional effectiveness of small group discussions, role-playing activities, and computer-based simulations on student competencies in dealing with classroom transactions.

Rationale

Some educators have suggested that computer-based simulations can play a role in teacher education, adding another dimension to the training of teachers. Lunetta (1977) has prepared a set of computer-based simulations that present problems in classroom management, discipline, or student relations. These simulations present the student, in the role of teacher, with a situation that requires an action. The student's action leads to a follow-up situation which, again, requires action. This process continues until the situation is resolved. Resolution of the problems may range from a very good solution to a very poor one. Student perceptions of the relative instructional effectiveness of the Lunetta computer simulations, role-playing activities, and small group discussions were assessed in this study.

Research Design and Procedures

The design of the study was a non-equivalent control group design. The subjects were 32 students enrolled in an introductory science education course and in one of three laboratory sections for the course. In each of the laboratory sections, the three treatment groups -- discussion, role-playing, or computer simulation -- were randomly created. Ten students were in the discussion group, 12 in the role-playing group, and 10 in the computer simulation group. The contents for all groups were based on eight of the Lunetta simulations. Such problems as suspected cheating, stolen laboratory materials, and suspected drug use were among the classroom situations presented. For the discussion and role-playing groups, the subjects were presented only the initial scenario and suggested alternatives. In addition to this initial information, the computer simulation group also was presented with additional probable results according to the Lunetta models. The study lasted for two weeks.

The independent variable was method of instruction with three levels of treatment:

Discussion Treatment. In this treatment, the instructor presented the problem situations orally to the group. The possible alternatives were also presented orally. Each member of the group was asked for his/her course of action. Discussion of the possible solutions followed. Instruction was accomplished in a two-hour session each of the two weeks.

Role-Playing Treatment. Subjects were given a problem situation along with a list of possible courses of action. The situations were acted out with a subject acting as teacher. Other group members played appropriate characters in the scenario. The scenes progressed until the problem situation was resolved in some manner. The group then analyzed and discussed the solutions. These subjects met for a two-hour session each of the two weeks.

Computer-Based Simulations Treatment. This group used the computer-based Lunetta simulations. The subject received an introductory session on the operation of the computer terminals and the mechanics of

using the simulations and was told to complete the simulations at his/her convenience within the two-week instructional period.

The dependent variable was student-perceived competencies and opinions concerning the effectiveness of the instructional methods. The measure was an attitude survey using a 5-point Likert scale format to assess 11 statements on a scale from "strongly agree" (1) to "strongly disagree" (5). All statements on the survey were written so that a rating of one was the most positive (strongly agree) rating. The data were analyzed using analysis of variance (ANOVA) to test for mean differences on each question. An alpha level of .01 was used to test the hypothesis of no difference among the means.

Findings

No significant differences among the means of the three groups were found for any of the 11 questions on the attitude survey. The authors assumed that any mean greater than three indicated an unfavorable response. In no case did any mean, either in each of the three groups or overall, equal or exceed 3.00. Based on the mean scores on the items, the authors report: (1) all groups appear to be satisfied with their instructional method; (2) each group expressed a desire to experience the activities of at least one of the other groups; (3) the problem situations represent "real world" situations; (4) the activities better prepare them to deal with students; (5) perhaps more time should be spent on such activities.

Interpretations

The authors conclude that the lack of significant differences may suggest that students perceive each of the instructional methods as equally effective. However, they argue that the time investment by faculty to prepare for a discussion or role-playing strategy as compared to the computer simulation may mean some differences exist -- the computer

simulation strategy required considerably less preparation time by the instructor. The authors caution against using the time factor as an argument to replace other methods of instruction with computer simulations; rather, they suggest that such activities can be an effective supplement to the curriculum. Additionally, the authors argue that the findings reported above are not generalizable beyond the sample because of the sample size and research design, but indicate that these subjects appear to have increased confidence as a result of the activities. Accordingly, the Lunetta programs may have promise in science education curricula.

ABSTRACTOR'S ANALYSIS

This pilot study attempts to answer an important question -- How can teacher educators best engage preservice students in realistic problem situations that they are likely to encounter in the classroom? Discussion strategies and role-playing activities have been used for some time to help preservice teacher education students gain such experience. Computer-based simulations are a new way for students to use the capabilities of the computer to explore various aspects of problem situations of teaching. With increasing use of computers, the question of the effectiveness of computer simulations as compared to other methods is one that needs to be addressed. Students also need to perceive that their instruction is effective, that the instruction is preparing them to become good teachers. This study is an initial attempt to assess student perceptions of instructional effectiveness of these three methods.

Although the authors have done an excellent job in describing their procedures and treatment conditions, there are some potential problems with the study. The authors acknowledge that the sample was small; 32 students divided into three treatment conditions results in very small treatment group sizes. Additionally, the subjects were already enrolled in the laboratory sections. There was no information reported that indicated equivalence of the laboratory groups prior to the random assignment of subjects in the laboratory groups to treatment conditions.

