DOCUMENT RESUME

ED 123 135 SE 020 964

AUTHOR	Helgeson, Stanley L., Ed.; Blosser, Patricia E., Ed.
TITLE	Investigations in Science Education, Vol. 1, No. 4. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION	Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	75
NOTE	70p.
AVAILABLE FROM	Ohio State University, Center for Science and Mathematics Education, 244 Arps Hall, Columbus, Ohio 43210 (Subscription $6.00, $1.75 single copy)
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	*Abstracts; Cognitive Development; *Educational Research; Instruction; Research; Researchers; *Research Methodology; *Research Skills; *Science Education; Teacher Education

ABSTRACT

This fourth issue of INVESTIGATIONS IN SCIENCE EDUCATION (ISE), designed to provide a new perspective for viewing research articles and to aid in the improvement of writing research reports, includes abstracts prepared by science educators, bibliographical data, the research design and procedure used by the investigators including data, purpose, and rationale of the research, as well as detailed notes offered by the abstractor. Articles in the ISE are selected primarily from such sources as professional journals and reports of government-funded projects. With this issue, ISE inaugurates the so-called Cluster Reviews. These are reviews grouped together according to common concern. Four clusters of studies analyzed in this issue are: Instruction, Student Attitudes, Teacher Education, and Cognitive Development. (Author/EB)

Vol. 1, No. 4 1974-75

INVESTIGATIONS IN SCIENCE EDUCATION

Expanded Abstracts and Critical Analyses of Recent Research

National Association for Research in Science Teaching

ERIC Science, Mathematics, and Environmental Education Information Analysis Center

The Center for Science and Mathematics Education
The Ohio State University

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC Center for Science, Mathematics, and
Environmental Education

Published Quarterly by

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

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

Vol. 1, No. 4, 1974-75

NOTES . . .

from the Editor

With this issue, Investigations in Science Education inaugurates the "Cluster Reviews" announced in the last issue. These are reviews grouped together according to common concerns. Four clusters of studies are analyzed in this issue: Instruction, Student Attitudes, Teacher Education, and Cognitive Development. The analysts have attempted, where appropriate, to consider related research in their critiques of the reported studies. By presenting such clusters, it is intended that a broader, more comprehensive view of the current state of research in science education can be presented.

Comments and suggestions for improving Investigations in Science Education are invited and publishable letters of response to the abstracts and analyses are encouraged.

Stanley L. Helgeson
Editor

Patricia E. Blosser
Associate Editor

6

INSTRUCTION

EJ 027 044

Gallagher, James Joseph, "A Comparison of Individualized and Group Instruction in Science; Effects on Third Grade Pupils." Journal of Research in Science Teaching, Vol. 7, No. 3:253-263, 1970.
Descriptors--Academic Achievement, Audiovisual Instruction, *Elementary School Science, *Evaluation Techniques, Grade 3, *Individualized Instruction, *Instruction, Science Activities

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Julian R. Brandou, Michigan State University.

Purpose

The purpose of the study was to compare the achievement and performance during instruction of groups of third grade students in two modes of instruction and a combination of the two. The two modes were audio-tutorial (A-T) and teacher-directed (T-D).

Hypothesis 1: (deals with performance)

A-T will be distracted less than T-D.

Hypothesis 2: (deals with achievement)

A-T will be greater than T-D as measured by:

a. definition of "interaction"
b. generalization across examples
c. identification of exemplars/non-exemplars
d. description by identifying interacting dyads

Rationale

The author notes that most contemporary elementary school science programs depend on pupil exploration or independent investigation of materials and phenomena at the outset of instruction. This follows from the assumption that these "common, first-hand experiences provide sensory input that is essential for meaningful learning."

Further, he suspects that the typical classroom setting may provide too much distraction for the pupil to effectively receive the necessary inputs. Previous research by Novak, Bridgham, Hill and Siemankowski is cited on audio-tutorial instruction.

Research Design and Procedure

The basic design is a posttest only applied to four randomly assigned treatment groups.

Sixty third grade pupils from a public school in a middle class suburban community were randomly assigned to four groups. Group one was given a two lesson audio-tutorial sequence; group two received the same lessons, teacher-directed; group three was given the first lesson via A-T and the second, teacher-directed; and group four received no instruction on the topic. Lesson one was an "exploration" from the Science Curriculum Improvement Study, Interaction unit, and lesson two was the subsequent "invention" lesson. The Iowa Test of Basic Skill was used to show that groups were similar and the school program already involved some individualized work in reading and mathematics. The author assumed the A-T was, thus, not a novelty.

On Hypothesis 1: Videotapes of the pupils were analyzed to determine "overt distraction." Five characteristic activities are discussed and the mean percentage of time, to the nearest tenth of a percent, is given.

On Hypothesis 2: A special test was developed consisting of items such as the following:

1. Tell what interaction is.

2. How do you know when interaction takes place?

3. Pictures of exemplars; (four items) and non-exemplars (one item) and the student was asked, "Does the picture show interaction?"

Response categories were created, and the number of responses in each category for each item was reported. Since the posttest was delayed to one week after the instruction and only two lessons were involved, the duration of the study was probably two weeks at most.

Summary data were presented for each item and for the four exemplar pictures as a group. These items were scored by total number of correct responses since the pupil could frequently describe more than one dyad of interaction for each picture.

Findings

1. Students taught by A-T were less distracted than those taught by the teacher-directed method.

2. Students taught by the combination of lessons did better on the definition of Interaction.

3. Students taught by A-T did better on generalization across several examples.

4. Students taught by either method were able to discriminate between exemplars and non-exemplars, irrespective of method.

5. Students taught by a teacher part or all (both) of the time(s) identified dyads more frequently than those instructed by A-T.

Interpretations

In a section titled "Discussion" the author emphasizes the relationship between the reduced distraction and the increased ability to generalize about the examples. He concludes that instructional mode selection should be done in relation to the outcomes desired and the relative effectiveness of the several available modes.

ABSTRACTOR'S ANALYSIS

Studies on the relationship of a variety of modes of instruction to a given set of conditions are likely to be as difficult to design as they are important. Wouldn't it be neat to describe the class, teacher, setting, social context and related variables and then, for a given outcome, to be able to select the most efficient, most permanent, or most whatever mode for the instruction? Unfortunately this study compounds the problem with only two lessons involved by choosing two quite different elements in the sequence. A variety of modes on "exploration" or "invention" might have be ı more productive. Nonetheless, the two selected, A-T and Teacher-d. ected, do show some benefit from personal observation time, especially where this time can be followed by a sharing and reshaping of these observations. This brings us into the whole realm of sequencing instruction to include not only contents and activities but modes of instruction as well.

The study seems well done and the author did not leap beyond his data. The written report is most adequate and the inclusion of the test as appendix a most welcome aspect. Of course, one would like more pupils, more lessons and, thus, more opportunity to generalize. Perhaps, the same results would show anyway on these carefully limited questions.

EJ 086 706

Olsen, Robert Charles, "A Comparative Study of the Effect of Behavioral Objectives on Class Performance and Retention in Physical Science." Journal of Research in Science Teaching, Vol. 10, No. 3:271-277, 1973.

Descriptors--*Academic Achievement, *Behavioral Objectives, Educational Research, *Group Instruction, Instruction, *Physical Sciences, Stience Education, Secondary Grades, *Teaching Procedures

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Russell H. Yeany, Jr., University of Georgia.

Purpose

This study was conducted to compare the achievement and learning retention of ninth grade physical science classes receiving instruction with a prior knowledge of the behaviorally stated objectives to equivalent classes taught the same material without prior knowledge of the objectives.

Rationale

During the last decade, new classroom materials and different methods of instruction have been investigated in attempts to increase learning effectiveness. One method of instruction which has received attention from science educators is the behavioral approach to learning. While advocates of behaviorally stated objectives have called for curriculum development using measurable objectives and the literature suggests many ways to utilize these objectives, the researcher believed that research on the effectiveness of using behavioral objettives in instruction is far from complete.

No underlying assumptions about the relationship of the treatment to psychological theories on previous research were stated.

Research Design and Procedure

Fourteen ninth grade physical science classes from two junior high schools in Ridgewood, New Jersey, were included in the study. With subjects in pre-assembled classes it was necessary to randomly assign intact classes to the two treatments.

The Sequential Test of Educational Progress was used to measure the entering performance of all subjects prior to the start of the experimental study. Pre-experimental data were also collected with the Otis-Lennon Mental Ability Test, Form J.

Eighteen behavioral objectives and 36 assessment tasks were constructured by the researcher and reviewed by a panel of science educators. A list of the action verbs used are included in the published report. The behavioral objectives for each of four chapters were given to the experimental classes, one chapter at a time. The assessment tasks, although not well defined, were provided to insure a more complete exposure and reinforcement for each objective.

The control classes did not receive the objectives but it was not made clear by the researcher whether they used the task assessment. The inference is that, since the tasks were objective specific, they probably did not.

At the conclusion of the three month treatment period, the Interaction of Matter and Energy (IME), also the text which was used, unit achievement test was administered to all classes. Three weeks later, the IME final test was administered and those items appropriate to the experimental objectives were analyzed.

The study can be diagrammed in the Campbell and Stanley notation as follows:

$$R^* \quad O_1 \quad O_2 \quad X_E \quad O_3 \quad O_4$$

$$R \quad O_1 \quad O_2 \quad X_C \quad O_3 \quad O_4$$

* random assignment was made at the classroom level

Analysis of covariance was used to test the effect of treatment on science achievement and retention of the ninth grade students. Initial scores on the STEP and Otis-Lennon IQ scores were used as covariates in the analyses.

Findings

Use of the ANCOVA procedures indicated the experimental group obtained higher adjusted mean scores than the control group in both achievement and retention ($p < .001$).

Interpretations

The author believed that the results of the study support the thesis that providing classes and teachers with behavioral objectives prior to instruction can enhance the performance on achievement tests. Also, the data suggest that the objectives and their accompanying assessment tasks will cause a resistance to forgetting.

ABSTRACTOR'S ANALYSIS

It may seem to be obvious that if a student has access to the instructional objectives learning will be facilitated. Study time should be more task oriented and the student should be better able to gauge his progress in relation to the objectives. But, the research does not always support the above assumption.

From an examination of 39 different studies it can be calculated that approximately 60 percent showed no significant difference in achievement when behavioral objectives were given to students and/or teachers. The age of the subjects in these studies ranged from fifth grade through college and covered a variety of content areas. The evidence is contradictory to say the least.

One dimension lacking in much of the research on the use of objectives is pointed out by a failure to look at a full range of student variables in order to ferret out the type of student who would benefit most from access to the objectives.

In the present study, Olsen may have indirectly identified one such variable. The average IQ of his sample can be calculated to be approximately 115, which is one full standard deviation above the mean of the population of all ninth graders. Whether these results are generalizable across all IQ levels is a question that will go unanswered because the researcher chose to use the variable as a covariate rather than a main factor in the design.

An effort must also be made to sort out the multiple treatment effects that tend to combine when a specific variable is being researched. For example, the investigator was researching not only the effects of providing behavioral objectives but also the use of assessment tasks. In order to examine their singular influence, two more treatment levels are called for (i.e., one group receiving only objectives and one group, only the assessment tasks).

The teacher dimension should never be ignored in any research related to instructional proceedings. Not enough is known about the teachers in this report. Were all teachers using both methods? Random assignment of teachers to treatment levels is usually a very sound decision. But, the fully crossed design with teachers as a factor, where it is possible, does allow one to look for the possible interaction between teacher variables and the method of instruction.

The study should be considered as a bright spot in the search for effective uses of behaviorally stated objectives. And, it provides another link in the matrix of research on the subject but does not resolve the issue.

Future research is needed to look at related factors such as: (1) the student entrance characteristics of mental ability, attitude toward school, personality type and previous success in school, (2) teacher's attitude toward structure and level of familiarity with the use of behavioral objectives, and (3) a whole host of environmental factors including the nature of the content, the taxonomic level of the objectives and the presence or absence of other instruction techniques. All of the above may in some way be related to the results of any instructional strategy and can therefore not be ignored when the effects of using behavioral objectives are being examined. The research must move beyond the two treatment group stage where all these variables are allowed to be random but unaccounted for.

Research on instruction should also look beyond the cognitive domain. If the method promotes cognitive growth at the expense of attitude toward instruction or the content, what have we gained? Or, if there are no significant differences in achievement with gains in attitude, are we not ahead?

EJ 086 723

Talley, Lawerence H., "The Use of Three-Dimensional Visualization as a Moderator in the Higher Cognitive Learning of Concepts in College Level Chemistry." Journal of Research in Science Teaching, Vol. 10, No. 3:263-269, 1973.

Descriptors--*Academic Achievement, *Chemistry, College Science, Critical Thinking, *Educational Research, *Instruction, *Instructional Media, Science Education, Visual Aids

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Milton O. Pella, University of Wisconsin.

Purpose

The problem was: Will cognitive learning of concepts in college level chemistry be moderated by training in three dimensional visualization?

> Null hypothesis: There is no significant difference between the formative test score means of a class of students who receive instruction using visualization as a moderator in cognitive learning of chemistry and a class of students who receive instruction using a didactic approach in learning chemical concepts.

Rationale

The underlying rationale is that learning is assumed to be a function of cognitive-perceptual factors, therefore, higher cognitive learning of chemical interaction is related to three dimensional visualization and this is moderated by manipulating concrete models.

Research Design and Procedure

Experimental design -- Comparative groups were formed from a total population of 102 college freshmen enrolled in chemistry. The class was divided into two groups and one was randomly selected as experimental.

Three dimensional ability was measured with Paper Folding Test V_3-2 and Surface Development Test V_3-3 by L. L. Thurston. The covariants were the scores on the combination of the Thurston Tests, ACT Science Reading, and ACT Mathematics tests.

Experimental factor -- One class was taught using individual molecular model kits from which three dimensional models of chemical species and interactions were contructed following the instructions of the teacher after the species or interaction had been discussed by the teacher.

The second class received teacher descriptions but did not use models. Essentially the presence or absence of models was the experimental factor, however, the problem relates to training in three dimensional visualization.

At the conclusion of each unit of the course formative tests of 50 items (15 analogy, 10 knowledge, 10 comprehension, 5 application, 5 analysis, 5 evaluation) were administered to both groups. Test reliabilities were (KR-21) 0.75 to 0.88.

Findings

1. Experimental group had significant pre-post test gains on Thurston instruments.

2. Control group achieved significantly higher scores than the experimental group on knowledge items.

3. Experimental group achieved significantly higher scores than the control on higher cognitive level items.

4. There was no significant difference between groups on comprehension level items.

Interpretations

1. Learning is a function of cognitive perceptual factors.

2. Null hypothesis is rejected.

3. Use of models results in greater achievement in freshman college chemistry.

4. Students with greater experience in visualization score significantly higher on application, analysis, evaluation, and analogy subtests.

5. Students who are not experienced in visualization perform at essentially the lower levels of knowledge and comprehension, and such performance at these levels has a destructive effect on achievement at higher cognitive levels.

ABSTRACTOR'S ANALYSIS

The total population of 102 students was used in the experiment so the population does not represent a larger population and the use of probability statistics is not proper. The calculations of significant differences are not possible when samples are not used. The only randomization employed was the selection of one half of the 102 to be experimental.