The provision of all the items on the attitude survey provides the reader with more detailed information about the content of the instrument than one normally sees in a journal article. But, the lack of any information about the validity or reliability of the instrument is bothersome. Indeed, since no significant differences among the means were detected by the ANOVA, the lack of reliability data makes the task of explaining the findings even more difficult. If the instrument was not reliable enough, ANOVA procedures may not be able to detect differences that may actually exist. Instruments of this type suffer from lower reliability indices when compared to tests of cognitive achievement. This problem, plus the fact that the instrument has only 11 items, makes the reliability question an important one.

The authors conclude that the lack of significant differences may suggest that the three methods are essentially equally effective, but that some differences may actually exist when one considers the much less preparation time required by instructors when using computer simulations. Because of the problems noted, particularly the reliability one, such a conclusion is probably not warranted. There is still the unanswered question of whether differences may exist. A larger sample size and a reliable instrument would give more confidence in attempting to answer the question.

As a pilot study, this study did provide experience in conducting a study of this type. Based on this study, other researchers can attend to the potential problems and design studies that can help answer the important questions raised in this study.

REFERENCES

Lunetta, V. N. "Human Transactions and Classroom Management: A Computer Based Unit in Teacher Education." Educational Technology, 17 (4): 35-37, 1977.

Bethel, Lowell J. and Others. "The Effects of a NSF Institute on Inservice Teachers' Views of Science and Attitudes Toward Environmental Science Education." Science Education, 66 (4): 643-651, 1982.

Descriptors--Elementary School Teachers; Elementary Secondary Education; *Environmental Education; *Inservice Teacher Education; *Institutes Training Programs; Middle Schools; *Program Effectiveness; Program Evaluation; Science Course Improvement Projects; Science Education; *Science Programs; Secondary School Teachers; *Teacher Attitudes

Expanded abstract and analysis prepared especially for I.S.E. by Carol Wareing, Stoughton Public Schools, Massachusetts.

Purpose

The purpose of this study was to investigate the effects of an environmental science education program on inservice teachers' views of science and attitudes toward environmental science education. Specifically, over a period of one academic year, would a program that stresses inquiry activities, science seminars, lectures, demonstrations, special reading assignments, and field trips foster positive attitudes toward environmental science and science education?

Rationale

Although it is believed that science educators have a major responsibility to foster both the acquisition of knowledge and positive attitudes toward environmental education, it appears that there are few study reports available which concentrated on the affective dimension of environmental education.

Research findings indicate that students' attitudes toward science are affected by their teachers' attitudes toward science. Since there was no evidence of longitudinal studies in this area, the authors undertook such an investigation.

Research Design and Procedures

A modified non-randomized, equivalent control group, pretest post-test design was utilized.

Fifty-one inservice teachers were selected from the central Texas region. To establish a pool of applicants, announcements were sent out to school districts within a 75 mile radius of the University of Texas at Austin. From a pool of 231 respondents, a participant group of 51 was selected based on classroom teaching experience and science background.

The final participant sample included teachers of grades four, five, and six who had an average of 4.6 years and a range of 2.0 to 31.2 years of teaching experience. The average age of the participants was 27.5 years, and they averaged 5.1 semester hours of undergraduate science.

The control group was generated by pairing comparable teachers with the treatment group teachers. Each participant was requested to select a teaching colleague in the same school and grade. These additional teachers served as the control group.

The program used as a treatment in this study took place over a period of one academic year, and consisted of thirty two-and-one-half-hour class sessions. Presentations were followed by laboratory sessions and post laboratory discussions. The participants also attended two field trips conducted by a geologist and a botanist. A total of 16 hours was devoted to field work by the participants.

Both the treatment and control groups completed the Views of Science (VS) instrument and the Environmental Education Questionnaire (EEQ) during the beginning of the first semester, at the end of the first semester, and again at the end of the second semester. Alpha reliabilities for separate samples of preservice elementary teachers were found to be 0.71, 0.84, 0.80, and 0.76 for the VS instrument. An Alpha reliability of 0.89 was found for the EEQ instrument.

Comparisons between the pretest means of the control and treatment groups for each of the instruments were calculated to identify any significant differences between groups prior to, during, and following the treatment. An analysis of variance (ANOVA) was performed on the data using the (SPSS) Statistical Package for the Social Sciences.

Findings

The analysis of data for the VS and EEQ tested the differences between the group means at the beginning, middle, and end of the course.

There was no significant difference between the group means for either the VS or EEQ at the 0.01 level of significance.

Comparisons between the means of the two groups at the end of 16 class sessions (the midtest) yielded a significant difference between group means for both the VS and the EEQ. The explained variance, using Hays Omega Squared, was 0.05 for the VS and 0.28 for the EEQ.

In comparing the means of the two groups at the conclusion of the 32 class sessions, significant differences were found between group means for the EEQ, but not for the VS. The explain variance was 0.29 for the EEQ.

Interpretations

Prior to the treatment, the control and experimental groups demonstrated similar views of science and attitudes as measured by the EEQ and the VS.

There was a significant difference between the control and experimental groups at the mid-point of the study on views of science. Those participating in the treatment program viewed science as more "tentative" than at the beginning of the study. It can be concluded that, after 16 weeks, the treatment was effective in bringing about changes in views of science and attitudes toward environmental education.

Given the difference between the groups on attitudes towards environmental science as measured by the EEQ, the 32-week treatment was deemed to be very effective in developing positive attitudes towards environmental science education.