The study lacks specificity; are concepts defined as per the educational psychologists or the scientists? (See Hemple - Philosophy of Natural Science.) Is the knowledge content that makes up the cognitive domain empirical, theoretical, or both (molecular models are theoretical); is the content classificatory or correlational; is the content qualitative, comparative, or quantitative?

Because of the lack of specificity of definition and description in terms of objectives and evaluation, the study does one thing: helps us to see the need for more precise problem statements, i.e., the relative effectiveness of verbal-pictorial and three dimensional analogous models in teaching a concept of compound as (a) the union of two or more elements, or (b) a homogenous substance composed of two or more elements present in definite proportions, or (c) "a substance made up of two or more atoms," or (d) "a substance made up of two or more ions," etc. Note that a and b are empirical; a is descriptive and b is quantitative and c and d are theoretical, c is more primitive than d. It is obvious that there are many concepts and laws in chemistry for which there is no application of molecular models as per those possible through the kits. The content of chemistry is not homogenous in terms of type or complexity as assumed in this paper.

The author exceeded the bounds established in the statement of the problem by his inferences and also by improper use of probability statistics. Note also, if the statistical treatment were proper, that the fact that F ratios vary in magnitude does not mean that those with the larger F are more significant. The existence of significance also does not indicate a cause and effect relationship. In the study it is stated -- the teaching methodology was a primary factor in the greater achievement.

References are made to "changing ways of thinking and 'critical thinking'." These were not part of the problem and are not reasonable inferences even if we knew what the terms meant.

This type of research is extremely valuable. However, there is serious need for greater precision. The results of this study merely indicate that somewhere in the chemistry content covered here the use of three dimensional models worked better than the other approach. It kind of misses the mark of saying those who possess ability in three dimensional perception learn more "chemistry" when models are used than those with a physical deficiency in three dimensional perception. The problem includes the factor "by training in three dimensional visualization" yet there is no evidence of training other than using the model kit.

This study fails to meet the criterion of generalizability because the population may not be assumed to represent any population beyond the 102 used.

STUDENT ATTITUDES

EJ 071 069

Milson, James L., "The Development and Evaluation of Physical Science Curriculum Materials Designed to Improve Student Attitudes." Journal of Research in Science Teaching, Vol. 9, No. 4:289-304, 1972.

Descriptors--*Attitudes, Curriculum, *Curriculum Development, *Curriculum Evaluation, *Educational Research, Intelligence, Physical Sciences, Secondary School Science, *Slow Learners

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

Purpose

The investigator intended to develop and evaluate physical science materials that were suitable for producing positive attitudes toward science instruction when used with below average secondary school students (referred to henceforth as lower ability students). The development of the materials was to be guided by characteristics of lower ability students as identified through a survey of the literature.

Rationale

Characteristics of lower ability students, as summarized from empirical descriptions, did not reflect any particular conceptual framework. Factors such as low IQ (75-90), reading difficulty, problems in abstraction, association, and generalizing were cited. While many frameworks exist in the literature linking these factors to instruction, none were cited.

Although not explicitly stated as such, the study assumed that a strong link exists between student interest, personality, and academic performance. The implication was that if school situations were changed, the attitude of the student would also change. These changes should, if the conjectures were correct, positively influence academic performance.

Research Design and Procedure

The two phases of the study were described separately. Phase one was concerned with the guidelines and development of materials for lower ability students and phase two, with the evaluation of these materials. During phase one, an elaborate set of guidelines was presented. How they were derived from the characteristics of lower ability students was not described. Specific information about the materials utilized and their evaluation in the study was not given; only general information was stated, which claimed that the "materials conformed to the guidelines." No other specific information about any particularly activity, lesson material, or teacher's classroom performance--expected or observed--was reported.

During the evaluation phase of the study, one set of materials concerned with the study of heat and temperature was utilized. Nine experiments and five reading assignments were involved. The specific natures of experiments and reading assignments were not described. The design utilized was identified as a "nonequivalent control group design." No random assignment of students to groups was made; instead, intact groups, as constructed by the school, were assigned to experimental or control status. (The n used in the analysis was the number of students). Out of the six lower ability classes, three were assigned to the experimental group and three to the control group. (This description here may be in error since, from the study, it was difficult to determine who was assigned to what). By chance, two teachers taught the three experimental classes and two taught the control classes. It was carefully pointed out that no teacher taught both types of classes--experimental and control. Evidently the classes must have been small since the total number of students reported in the experimental group was only 23 and in the control group, 26. No comments were made about high attrition rate during the several weeks the study took place. The students were mostly ninth and some tenth grade high school students.

Before and after their participation in the sample unit on Heat and Temperature, the students were given some form of Osgood's Semantic Differential. Five concept areas were assessed: science classes (sc), science laboratory (sl), science teacher (st), the school (s), and specific science topic (c). The specific set of adjectives used for each of these concept areas was not reported nor was the procedure described as to how the "factor scores" were calculated. Some simple arithmetic suggests that the scores were raw summative scale scores where the responses of the students, which probably varied from one to seven, were simply added. High scores were hence indicators of positive attitudes and low scores, negative ones. However, this procedure is only this reviewer's assumption.

Analysis of data included ANOVA to test pre-to-post differences within the experimental group and within the control group. Between group differences were similarly assessed. Each of the concepts included the three typical semantic factors; evaluation, power, and activity. No intercorrelation was reported to establish that the observed scores for these factors within the pretest, or within the posttest, were independent. However, independence of pre-to-post data was tested using a regression technique. No achievement measure was reported, even though achievement was a consideration in the rationale.

Findings

Significant differences were found for two of the concept areas within the experimental group; lower ability students within the group using the specially prepared materials displayed a more positive attitude, as assessed by the semantic differential administered, towards the science class (sc) and the science laboratory (sl). No pre-to-post differences were found for their attitude on the other factors (st, s, c). No pre-to-post differences were found for lower ability students not using the specially prepared materials. According to the data reported, pretest scores did not predict the posttest scores according to regression technique. This was interpreted to indicate that the change in attitudes assessed was the result of the materials and activities utilized with the experimental group and not of other external factors.

Interpretations

Only one major interpretation was proposed and it generally restated the findings. That was, that the materials did significantly improve the attitudes of lower ability students toward science class and science laboratory type work and that this change was due to the materials and not due to other external factors. The author suggested that the study demonstrated an effective approach for developing marerials designed to improve attitudes of lower ability students.

ABSTRACTOR'S ANALYSIS

Assessing attitudes and attitude changes remains a difficult task without many clearly defined conceptual frameworks. Reviews of the literature in this area reflect this problem and tend to reveal more conflict than consensus. Clearly the problem of suitable instrumentation is a major contributor to this confusion. Most researchers in this area consider the semantic differential technique far less appropriate than are other techniques. In this case, the population consisted of mostly lower ability students with reading difficulties and, judging from the IQ range, low verbal comprehension as well. Their complex, emotional responses were, however, measured by a verbal instrument requiring the student to perform some rather high level verbal discrimination. Research of this type would strongly benefit from other, more relevant techniques.

While instrumentation in attitude studies is always a first problem to be solved, a second, more subtle but still important, problem also needs attenrion. This one is concerned with the actual treatment. Charters and Jones (1972) suggest that too often research is assessing "non-evenrs." That is, from the student's perspective, nothing different has occurred. In this study, something different did occur. The experimental group did something different than the conrrol group. However, what that was remains unknown as no descriprions of student activity were provided. Control group classes did something, but not even the name of the materials used was provided. Since the point of the study is to eventually lead to some association between class activity and student attitude, it is fundamentally necessary thar the class activities be described for not only the experimental group but also for the control group.

Many sets of data were presented in the study; eleven tables in all. However, only two are really relevant to the study's main premise. Since no variance measures, such as standard deviation, are reported, many in-depth reflections remain beyond review here. Without this information, the validity of comparing scores from experimental to control remains in question, since the two groups are, to begin with, non-equivalent. Are there variances within a suitable range that these sets of scores can be compared in their raw form using ANOVA? Also limiting further interpretation of the data is the complete absence of any explanation of what the semantic descriptors were and how student responses were combined into some score. The lack of testing for independence of the factor scores raises some question about the importance of the significant differences found. Were there really six significant differences within the pre-to-post experimental data or was there only one significant difference measured six times? Without any test of independence the answer will remain unknown.

The regression technique attempted to establish that the posttest factor could not in most cases be predicted by some combination of the pretest factors. Results of this analysis were reported in Tables 6 and 11. It should be pointed out that each table reports the results of 25 regression analyses for any one posttest factor. All pretest factors <u>except</u> the corresponding pretest factor were combined to see if they could predict the outcome. It can be also stated as follows:

$$(p''E_{sc} = f\ (p'E_{sl},\ p'E_{st},\ p'E_{s},\ p'E_{c})$$

where $p''E$ is a posttest factor, $p'E$ is a pretest factor, and f represents some function that associates them and is equivalent to R^2. Notice that the pretest factor (sc) is omitted since the regression is trying to predict the (sc) posttest factor. In examining these data, the degrees of freedom should be 4, 18 <u>not</u> 1, 18. Some other checks of the F's reported support a suspicion that inappropriate regression information is included in these tables. Furthermore, keeping in mind that the regression represents predicting <u>without</u> the corresponding pretest factor, the lack of significance <u>may</u> mean that pretest scores do predict posttest scores, but one to one. If this were true, then the interpretation provided is as misleading as is excessive manipulation of data. A simple, straight forward analysis would do much to clarify many questions. Correlation would no doubt be amply suitable, considering the generally high level of random error probably existing in data of this type. The existence of this random error would contribute to independence as random numbers neither correlate nor predict one another.

Assuming at face value that the data have meaning, more should be said concerning where the differences occurred and what they mean. Specifically, the control group exhibited <u>no</u> change in attitude while the experimental group did. What were the classes in the experimental group doing? The experiments included in the heat and temperature unit include many interesting manipulations of equipment. Lower ability students involved in this group would normally be reading books and listening to the teacher. These students find book work too often difficult even though they get lots of it anyway. A change in pace is a welcome relief. The shifts in attitude for class and laboratory are not unexpected. However, the topic of those activities would seem to be of little student concern since no shifts on that factor (c) occurred. Similar lack of shifting in school (s) or teacher (st) factors also support the notion that the change in instructional pattern affected their attitude toward that instructional pattern, but probably had little effect on their attitude toward science.

REFERENCE

Charters, W. W., Jr. and John E. Jones. "On the Risk of Appraising Non-Events in Program Evaluation," Educational Researcher, 1:5-7, November, 1972.

EJ 064 264

Schwirian, Patricia M. and Barbara Thomson, "Changing Attitudes Toward Science: Undergraduates in 1967 and 1971." Journal of Research in Science Teaching, Vol. 9, No. 3:253-259, 1972.
Descriptors--*Attitudes, College Science, *Changing Attitudes, Education Majors, *Student Characteristics, Socioeconomic Influences

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Ronald D. Simpson, North Carolina State University.

Purpose

The purpose of this study was to examine changes in attitude among undergraduate university students toward "science and the scientific institution" between 1967 and 1971. The Schwirian Science Support Scale was used to measure these potential changes.

Rationale

During the time this investigation was being completed and this article was being written, there was a definite "feeling" among science educators that our country was experiencing a backlash toward science and technology. Elementary and secondary school students were fighting pollution and "ecology" became a household word almost overnight. Large sums of money had been pumped into scientific endeavors all through the sixties; yet the mood of our country seemed to be a troubled one, with more complex problems seemingly than ever before. The investigators in this study, therefore, sought to generate comparative data that would substantiate or refute the notion that students in 1971 possessed less positive attitudes toward science than their counterparts in 1967.

Research Design and Procedure

The Schwirian Science Support Scale (Tri-S Scale) was administered to 398 undergraduates (mostly education majors) at a large midwestern university in 1967. Different undergraduates taking the same course at the same institution in 1971 were administered the Tri-S Scale again, a total of 153 students this time. Scores were based on a 20-item, shortened version (Form A) of the original 60-item scale. Since the Tri-S is a 5-point, Likert-type scale, a 20-item form would produce a minimum score of 20 and a maximum score of 100.

The data analysis consisted of nine two-way analyses of variance. The major independent variable, time of administration (1967 and 1971), was always one of the factors. The second factor in each ANOVA consisted of each of the nine following contingent independent variables: (1) student's age; (2) student's sex; (3) student's religious preference; (4) father's education; (5) mother's education, (6) father's occupation; (7) student's academic major; (8) student's hometown size; and (9) the type of high school from which the student was graduated.

Findings

The single, one-way analysis of variance performed between 1967 and 1971 respondents exhibited no statistically significant difference. In fact, the 1967 and 1971 group mean scores were 75.47 and 75.14 respectively, indicating almost identical reactions to the Tri-S Scale by the two groups of students compared in this study.

When data were analyzed by ANOVA relative to the nine contingent independent variables, only two produced statistically significant differences--father's occupation and student's hometown size. Aside from this, no significant differences in science attitude scores were observed, either as a function of the major independent variable, time, or as a function of the other seven contingent independent demographic variables.

Interpretations

This study was designed to investigate potential changes in attitudes toward science between students taking a similar course at a large midwestern university in 1967 and 1971. The attitudes of these two groups of education majors toward science and the scientific enterprise did not change as a function of time as measured by the Tri-S Scale. On this count the investigators concluded: "Perhaps the criticism of science that we hear from among the young is not so much a criticism of the value of science itself, but of the lack of responsiveness of the scientific institution to their cry for assistance in the amelioration of what to the youth seem to be mankind's overwhelming problems."

Data from this study indicated that, in general, the "higher" the occupational level of the student's father (laborer, blue collar, white collar and professional), the more positive was the student's attitude toward science, with leveling-off noted between the offspring of white collar workers and professional men. Statistically significant differences between mean Tri-S scores were also seen relative to the size of the hometown in which the students lived while they attended high school. Students in this study from farms or from communities of over 100,000 in population exhibited little difference in attitudes toward science in 1971 than they did in 1967. Conversely, the students from medium and small-size communities showed a more positive attitude toward science in 1971 than they did in 1967. The authors of this paper attributed increases in positiveness of scientific attitudes to recent increases in coverage of scientific and technological events through mass media and improved science education programs in the schools. Small and medium size communities, the investigators stated, have experienced these advancements to a larger degree than have farm-isolated and big-city environments.

ABSTRACTOR'S ANALYSIS

One of the most active yet challenging areas of educational research today is that of measuring affective outcomes of students and searching for relationships that potentially exist among attitudes, values, and cognition. Studies like this one are very worthwhile and more are needed. This investigation not only considered relationships among attitude toward science and nine demographic variables, but sought to establish a data base over time whereby changes in attitude could be quantified. Statistical procedures used in this study appear appropriate, and the investigation was communicated to the reader in a clear, well-written manner.

As with any research effort, there is always room for constructive criticism. Investigations where the feelings of people are examined are particularly difficult to design, control, and make generalizations from. Having been involved in research of this type, this abstractor can appreciate the many difficulties associated with investigations like this one. Having appropriate populations available for study, having valid and reliable instruments available for quantifying the affective entities in question, and drawing conclusions from data that are usually very global and complex represent only a few of the typical problems faced. Since the phenomenon of "attitude" is essentially a psychological construct, research efforts designed to approach this topic experimentally are seldom, if ever, as "clean" as research where more discrete variables are being studied.