Analysis of the data from the VS instrument revealed no significant differences at the conclusion of 32 weeks. This result was surprising to the investigators.

There was evidenced a pattern of changes in means over time for the two groups on both the VS and the EEQ. For the treatment group, there existed a steady increase in philosophical views of science and attitude toward environmental science education from the beginning of the study to its conclusion. The pattern of attitudes toward environmental science education for the control group showed little change from pretest to post-test.

The participation of the control group in two inservice workshops with members of the experimental group may have had a causative factor relative to the positive increase demonstrated by the control group during the second half of the program.

The results of the study indicate that inservice programs of a year's duration can have a positive effect on teachers' attitudes.

ABSTRACTOR'S ANALYSIS

In their review of the literature, Bethel, Ellis, and Barufaldi found that preservice and inservice teachers generally indicated negative, or at best, neutral attitudes toward science. Many teachers judged themselves unqualified to teach science. Since the research likewise indicates that students' attitudes toward science are affected by their teachers' attitudes, Bethel et al focused on the perceived need to improve teachers' attitudes toward science by upgrading their skills in the area of environmental science.

This study contributed favorably to the existing body of knowledge on the subject since the inservice treatment was successful in effecting positive changes in attitudes toward science.

Although the experimental group demonstrated growth in positive attitudes as measured by the midtest and final test, there was evidence of a pattern of change in the control group's attitudes as well. Participation of the control group in two of the inservice sessions was suggested by the authors as a possible causitive factor. Such a suggestion might warrant a repeat of the study with a control group totally absented from any treatment effect. Groups differentiated according to full, partial, and no treatment would lend further strength to the position that inservice programs of a year's duration can have a positive effect on teacher's attitudes.

A rather unconventional methodology was used in selecting the control group. Control group pairing to the treatment group was accomplished through colleague selection. Generalizations from such an investigation must consider the modified non-randomized, equivalent group with the pretest post-test design used.

In reporting on the study, attitudes at the outset were not clearly outlined in terms of their relative strengths, or weaknesses. Likewise, the suggestion that science was viewed as "tentative" was neither explained nor valued.

Given the findings from this investigation, future research efforts might consider the following: Do positive attitudes toward environmental science result in transference to environmental consciousness and behavior changes? What are the long-term effects of a given inservice, and are periodic updates or review workshops suggested for reinforcement? What are some of the specific antecedents of attitudes toward science, and are there further implications for inservice education? What specific aspects of the inservice program had the greatest effects?

Smith, Walter S. "Career Education Attitudes and Practices of K-12 Science Educators." Journal of Research in Science Teaching, 19 (5): 367-375, 1982.

Descriptors--*Career Education; *Elementary School Science; Elementary Secondary Education; Science Careers; Science Education; Science Instruction; *Science Teachers; *Secondary School Science; *Surveys; Teacher Characteristics; *Teacher Attitudes

Expanded abstract and analysis prepared especially for I.S.E. by Gerald H. Krockover, Purdue University.

Purpose

The purpose of this study was to survey K-12 science educators regarding their attitude toward and practice of career education in their science teaching.

Rationale

In 1978-79 the United States Office of Education engaged seven professional associations, the National Science Teachers Association being one of these, to assess teacher attitudes and to improve career education instruction. Based upon the survey the answers to the following questions were sought:

1. What importance (i.e., value) is career education as part of a student's education?
2. Should career education be incorporated into regular science courses or delivered separately?
3. Is career education perceived as narrowly vocational or broadly inclusive of a range of goals?
4. Are K-12 science educators willing to attempt to include career education concepts in existing science courses?
5. How easy do K-12 science educators perceive to be the inclusion of career education in existing science courses?

6. What problems do K-12 science educators encounter when attempting to incorporate career education in their science courses?
7. How frequently do K-12 science educators incorporate career education in their science courses?
8. Among K-12 science educators, what is the relationship of willingness to incorporate career education, perceived ease of incorporation, and frequency of incorporating career education in science courses?
9. Do subgroups of K-12 science educators, based on teaching experience, level of teaching, size of community, sex, age, or highest degree, differ in attitudes and practices represented by the preceding questions?
10. What information do K-12 science educators feel they need in order to improve their career education instruction?
11. From what sources has career education information come to these K-12 science educators and from where do they believe the information ought to come?

For purposes of this study, vocational education was subsumed under the broader term, career education, which is considered to be an educational program that focuses on individual career development from childhood through adulthood.

Research Design and Procedures

Representatives of the seven professional education associations met to outline the content of and suggest items for a common survey. A private consultant developed the survey which was reviewed and refined by the seven representatives. No other data were collected to establish reliability or validity. Demographic data were also collected. Four hundred randomly selected NSTA members were sent surveys and 190 (47.3%) responded to the first mailing. In response to two subsequent requests, 46 (11.5%) and 63 (15.8%) surveys,

respectively, were returned for an overall return rate of usable surveys of 74.8% (299 of 400). The results of the three sets of returns were compared for six key variables. No significant differences were found among the responding groups for any of the variables. Because the three groups did not differ significantly, their responses were pooled together for further analyses and were considered representative of K-12 science educators who were members of NSTA.

Various attitudes toward career education were assessed in the group of 29 items. For each statement, respondents indicated their definite agreement (scored as 4), agreement (3), no opinion (2), disagreement (1), or definite disagreement (0). The data collected were then factor analyzed.

Findings

The first factor dealt with how strongly educators felt career education was important in a student's overall education. K-12 science educators expressed a slightly positive attitude toward the importance of career education. Factor 2 dealt with whether career education ought to be taught as a separate course or included in existing courses. The science educators responding felt that career education ought to be taught in existing courses. The third factor dealt with whether the goals of career education ought to be limited to vocational education or be more inclusive. The K-12 science educators strongly affirmed the inclusive view of the purposes of career education.

Subgroups were also examined with regard to the three factors including: grade level taught, years experience in present position, age, sex, highest degree, and size of community. No significant differences at the $p < 0.05$ level were found among subgroups based upon size of community, sex, years of experience in present position, or in their attitude toward the inclusiveness of career education.

Science educators who taught at different grade levels, differed in the first factor ($p < 0.05$). Using the Scheffé method, it was found that junior high science educators valued the importance of career education more strongly than did senior high science educators. Teachers aged 31 to 55 more strongly favored incorporation of career education into existing courses. Teachers possessing a specialist or doctorate degree favored more strongly incorporation of career education into existing courses.

Two problems impeding the incorporation of career education into existing science courses were too little time (52.7%) and lack of knowledge of instructional materials (47.8%). The primary source of career information for K-12 science educators was from outside their school districts (professional publications, 41.7%).

ABSTRACTOR'S ANALYSIS

This study made an attempt to ascertain the attitude of 400 K-12 science teachers who were members of NSTA regarding career education and its role in the science education curriculum. Many questions can be asked regarding this study and its "results." Is this study a representative sample of K-12 science teachers across the United States? For secondary teachers, maybe; for elementary teachers, probably not. Is this study a representative demographic representation of K-12 science teachers in the United States? For white males who teach in secondary school, probably yes; for minorities, females and elementary school teachers, probably not.

What about the survey used? Is it either reliable or valid? Seven "representatives" got together and suggested items and these were put together by a "private" consultant into a survey. Would any researcher at a major educational institution approve a graduate student's research using this format? Would George Gallup want to use this poll? Probably not.

It is also interesting to note that there s no comprehensive bibliography included with this study. As a result, it is impossible for someone with an interest in pursuing one of the subtopics to readily locate any references. It is unfortunate that references were not included since the author probably has a wealth of information that would be of significant use to the readers of this study.

Finally, one has to be concerned over the lack of return of the first mailing of the survey (190/400, 47.3%). Two subsequent mailings were used to add another 109 returns. Are these 109 additional "late" responses as valid as the 190? Are the first 190 those who had career education concerns?

When an organization such as NSTA agrees to participate in a survey, it has the responsibility to make sure that it is a valid and reliable one. Could there be an underlying reason why the National Council of Teachers of Mathematics is not one of the participants listed?

Career education is too im portant to allow it to flounder due to questionable survey methods and techniques. It deserves better treatment. Maybe it will receive it in future surveys.

Richmond, James M. and Neil Baumgart. "A Hierarchial Analysis of Environmental Attitudes." Journal of Environmental Education, 13 (1): 31-37, 1981.

Descriptors--Attitudes; *Environment; *Environmental Education; Foreign Countries; *Measurement; *National Surveys; *Secondary Education; *Student Attitudes

Expanded abstract and analysis prepared especially for I.S.E. by Charles L. Price, Indiana State University-Evansville.

Purpose and Rationale

A national survey was conducted in England in 1976 (Richmond and Morgan, 1977) to assess the environmental knowledge and attitudes of fifth-year secondary students. The survey provided descriptive information on student responses and examined relationships between outcome measures and various demographic variables including gender, school type, and school size.

In this 1981 study, a portion of the 1976 data was re-analyzed. In order to determine if attitudes toward environmental education might reveal a hierarchial nature, an examination of the responses to items regarding beliefs was undertaken.

A hierarchial belief structure had not been postulated in the original study. The present analysis therefore represented an empirical effort rather than theory-based research. Nonetheless, the concept of a hierarchial belief system has appeared in environmental education research. For attitudes in particular, it has been postulated that some concerns about the environment are almost universally held, while others are held only by persons with a particular commitment to environmental issues. The validation of an environmental belief heirarchy would have important implications for the sequencing of instructional material.

Research Design and Procedure

The instrument developed for the survey (Richmond and Morgan, 1977) was comprised of three forms (A, B, and C) each, with three parts (1, 2, and 3). The purpose for having three forms of the instrument was to maximize content coverage without making unreasonable demands on respondents. Hence the three forms were not designed as parallel versions. Parts 1 and 2 of each form dealt with factual and conceptual knowledge. Part 3 contained 15 statements of belief. This study dealt solely with analysis of responses to belief items.

Each belief item appeared as a statement with the following response options: a. agree; b. disagree; and c. no opinion. The "environmentally positive" response to each statement was determined by a panel of environmental experts, with unanimous agreement necessary for the inclusion of an item in the instrument. Both positive and negative statements were used and the responses were scored dichotomously. Responses which indicated concern for the environment (that is, agreement with a positive statement or disagreement with a negative one) were scored as (1) and other responses as (0). The facility index of an item was defined as the percentage of respondents agreeing with the experts.