In my opinion, weaknesses associated with this study fall into two major categories. First, the sample--particularly the 1971 group --is restrictive. The investigators, through ANOVA, search for differences among various demographic characteristics, yet in some groupings there are less than 10 students represented. In fact, the investigators stated at the beginning of their paper that the purpose of the study was to examine change in attitudes toward science which had occurred among "undergraduate university students" between 1967 and 1971. Table I in their article, however, shows that out of the 153 respondents, 148 are education majors (presumably elementary education), one is a liberal arts major, and three are "other." Since all these students are enrolled in the same course at the same university and are largely majoring in the same thing, it is difficult for me to label this group "undergraduate university students." In short, comparing a rather select group of majors from only one university enrolled in an elementary education course in 1967 with a similar but much smaller group enrolled in the same type of course, only four years later, hardly speaks for "change in attitudes toward science of undergraduate university students."

The second category of concern by the abstractor associated with this investigation is the instrument used in measuring "attitudes toward science and the scientific institution." While earlier research by Schwirian (2) reported favorable reliability estimates with the Tri-S when used with university undergraduate students, a research report by Simpson, et al (3) questioned the validity of this instrument.

When the Tri-S Scale was used with 618 high school students, it was found to contain many items too abstract for eliciting student feelings toward science. The Tri-S is based on five basic cultural values identified by Bernard Barber (1) that are considered necessary before science as an enterprise can flourish in any society. Whether or not this instrument actually measures a student's attitude toward science as an object of affect is still questionable.

Before investigators of attitude change can hope to assess the attitudes of any population toward any object, they must be assured that the instrument or technique used contains tangible or concrete items to which respondents can attach their feelings. Moreover, these stimuli should represent valid or "real" dimensions of the attitude object and be structured to discriminate a wide range of feelings. Close examination of the Tri-S Scale by this abstractor raises serious questions concerning the appropriateness of this instrument to accomplish this task--namely that of carefully determining how students feel toward the products and processes of science.

REFERENCES

1. Barber, Bernard. Science and the Social Order. Collier, 1961.

2. Schwirian, P. M. "On Measuring Attitudes Toward Science," Science Education, 52:172-179, 1968.

3. Simpson, Ronald, D., John W. Shrum and R. Robert Rentz. "The Science Support Scale: Its Appropriateness with High School Students," Journal of Research in Science Teaching, 9:123-126, 1972.

EJ 061 298

Simpson, Ronald D., John W. Shrum and R. Robert Rentz, "The Science Support Scale: Its Appropriateness with High School Students." Journal of Research in Science Teaching, Vol. 9, No. 2:123-126, 1972.

Descriptors--*Attitudes, *Biology, Educational Research, Factor Analysis, *Measurement Instruments, Secondary School Science, *Validity

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Robert L. Steiner, The Ohio State University.

Purpose

The authors' stated purposes were twofold: (1) to see if differences in teacher science support, as measured by the Science Support Scale, Tri-S, (Schwirian, 1968) produced differences in student science support, and (2) to assess the appropriateness of using the Tri-S as an instrument for measuring attitudes toward the support of science among high school students.

"It was hypothesized that teachers with 'high' support of science would produce students with significantly higher support of science and that the Tri-S would be a valid and reliable tool to measure these attitudes."

Rationale

The need of greater emphasis on the affective domain in science teaching and learning is often expressed by science educators. The study of the affective domain entails at least two significant aspects: (1) accurately assessing or measuring effective attributes, and (2) effecting changes in affective attributes. Both are important aspects to study, but it is difficult to do the second without having adequate instrumentation to be able to do the first.

The assessment or measurement of affective attributes necessitates the use of appropriate, valid, and reliable instruments and techniques. Instruments used to measure achievement in non-cognitive areas should be carefully analyzed before they are recommended for use with specific populations.

Non-cognitive outcomes are learned, just as are cognitive outcomes. We should be concerned with those factors which contribute to this learning. Teachers with a different level of a non-cognitive attribute might produce students with different levels of the non-cognitive attribute.

Research Design and Procedure

The study was basically a correlational study lacking a specific experimental treatment from which a causal relationship might be implied. The authors looked at the possible relationship between student scores on the Tri-S and the scores of their teachers whose class they had been in for the academic year.

All tenth grade biology teachers (N = 24) in a large metropolitan school system (9 high schools) near Atlanta were administered the Tri-S at the beginning of the 1969-70 school year. One of each teacher's biology classes was randomly selected and the Tri-S was administered to the class during the first two weeks of September and again during the first two weeks of May. There were 618 students in the 24 selected classes who completed both administrations of the Tri-S.

The 24 teachers were separated into four groups of six each, according to their score on the Tri-S. The groups ranged from a "high" science support (group 1) to a "low" science support (group 4) with two intermediate groups (2 and 3).

Student scores were analyzed for the following: (1) gain from pretest to posttest, (2) differences between students of the four teacher groups on each of the five subscales of the Tri-S and on the total scale, (3) reliability estimates for each of the subscales and the total scale, and (4) a factor analysis of the inter-item correlations.

The reliability of the pretest administration of the Tri-S was determined using Cronbach's coefficient alpha, an estimate of internal consistency. The factor analysis of the inter-item correlations was used to examine the construct validity of the instrument.

Findings

The students of the four different teacher groups were not significantly different on four of the five subscales and on the total Tri-S. The F value obtained was significant at the 0.01 level on one of the subscales.

Duncan's Multiple Range test indicated that students of the teachers in the high group (group 1) scored significantly lower than the students in the other teacher groups. The students in the low teacher group (group 4) scored significantly higher than the students in teacher group 3.

A rank order correlation coefficient between the teachers' scores and the posttest class mean scores of their students was only 0.128.

The subscsle reliability estimates using the Cronbach's coefficient alpha ranged from 0.338 to 0.558 which were all lower than the results obtained by Schwirian.

Factor analysis did not yield the predicted five subscales of the instrument. The item factor loadings were generally low. Some items loaded on more than one factor. The first factor was the largest, with almost one half of the items loading on it, and only a few items loaded on the last two extracted and rotated factors. In no case did all items of the theoretical subscale load on a single factor.

Interpretations

Student scores did not significantly increase from pretest to posttest, suggesting that either science support as meaaured by the Tri-S was not increased by an introductory biology class or that the Tri-S was not sensitive in measuring gains in affective achievement relative to cultural values within the Tri-S.

Students of high science support teachers did not score significantly higher than the students of low science support teachers, suggesting that teachers with high science support were no more effective in communicating this attitude to students than were teachers with low science support scores. In fact, no teachers were particularly effective in communicating science support attitudes to the students. Another possibility suggested by the authors was that the Tri-S was possibly not measuring increases in science support.

Based on teacher feedback and reliability estimates on the subscales and on the total instrument, the authors suggested that students had difficulty comprehending many items on the scale. The factor analysis of the student responses indicated that not all items on the subscales were being interpreted in the same manner by all students.

The authors suggest that the Tri-S is not a reliable instrument to measure achievement of high school students in the affective domain. Apparently "when the Tri-S is administered to tenth grade students, it is not a valid assessment of the five cultural values which is purports to measure."

It appears "that the Tri-S is either not a valid tool for assessing growth in science support among high school students or that the teachers in this study were not effective in bringing about significant attitudinal changes during the introductory course in biology."

ABSTRACTOR'S ANALYSIS

The authors have addressed themselves to a problem which is quite prevalent in science education research within the affective domain. That is, the problem of reliable and valid assessment. Numerous instruments developed and piloted in dissertation research are either never used again or are assumed to be valid and reliable instruments for use with populations other than the one used in the instrument development. Perhaps if fewer instruments were developed and more time were devoted to verifying, modifying, improving and increasing the usability of already-developed instruments, science education research would be ahead.

The authors' have critically analyzed one such instrument, the Tri-S, to determine its appropriateness with a population other than the one with which it was developed. The instrument was developed using undergraduate college students, but the instrument developer suggested that the instrument might be appropriate for use with high school students. The authors examined the use of the Tri-S with high school biology students in a large metropolitan school system.

The authors' analysis of the Tri-S represents a useful and necessary piece of research, but the second part of their work, to determine if teachers with different science support scores produce students with different science support scores, is questionable.

At best, the authors' could only look for a relationship between teacher and student scores. To hypothesize a causal relationship is totally outside the realm of this study as it was carried out.

There does not appear to be any theoretical basis for expecting that students will increase their scores on the Tri-S. There is no indication that an objective of the biology teachers, or courses, was to increase science support. Similarly there is no reason to expect that teachers with high science support scores will produce students with higher science support scores. In other words, there really was not an experimental treatment from which predictions could be made and a causal hypothesis tested.

Because the student scores on the Tri-S did not increase from pretest to posttest, the authors' concluded that "either 'science support' as measured by the Tri-S was not increased as a result of an introductory biology course or that the Tri-S was not sensitive in measuring gains in affective achievement relative to the cultural values within the Tri-S."

It is perfectly possible for the instrument to be valid and reliable, and for the course not to produce any change in science support. It is also possible for the instrument not to be valid and reliable and for the course to produce changes in science support. However, since there does not appear to be any controlled effort to produce a change in science support, there is no reason to expect one, making this part of the research rather questionable.

The question of the validity and reliability of the Tri-S with high school students could be examined without being concerned whether tenth grade biology teachers produced any differences in student science support.

The authors' found the reliabilities of the subscales and the total Tri-S to be lower than Schwirian's original results. Based on teacher feedback, they concluded that students had difficulty comprehending many items on the Tri-S. The results are not surprising since the range of student abilities in the tenth grade of high school would be much greater than in undergraduate college classes. The authors' use of the Cronbach coefficient alpha would be a better estimate of the internal consistency than the split half method used by Schwirian. Whereas the Cronbach coefficient alpha is the average of the split halves, the split half method used by Schwirian could vary depending on the particular split used to determine the correlation.

Factor analysis is a method increasingly being used to help determine the construct validity of instruments. The authors' findings support their conclusion that the validity of the Tri-S used with high school students is questionable. They suggest that certain emotionally charged words such as homosexual, sex deviants, godlessness appeared to distract students from the meaning of the subscales in which the items were found. That less mature high school tenth grade students would be more reactive to words like the above than would college students would not be surprising.

The statistical analyses used in this study should have been confined to those which indicate relationships between the teachers' scores in the Tri-S and their students' class mean scores on the Tri-S. A number of computer programs allow the researcher to look for linear, as well as higher order, relationships between variables. This type of analysis would have been more appropriate and powerful for the researchers to use than was the analysis of covariance technique used.

Although the relationship between teachers' and students' science support would best be examined utilizing correlational analysis, many correlational studies do use an analysis of variance technique as the authors have done. Since the authors were comparing the four teacher-student groups on each subscale and on the total Tri-S using the pretest scale scores as a covariant, perhaps a multivariate test would have been more appropriate, rather than the six univariate tests. The univariate tests would have been appropriate if the subscales and the Tri-S were relatively independent, that is, with small intercorrelations. The authors do not report the intercorrelations between subscales and the Tri-S, so it is impossible to determine if a multivariate test was more appropriate.

The use of the Duncan Multiple Range test to determine if significant differences existed between the individual student groups on the one subscale of the Tri-S for which a significant F-value was found is somewhat questionable. Glass and Stanley (1970) have indicated that "Duncan's procedure has been the most popular in the behavioral sciences and education; but until mathematical statisticians resolve their differences or satisfy one another that the derivation of the procedure is valid, it might be wise for persons in applied fields to observe a moratorium on Duncan's procedure" (p. 383). The Tukey post-hoc multiple comparison procedure would have been more appropriate.

Most of the values in the tables of the report are given with at least three and sometimes five significant figures, yet the original measures of the variables did not have this many significant figures. As science educators, we should use care that our reports do not imply greater precision and accuracy than is warranted.

I believe that the study is lacking with respect to the second problem due to the absence of a theoretical basis and treatment for predicting any change in student attitudes toward the support of science, but, overall, the authors have addressed a serious problem. Their conclusion that the appropriateness of the use of the Tri-S with tenth grade students is questionable is supported by their study and should be heeded by others planning to use the scale. Another study raising some concerns regarding the use of the Tri-S has been carried out by Wideen (1971). The current study clearly points out the need to carefully pilot instruments with populations with which the instrument is to be used. It also suggests the desirability of improving existing instruments instead of starting from scratch to develop a new instrument.

REFERENCES

Glass, Gene V. and Julian C. Stanley. Statistical Methods in Education and Psychology. Englewood Cliffs: Prentice-Hall, 1970.

Schwirian, Patricia M., "On Measuring Attitudes Toward Science." Science Education, 52:172-179, 1968.

Wideen, Marvin F., "Will Students Fake a Science Support Scale When Given Direct and Indirect Instructions to do so?" Paper presented at the 44th Annual Meeting of the National Association for Research in Science Teaching, Silver Spring, Maryland, March 23-25, 1971.

TEACHER EDUCATION

EJ 022 704

Merkle, Dale G., "A Leadership Workshop on Elementary School Science; An In-Depth Evaluation." Journal of Research in Science Teaching, Vol. 7, No. 2:121-133, 1970.

Descriptors--*College Teachers, *Elementary School Science, Inservice Education, *Program Descriptions, *Program Evaluation, *Teacher Education, Workshops

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

Purpose

During the summer of 1968, an NSF-sponsored Leadership Workshop on Elementary School Science was held at Michigan State University. The 30 participants included elementary science consultants, science supervisors and college and university professors currently teaching science and/or science education courses for pre-service elementary school teachers.

The purposes of this workshop were to determine the relative effectiveness of different workshop activities and to determine how these activities related to subsequent changes in the behavior of the workshop participants. A further objective was to follow up the effectiveness of the participants as they engaged in post-workshop implementation activities.

The four-week workshop provided the participants with an orientation to the AAAS Science-A Process Approach materials and the Science Curriculum Improvement Study material. The participants used these materials in micro-teaching experiences with children as well as in observing experienced teachers using these two curricula with children.

The workshop participants received training in both group process skills and change-agent skills. These skills were considered important since the participants were also involved in preparing appropriate materials and activities for future orientation sessions and in-service programs. During the final week of the workshop nine participant teams were formed and these teams instructed elementary school teachers for three days at the Michigan Education Association Camp.

A mid-winter evaluation conference was later attended by 28 of the college teachers. Through group and individual interaction sessions the staff obtained feedback from the participants.

Research Design and Procedure

The seven instruments used in this study were developed by the staff of Michigan State University. The instruments or questionnaires were generally intended to measure change in participant response over time or to seek evidence of linear relationships among the instruments.

Instruments A, B, and E were administered in the pretest, posttest, and mid-winter conference. Instruments C and D were used in the pretest and posttest. Instrument F was used in the posttest and mid-winter conference. Instrument G was used only in the mid-winter conference.

Instrument A, Knowledge of Program Characteristics and Program Implementation, was a fifty-seven item multiple choice instrument which assessed knowledge of the AAAS and SCIS programs. Instrument B, Attitude Toward The AAAS and SCIS Programs and the Content of These Programs, combined sixteen Likert Scale items plus two questions related to program preference. Instrument C, Analysis of Personal Behavior in Groups, was an attempt to measure the effectiveness of the group process sessions using a weighted ranking. Instrument D, Knowledge of Change-Agent Strategies, was used to measure change-agent skills of the participants. Instrument E, Satisfaction of Perceived Needs, was used to determine participant needs in areas of school community expectations and the new science curricula with responses scored on a one to five weighted basis. Instrument F, Evaluation of the Divisions of the Workshop, called for the participants' opinions on statements related to five components of the workshop. Instrument G was a questionnaire which asked for evidence of change in such areas as number of recent formal class sessions devoted to AAAS and SCIS in contrast to similar activities the previous year.