The target population was all fifth-year students enrolled in the secondary schools of England in 1976. The sample was drawn to insure proportional representation of the major types of schools. A total of 500 schools was selected and the instrument was administered to a sample of about 30 students in the fifth year at each school. Schools were contacted by mail, and 383 schools (76.6 percent) returned completed answer sheets from 11,008 students.

Each of the three forms (A, B, and C) contained 15 belief items. Four items were common to all forms while other items were distributed over other areas of environmental concern. Because the forms were not parallel, the responses to each were analyzed separately for possible hierarchial structures. Since the sample was large, it was decided to divide the respondents to each form randomly into four groups and to treat these as replications for analysis purposes. Each group had at least 900 subjects.

Hierarchial structure determination was made with the use of a contingency table for two items. For each form of the instrument, relationships between all pairs of items were examined. Of particular interest were situations in which respondents (1) did not support the environmentally positive (preferred) response to item a and did support the preferred response to item b and (2) supported the preferred response to item a and did not support the preferred statement for item b. If condition 2 was $\gg$ than condition 1, it could be argued that item a preceded item b.

Implementation of the procedure was accomplished by converting the phi-coefficient for each cell to a z-transformation. The following rules were applied to locate the structures described in the results:

1. The order of items was determined by the facility index (proportion of Ss correct).
2. Items linked in the hierarchy were associated at a 0.01 probability level.
3. Only linkages which emerged in three of the four replications were included.

Findings

Form A: Of the fifteen items in Form A, eleven met criteria to be included into hierarchial structures. Two major branches of hierarchy were determined: (a) one based largely on population issues, (b) the other based on concerns of pollution. The beliefs most commonly held were those about the desirability of maintaining environmental quality and excercising environmental safeguards. Above these were items concerning sanctions, controls and incentives; the highest hierarchial items were those involving conflicts between environmental safeguards and alternative benefits (lower costs, energy supplies, individual freedoms).

Form B: Nine of fifteen belief items scaled within a hierarchy. The item at the bottom of the structure was one of restoration of the environment. Those which followed related to beliefs about the desirability of conservation of animals, plants and land. The higher level implied control of enforcement, and finally, conflict situations.

Form C: There were nine items in the hierarchy, organized into two branches. The item at the bottom of the hierarchy was one involving control through regulation. In one branch were items related to alternatives to traditional sources or uses of energy. The other branch had a mix of items mostly related to energy use and its relationship to the environment. In both branches, items at the top of the hierarchy involved situations in which environmental concern was in conflict with other concerns of personal freedom, or economic or social benefits.

Interpretations

In general, items expressing opinions on the desirability of preserving the natural environment, on repairing damage, on avoiding pollution, and so on occurred lowest in the hierarchies. Items which followed advocated control of the use of the environment through incentives or through regulations appeared next. At the top of the hierarchies were items suggesting a conflict between preservation of the environment and other potential benefits of economic, social, and personal nature.

The data re-analysis suggested that students were unlikely to react positively to items higher in the hierarchy unless they also reacted positively to lower order items. Students would appear unlikely to favor controls over the use of the environment unless they initially expressed concern about environmental quality. It would follow, therefore, that environmental education curricula should consider the nature of hierarchial belief structures.

ABSTRACTOR'S ANALYSIS

This research is a re-analysis of data gathered at a previous time. Some of the comments in this review will be directed at issues in the original study (Richmond and Morgan, 1977), while other remarks will be addressing the 1981 study.

Research Design

The purpose of this study was stated as an attempt to identify possible hierarchial structures in belief statements regarding environmental concerns. Three such structures were derived from analyses of data. The hierarchial analysis, "was not considered at the time the instrument was constructed" (Richmond and Baumgart, 1981, p. 32). For this reason, the study could not be considered a validation of previous theory, but rather an attempt to uncover an empirical ordering of items.

The search was warranted "given the large sample of subjects and the care that was exercised in sampling" (ibid.). The fact that over 11,000 forms were completed by Ss selected using stratified random sampling techniques is significant. Some educational researchers have pointed out (Isaac and Michael, 1977, p. 69), that in a descriptive study of this nature the correct unit of analysis should be the classroom, rather than the individual. If this were the case, the Ss would be 500.

As previously stated, the research instrument(s) were not developed a priori with hierarchies of belief structures. Responses to belief items were gathered with three different instruments not designed to be parallel in nature. There were, however, four common items on all three forms. Of the four items, only one met the criteria to appear in all three of the belief hierarchies. This can be attributed to the fact that each of the forms had different items representing various beliefs. With these considerations, it is difficult to make generalizations as to definite belief hairarchies.

For these reasons, experimental research is needed. Belief items based on the hierarchies formulated from this study could be developed. Since the data for this research were gathered in Great Britain, a comparison of beliefs of English students regarding environmental issues with those of students in the United States and in other countries is suggested.* In addition, longitudinal studies of beliefs regarding environmental issues would appear to be promising. Do the beliefs and values of students with regard to the environment remain relatively stable or is there preference for the most immediate and publicized environmental concerns?

These issues should be of vital concern of those developing environmental education curricula. Richmond and Baumgart have provided a good foundation on which to base additional research.