The first four hypotheses of the study tested the significance of the difference between pretest and posttest scores of the 30 participants on each of the instruments A, B, C, and D. A one-tailed test of significance was used with the paired t-test.

The last four hypotheses were attempts to show significant Pearson product-moment correlations among different measures. The tesearcher indicated the possible presence of these correlations would assist future designers of workshops in the selection of activities and participants.

Findings

The tests of the first four hypotheses indicated that there were significant differences between pretest and posttest scores on instruments A, B, C, and D. The workshop was described as a successful vehicle for increasing the knowledge of the participants in the topics of the workshop and for creating a positive attitude toward the AAAS and SCIS programs. The participants also showed marked gains in knowledge of change-agent skills. The final four hypotheses dealt with correlations among the instruments and findings indicated no significant relationships.

The study also included descriptive data that were not statistically treated. For example on post-workshop Instrument F, the participants ranked the participant-run three-day workshop as the most important activity. The participants ranked the sessions on Change-Agent Strategies and Group Processes low in every category of measure.

Instrument E provided data to contrast the perceived pretest needs of the participants with the posttest satisfied needs of the participants. The researcher indicated that those needs considered to be the most outstanding needs were satisfied almost fully.

Instrument G provided descriptive data indicating that the participants had increased the number of class sessions devoted to AAAS and SCIS and increased the number of workshops as contrasted with efforts the previous year. The participants reported an increase in colleagues using the new science curricula as well as increases in the number of consulting requests.

Interpretations

The researcher indicated that cognitive and affective changes were brought about in this leadership workshop.

The information obtained on Instrument F indicated that a three-day workshop run by participants should be included in subsequent summer workshops. Future workshops should consider giving less time for group process skills and change-agent sessions.

From the participants' mid-winter conference responses on Instrument G, the researcher concluded that a definite increase in the use of the processes of science and a definite re-alignment of many pre-service programs had begun as a partial result of the leadership workshop.

ABSTRACTOR'S ANALYSIS

In this descriptive study of a NSF leadership workshop, all evaluation instruments were locally constructed. No evidence was presented regarding reliability or validity of the seven instruments. The lack of significant correlations among the instruments may possibly be related to low validity and low reliability of these measures.

The use of a one group, pretest, posttest design served only to indicate change in certain cognitive and affective areas without excluding alternative explanations such as increased sensitization due to pretests. The brief description of Instrument C implied that some type of ranking was involved in the responses which would suggest the use of an appropriate nonparametric statistic.

This research study exemplifies the problems of extending a study over too many areas using unproven instruments. Helgeson (3), after a review of 138 NSF Teacher Institute research studies, pointed out the difficulty in pinpointing the precise nature of the impact of institutes on education.

In fairness to the author of the workshop article, it should be noted that in 1968 the NSF placed only limited emphasis on formal evaluation of workshop effectiveness. Nevertheless the workshop staff did initiate evaluation activities in several areas previously slighted. The innovative areas investigated at the workshop included:

1. identification of participant perceived needs and subsequent degree of needs satisfaction,

2. formative reactionnaires administered after each major activity,

3. evidence of attitudinal change,

4. effectiveness of participants in initiating changes in the colleges or schools,and

5. emphasis on change-agent strategies.

The following cursory review of these five research areas may suggest alternative means of evaluating programs for inservice educators.

Perceived Needs. Hall, Wallace, and Dossett (1) described a Concerns-Based Adoption Model with seven operationally defined stages of concerns believed to be generalizable to most educational innovations. A Concerns Questionnaire based on the model is currently being validated with a sample of SCIS science teachers. The availability of a valid instrument would be helpful in seeking a match between workshop participants' concerns and the subsequent instruction.

Formative Evaluation. Rowe and Hurd (7) described an engineering research approach for use in inservice programs to diagnose sources of resistance to innovation. The measures used in the study included two three-part scales labeled Not Interesting, Interesting, Very Interesting and Not Useful, Useful, and Very Useful. The instruments were administered after each of six inservice sessions. The results supplied immediate feedback on the status of the session. The use of the Contingency Coefficient as described by Siegel would be useful in detecting significant differences in the perceived usefulness of specific activities among different groups such as college teachers and curriculum supervisors.

Attitudinal Change. Moore (6) used the Science Teaching Attitude Scales in a two-year study of attitudinal changes in a NSF-CCSS group of elementary school teachers. The results of the study indicated that the workshop had an observable long-range positive effect on science teaching attitudes but not on attitude toward science. The careful development, field test, and validation of the instrument by Moore (5) serves as an excellent model.

Levels of Use. Hall, Loucks, Rutherford, and Newlove (2) operationally defined eight sequential levels of use of an innovation. An interview technique is presently being used to classify users of specific innovations. A workshop director would be materially assisted in building flexibility into a program if he or she had prior knowledge of both the concern stage and level of usage of participants interested in new science curricula.

Change-Agent Strategies. Mahan (4) described a four-year effort of the Eastern Regional Institute for Education (ERIE) to implement AAAS Science - A Process Approach materials in fifty-three elementary schools. Sixteen guidelines for curriculum installers emerged from the results of two contrasting installation strategies used by ERIE. The guidelines and supporting evidence should be on the reading list of all aspiring change-agents.

REFERENCES

1. Hall, G. E., R. C. Wallace, and W. F. Dossett. A Developmental Conceptualization of the Adoption Process Within Educational Institutions. Austin, Texas: Research and Development Center for Teacher Education, University of Texas, 1973.

2. Hall, G. E., S. F. Loucks, W. L. Rutherford, and B. W. Newlove. "Levels of Use of the Innovation: A Framework for Analyzing Innovation Adoption," Journal of Teacher Education, Vol. 26: 52-56, Spring, 1975.

3. Helgeson, Stanley L. "Impact of the National Science Foundation Teacher Institute Program." Information Bulletin, No. 2, 1974-75. Columbus, Ohio: ERIC/SMEAC.

4. Mahan, James M. "Overview of a Systematic Effort to Engineer and Monitor Curriculum Change: Emerging Guidelines and Encouraging Findings for Curriculum Installers." Paper presented at the annual meeting of the American Educational Research Association, February 6, 1971, New York.

5. Moore, Richard W. "The Development, Field Test, and Validation of Scales to Assess Teachers' Attitudes Toward Teaching Elementary School Science," Science Education, Vol. 57, No. 3:271-278, 1973.

6. Moore, Richard W. "A Two-Year Study of a CCSS Group's Attitudes Toward Science and Science Teaching," School Science and Mathematics, Vol. 59:288-290, 1975.

7. Rowe, Mary Budd and Paul DeHart Hurd. "The Use of Inservice Programs to Diagnose Sources of Resistance to Innovation," Journal of Research in Science Teaching, Vol. 4, No. 1: 3-15, 1966.

EJ 034 786

Roberts, Richard A. and Jacob W. Blankenship, "The Relationship Between the Change in Pupil Control Ideology of Student Teachers and the Student Teacher's Perception of the Cooperating Teacher's Pupil Control Ideology." Journal of Research in Science Teaching, Vol. 7, No. 4:315-320, 1970.

Descriptors--*Cooperating Teachers, Class Management, *Elementary School Teachers, Inquiry Training, Preservice Education, *Student Teaching, *Science Education, *Teacher Education, Teacher Behavior

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Vincent N. Lunetta, University of Iowa.