REFERENCES

Isaac, Stephen and William B. Michael. Handbook in Research and Evaluation, San Diego, CA: Edits Publishing, 1977.

Richmond, James M. and Richard F. Morgan. A National Survey of the Environmental Knowledge and Attitudes of Fifth Year Pupils in England. Columbus, OH: ERIC/SMEAC Information Reference Center, 1977.

*Perkes, A. Cordell. "A Survey of Environmental Knowledge and Attitudes of Tenth and Twelfth Grade Students from Five Great Lakes and Six Far Western States." Unpublished Ph.D. dissertation, The Ohio State University, 1973.

Bohl, Walter Benson. "A Survey of Cognitive and Affective Components of Selected Environmentally Related Attitudes of Tenth and Twelfth Grade Students in Six Midwestern, Four Southwestern and Twelve Plains and Mountain States." Unpublished Ph.D. dissertation, The Ohio State University, 1976.

TEACHER EDUCATION

Ritz, William C., J. G. Cashell, M. F. Felsen. "The Effectiveness of Science Supervisors and Their Membership Status Within a Faculty." Journal of Research in Science Teaching, 18 (3): 229-40, 1981.

Descriptors--Evaluation; *Job Performance; *Personnel Evaluation; Questionnaires; Science Education; *Science Supervision; *Science Teachers; *State Supervisors; Supervisor Qualifications; Supervisory Methods

Expanded abstract and analysis prepared especially for I.S.E. by William Capie, University of Georgia.

Purpose

Two purposes were cited for the study:

1. To examine "whether or not there are differences in perceptions of supervisory effectiveness between science supervisors and the science teachers with whom they work," and
2. To examine "one variable which may moderate the relationship between teachers and supervisors and which affects perceptions of the effectiveness of supervisors."

"Science supervisors" was broadly conceived to include department heads, specialists and consultants as well as individuals with the job title.

Rationale

The study explored both formal and nonformal aspects of the supervisor's rule. The principal rationale for the work lies in Jackson's model of group membership (1959). Jackson described an individual's relationship with a group in terms of two independent dimensions--the individual's attraction to the group and the group's acceptance of the individual. An extreme of high attraction and high acceptance is characterized by "psychological membership" while a combination of low acceptance and low attraction is characterized by a "war relationship." The assumption for the study is that group membership will enhance the effectiveness of supervisors.

Research Design and Procedure

The study involved the analysis of two surveys administered to 401 educators in New York. Responses were received from 143 of 550 "supervisors" and for 258 of 760 teachers to whom they were provided.

The Science Supervision Rating Scale (SSRS) consists of 26 activities which respondents rated on a scale ranging from "This is among the activities I perform least effectively" to "This is among the activities I perform most effectively." An orthogonal factor analysis procedure was used to create four subscales. The subscales are Instructional Intervening defined by 3 items; Interpersonal/ Supporting defined by 6 items; Management/Planning defined by 6 items; and Socializing defined by 2 items. Thus, nine items apparently failed to load on any factor. Subscale reliabilities (coefficient alpha) ranged from 0.55 to 0.80 with the total yielding a coefficient of 0.91.

The Person-Group Relationship Scale (PGRS) was used in the study (Apparently a specialized form for supervisors was used, but the report is not clear in this regard.) After supervisors and teachers responded to the instrument, factor analysis was used to construct two subscales dubbed "attraction" and "acceptance." Coefficient alpha for both six item scales exceeded 0.9.

ANOVA procedures were used to compare the responses of teachers and "supervisors" on the four SSRS scales and the two PGRS scales. A canonical correlation analysis between the two instruments was also completed.

Findings

Significant differences ($p < 0.05$) between the perceptions of the two groups were found in the following four areas: Instructional Intervening, Interpersonal Supporting, Socializing, and Group Membership Acceptance. Seven SSRS items were ranked differently by the two groups; interpersonal items were ranked more highly by teachers while supervisors ranked areas such as "observes" and "conducts demonstrations" more highly.

The maximum two canonical variables were identified between the sets of variables. The stronger (R=0.62) was characterized by the acceptance and supporting dimensions, although the attraction and planning scales were also moderately correlated with the variable. The second canonical variable which was characterized by attraction and socializing was not strong (R=.17).

Interpretations

The authors claim that the results "support the commonly held view that teachers and their supervisors hold quite different perceptions of how effective the supervisor is." They explore a number of possible reasons for the differences including insensitivity of some supervisors about the importance of the interpersonal dimension, the absence of expectations and rewards in this area, and limited training in the area.

ABSTRACTOR'S ANALYSIS

This area of research certainly is important and quite pertinent today with increasing interest in the quality of schools and science teaching. Concern with the formal and informal dimensions of supervision/administration is well-founded with a history at least into the 1950's when the Ohio State Leadership Studies identified the areas of personal characteristics and task-orientation as important for administrators. This study attempts to deal with the issue of effectiveness by exploring the perceptions of supervisor effectiveness held by supervisors and those for whom the supervisors are responsible. Unfortunately, the report has several critical omissions that make an assessment of the findings and their interpretation quite difficult.

The authors acknowledge that they are dealing with perceptions which may well be flawed. In fact, self perception measures of school administrators have been shown to be much higher than similar measures by school faculties. A more serious problem is the constitution of the group of supervisors. The report indicates that the group is diverse, but it does not suggest the frequencies of the various job titles within the group. The readers can suspect that there are far more department heads than central office personnel. However, even with this insight, the role and responsibility of middle school and high school department heads is quite different. Also, there is no indication of how many teachers are associated with each supervisor. There may be some rationale for using the mean of teacher responses for each school or school system as the data to be compared to a particular administrator. These difficulties are compounded by the low response rates reported in the study.

Beyond the question of "'Who' is describing 'Whom'?" is "What are they saying?" The report contains no means for any of the response scales. This omission is serious for a number of reasons. First, a basic concern in the study is "How effective (on a five-point scale) are supervisors?" Second, small differences in the mean scores could change the relative ranking of items which are reported and, although the items were ranked differently by the two groups, the importance of these differences is difficult to assess. Thirdly, the sample is sufficiently large that very small differences in item or scale scores could be statistically different without being important. With no notion of whether grand means were 2.4 or 4.2 with standard deviations of 0.2 or 2, the value of the comparisons is limited.

The use of factor analysis to identify items for the subscales appears to have had several consequences. One consequence is reducing the items to a more manageable set of constructs, or supervisory functions. While some may quarrel with the inclusion of any two item factors such as socializing, the supervisory functions do have a logical face validity. A second consequence is that

several items did not fit into any subscale. Thus, "evaluation of teaching," "observation of classroom teaching" and "consulting with administrators" were not included in any analysis. Surely, these are supervisory practices worthy of some inspection. Their importance may have been lost in a sample heavily weighed by department heads, however. The use of factor scoring coefficients to develop the scale scores would have included all items in the analysis although it would not have enhanced the importance of these low loading items in the ANOVAs.

The use of the factor analysis derived scale scores in the canonical correlation procedures may have been somewhat redundant. The items of both scales could have been entered as sets of variables and the canonical coefficients would then have indicated the relative contribution of each item to the resulting canonical variables. Although the many variables may have made interpretation more difficult, the additional information provided by the analysis would have made interpretation much richer.

The authors are to be commended for pursuing the area of supervisor effectiveness. "Educator effectiveness" is "in" for the nineteen-eighties. With teacher effectiveness research there is a convenient criterion variable in pupil achievement to use in studies. With administrator effectiveness there are also plausible criterion variables such as ADA, school achievement, teacher absences, and so on. Unfortunately, with supervisors products/outcomes to be related to job performance are rare. Consequently, other methodologies must be considered for identifying and understanding the aspects of a supervisor's job performance which are related to effectiveness. For example, identifying high performers and low performers through nominations or other similar procedures may provide samples of supervisors to analyze and contrast for performance differences. However, self-reports are likely to contribute little more than helping to establish a need for supervisors' professional improvement.

REFERENCE

Jackson, J. M. "A Space for Conceptualizing Person-Group Relations." Human Relations, 12: 3-15, 1959.

Van Koevering, T. E. and N. J. Sell. "An Analysis of the Effectiveness of Energy Education Workshops for Teachers." Science Education, 67 (2): 151-158, 1983.

Descriptors--Course Content; Elementary Education; *Energy; Environmental Education; *Inservice Teacher Education; Junior High Schools; Knowledge Level; Learning Activities; *Program Content; Questionnaires; Science Education; *Teacher Workshops

Expanded abstract and analysis prepared especially for I.S.E. by David Stevenson, Truro, Nova Scotia.

Purpose

The authors prepared and presented workshops on energy education with the goal of increasing both information and activity on the subject of energy within classrooms of elementary and junior high schools. They gathered data about the increase in teacher knowledge and interest as part of their program.

Rationale

During June of 1980, a total of 135 elementary and junior high school teachers took part in 15 hours of lectures, discussion, audio-visual presentations, and hands-on activities during two- or three-day workshops. The sessions were free to the teachers (due to federal funding). One graduate credit was given for workshop attendance. The teachers received an energy activity booklet said to be appropriate to the grade level of the recepient. A follow-up assignment was due in one month.

No reference is made to energy education literature or to other research.

The authors make no statements of assumption or hypothesis. They state, however, that four questions are addressed in the data analysis:

1. What is the teacher's perception of their (sic) familiarity with each topic prior to the workshop?
2. Which topics (of those familiar to the teacher) had previously been incorporated into the classroom activities?
3. Which topics did the teachers identify as ones where they had made significant gains in knowledge due to the workshop?
4. Which new topics would likely be introduced into the classroom due to the workshop?

Research Design and Procedures

The authors report their work as a pre-post design study with the energy workshops as an intervening treatment. A questionnaire of 45 idea statements was administered to the 135 subject teachers before and after the workshops.

Two responses were made at each administration: (1) an indication of level of knowledge about each statement topic on a Likert scale and (2) an indication of degree of classroom use (or potential use) of the statement topic on an A to C (none to frequent) scale. In addition, subjects were asked as part of the post-workshop questionnaire administration to list the 10 idea statements which they considered "winners," and the 10 "losers."

For each of the knowledge responses to the 45 statements arithmetic means and standard deviations were calculated for the (1) pre- and (2) postworkshop administrations, and (3) for the difference between the two responses. The statistical significance of each difference was calculated using a t test.

The authors also made interval categories of pre- and postworkshop means in order to investigate the perceived knowledge of teachers. The five categories were labeled A to E for each categorization. A to E categories of different interval size were used for the difference between the pre- and post means.

An additional categorization was used to consider the effect of the workshops on teacher willingness to use the topics within the curriculum. The sums of B and C (i.e., positive) responses were found for subjects for each item in the two separate administrations, and were formed into five (A to E) categories. Those statements which gained from a preworkshop rating of D or E to a post-rating of B or A were regarded as "winners;" "losers" were those which retained a D or E rating.

Findings

The authors present the findings on knowledge about the topics in a table with means, standard deviations, and ratings listed for each of the 45 statements as found before the workshop and after. The differences between the means are given also, followed by the standard deviations, t-test values, and ratings.

A second table presents the findings for use (or potential use) of the topics in the classroom, with columns for the pre- and postworkshop results on summing B and C responses, and the subsequent category ratings. The table also gives the "winners" and "losers" as rated by teachers, as well as through the authors' analysis of the categories.

Within the written part of the report, the authors state that a statistically significant gain at the 0.001 level of confidence was found, for each of the 45 idea statements, between the responses for pre- and post questionnaire administrations.

Interpretations

The authors concluded that workshop participants, prior to workshops, were most knowledgeable about the topics often given media attention and usually found popular with the general public. These

were the topics teachers discussed with students. They state that it is reasonable to assume an increase in knowledge will result in an increase in the number of topics which teachers would discuss with students.

The "Winners by Analysis" are topics which authors expect teachers to include as a result of the workshops. "Losers by Analysis," on the other hand, are considered by the authors to be excluded by teachers from instruction even though knowledge about the topics increased significantly.

The authors state that it can be inferred that well-taught workshops can contribute both to knowledge and to the variety of topics taught. But, such workshops are not likely to result in the use of all information learned. They feel attention to political and technical topics should be avoided, while more should be given to electrical energy, and to energy use in other countries.

ABSTRACTOR'S ANALYSIS

The topic of energy education emerged from the "crisis" period of the mid-1970's and has continued an existance as part of scientific and social issues, and yet somewhat apart from the core of science topics in need of attention.

Here one may expect to find a natural topic through which to explore the intersection of science, technology, and social issues. Indeed, the concepts which could be explored lend credence to the integration of several school subjects.

Van Koevering and Sell have reported on one of many responses to the need for increased training for teachers to encourage wider and deeper discussion and exploration within the newer and slightly peripheral science topics. Most of the responses are likely to go unreported in the science education literature, for the ability and willingness of science educators to conduct studies as part of the projects likely are limited.

In this case the authors were fortunate to find a number of teachers willing to improve their knowledge and to express their interests in the use of the knowledge within the classroom. One may wonder if the incentives, evenly spread across the continent, could be used to invite similar levels of interest in educational topics generally. It is normal to find a limited number of reports about situations in which teachers are encouraged to receive 15 hours of instruction after which positive results are expected.

Perhaps a closer look is in order to gain a measure of the findings which flow from the work of Van Koevering and Sell.

We are given 135 teachers who participated in two or three day workshops reasonably close to their homes or schools. The time(s) of year is not specified. The sessions are varied in the use of instructional techniques, and tangible rewards are available for participants.

The origin of the questionnaire is not reported. The preworkshop responses indicate a fairly homogeneous group of subjects in terms of knowledge of the topics, and the postworkshop responses suggest, in general, little deviation from that pattern. If the questionnaire has captured the perception by teachers of their knowledge one may feel confident that the subjects, as a group, are returning to teaching assignments better prepared for their teaching roles.

The reader is left, however, with a doubt about the format of the 45 statements to which responses were made. For example, "Electromagnetic radiation" may be the statement, or may be a summary of it. It would be helpful to readers to know if the small mean, both pre-and postworkshop, results from lack of knowledge about the topic or about the statement used to ask about the topic.

In spite of this limitation, it is clear that answers were found for the questions which were posed under the article heading of "Data Analysis." The second question requires some comment and a question. Does a positive response to the statement "have frequently discussed (or will frequently discuss) this topic with my students" mean the teacher had a choice of actual as compared to wishful, or is the format

of the statement (as taken from the report) a shorthand way of saying there are two forms of the statement - one pre-and the other postworkshop? A misunderstanding in this point could influence thinking about responses to the questionnaire.

The investigation of the likely increase of use of a topic after the workshops, and the rating of "winner" and "loser" topics presents a study in contradictions. The teacher selections of successful topics did not match the responses when the categories were calculated. This may not seem unusual to readers, and may indicate the chance selection of program should teachers indicate their choices of topics. On the other hand, the selection of five "winning" topics by analysis seems to have produced some surprises for "the law of thermodynamics" and "electrical power pricing patterns" were among them.

Van Koevering and Sell have reported on successful workshops designed to improve teacher knowledge about energy. The results of the workshops would seem to be positive in the short term. Beyond that the analysis has nothing to say. Will the subjects continue their positive attitudes over a period of one year? Will the program be strengthened in schools? To those science educators wondering if the task is worth the effort, we hold this report to be positive. However, a caveat is in order due to limited reliance on questionnaire results. One may wish, as well, for some knowledge of school system support for the topic being reported.