Purpose

The study investigated change in the pupil control ideology (PCI) of elementary student teachers and the relationship of this change to the student teacher's perception of the cooperating teacher's pupil control ideology. The research hypotheses tested were:

1. Is there a change in PCI during student teaching?
2. Is there a difference in student teaching grade level between those student teachers showing no increase in custodialism and those showing greatest increase in custodialism?
3. Is there a difference in the perceived PCI of cooperating teachers between those student teachers showing no increase in custodialism and those showing greatest increase in custodialism?
4. Is there a difference in mean socialization pressure experienced by those student teachers showing no increase in custodialism and those showing greatest increase in custodialism?
5. Are there differences in the change in PCI of student teachers in situations of low, medium, and high socialization pressure?

Rationale

The authors base the need for the study upon the fact that "successful implementation of new curricula...requires careful consideration of...the teacher-pupil relationship." The contemporary emphasis upon inquiry teaching requires that the student's role change "from one of dependency upon the teacher to one of competency in directing his or her own inquiry activities." Yet, the authors surmise that teachers may be reluctant to relinquish control over student activity, and this may inhibit the effectiveness of an inquiry approach to teaching. The authors write that a review of the literature revealed that the behaviors and attitudes of the cooperating teacher have a significant influence on the socialization process of preservice teacher education.

The report cites a very limited number of relevant research studies. It reports one study by Meyer Horowitz (1) indicating that the differences the student teacher perceives between his views and those of his cooperating teacher are even more crucial than are the actual differences.

Research Design and Procedure

Pupil Control Ideology (PCI) was measured on a twenty-item instrument developed by Willower, Eidell, and Hoy (2). The PCI form was administered to preservice elementary teachers (K-6) in the last semester of their senior year. During that semester the students participated in eight weeks of on-campus methods instruction followed by eight weeks of student teaching. The PCI form was administered as a pretest during the fifth week of on-campus methods instruction. It was administered again, as a posttest, during the last week of student teaching. A modified version of the form was developed to assess student teachers' perceptions of their cooperating teachers' pupil control attitudes. This form was administered during the final week of student teaching.

> "The difference between the student teacher's pretest PCI Form score and the Perceived PCI Form score of her cooperating teacher was used as a measure of the pressure felt by the student teacher to bring her views on pupil control into agreement with those of the cooperating teacher..."

Subsequently, the relationship between socialization pressure and the change in student teacher pupil control ideology was examined.

Findings

1. A t-test showed changes in student teachers' PCI that were significant beyond the .01 level from pretest to posttest. The direction of mean change on a humanism-custodialism scale was toward custodialism.

2. A Kolmogorov-Smirnov test for differences between the student teaching grade levels and change in PCI did not permit rejection of the null hypothesis.

3. An analysis of variance showed a relationship between perceived teacher PCI and change in student teacher PCI that was significant beyond the .05 level.

4. An analysis of variance showed a relationship between mean socialization pressure and change in student teacher

PCI that was significant beyond the .01 level. (Mean socialization pressure was defined as the difference between the student teacher's PCI pretest score and the perceived PCI score of the cooperating teacher).

5. An analysis of variance showed a relationship between mean change of student teacher PCI in classrooms of low, medium, and high socialization pressure that was significant beyond the .01 level.

A Scheffé test applied to the data indicated significant differences beyond the .01 level between the low and medium socialization pressure groups and between the low and high socialization pressure groups. No significant differences were shown between the medium and high socialization pressure groups on change in PCI.

Interpretations

The investigators found that the student teachers' views on pupil control changed in a custodial direction during student teaching. They also found a significant relationship between the change in PCI and the differences existing between the student teacher's own views and his perception of the cooperating teacher's views. They report that "...there was an increase in the mean change in PCI at each progressively higher level of socialization pressure." In further interpretation however, they say "the mean change in PCI at medium and high levels (of socialization pressure) was not significant. This lack of significance indicates that there may be an upper limit to the socialization pressure which can bring about change in views."

The investigators interpret their findings to show the significance of the socialization process for persons entering the teaching profession. "If a humanistic pupil control ideology is desirable, care must be taken to select cooperating teachers whose attitudes and beliefs toward pupil control are consistent with this desired ideology."

ABSTRACTOR'S ANALYSIS

This research study was timely and significant. It investigated an issue of concern and interest in teacher education, and it provides an appropriate base for a number of related studies needed. While the investigators have clearly found significant differences in important variables, the description of the research at times lacked precision and clarity, and it provided a very limited discussion of implications.

The study showed the significant effects of the attitudes of cooperating teachers on student teacher attitudes regarding pupil control, but it did not provide any direct insight regarding means to solve the problem that was observed. Can cooperating teacher models be developed that are more compatible with optimal pupil control ideology? Can teacher education programs provide an instructional treatment that will facilitate more optimal and lasting pupil control ideology? What is an optimal pupil control ideology? While these questions were not clarified in the study, they are very appropriate for subsequent investigation.

The comments that follow relate primarily to the clarity of the written report:

1. The rationale for the study included the assumption that successful implementation of new curricula demands careful consideration of the changes new strategies require in the teacher pupil relationship. The correct use of an inquiry approach in science teaching has "the effect of reducing the status difference between student and teacher." What is meant by "an inquiry approach to instruction"? Inquiry teaching has been defined in a variety of ways by different writers, and references should have been cited that elaborated on and supported the authors' position. Furthermore, what evidence did the authors have to support their assumption that inquiry teaching is related to low pupil control ideology, i.e., of a humanistic orientation? What evidence supported the authors' contention that "most teacher education programs approach the issue of pupil control from a liberal or humanistic point of view..."?

2. More information regarding the nature of the PCI instruments would have been particularly helpful. While a reference was cited, the reader was provided with no comments on reliability or validity. Insufficient data were available to permit interpretation of the scores that were reported in tabular form. Have any studies to date examined the relationship between paper and pencil PCI scores and actual classroom behavior? Pre- and posttest means and standard deviations were reported for the PCI Forms, but no such data were reported for the Perceived PCI Forms for cooperating teachers. The reader was told that the PCI Form means were 40.7 and 45.7 on the pretest and the posttest respectively, but he also knew that it was a 20-item instrument. How were the scores determined? What was the meaning of the 5 point difference?

3. Very inadequate information was provided regarding the nature of the sample of student teachers. In what kind of university were they students? How did their scores relate to national norms? In what kinds of school environments did they do their student teaching during the time of the study?

4. What was the nature of the methods instruction that preceded the pretest? Was there any particular treatment designed to develop a commitment to inquiry teaching or to a humanistic PCI? Further research studies should

investigate the effects of alternative treatments upon pupil control ideology and related measures. Comparison with this study will lack meaning, however, since the treatments were not described.

5. The presentation of the results of the study omitted some important details beyond those previously mentioned (such as perceived PCI scores). In the discussion of results relating to hypothesis 3 (page 317), for example, the narrative and table indicated rejection of the null hypothesis, but no comments or data are provided showing the direction of influence caused by the perceived PCI of cooperating teachers. The reader of a research report should not have to surmise this kind of information.

6. In general, the implications of the study were not discussed as thoroughly as seems justified by the interesting results. Specific suggestions could have been directed to the nature of methods instruction, to techniques for long term attitude change, and to the development of cooperating teacher models. A variety of significant questions for further investigation could have been elaborated, such as the need to investigate the effects of alternative treatments upon teacher attitude and subsequent teaching style and behavior.

7. While the results of this study permitted some powerful conclusions, the authors' concluding remarks provided a restatement of some initial assumptions that were <u>not</u> evaluated in this particular study. These remarks were written in a manner that could readily cause readers to perceive them erroneously as conclusions supported by the study. The paper concludes, for example, with the following final sentence:

    > If a teacher feels that both status and authority are threatened unless students are rigidly controlled, the probability of the use of curricular materials emphasizing inquiry is greatly reduced.

    This concluding statement is an assumption, and this research study provided <u>no</u> evidence to support or refute it.

REFERENCES

1. Horowitz, Meyer. "Student-Teaching Experiences and Attitudes of Student Teachers," <u>Journal of Teacher Education</u>, Vol. 19, No. 3:317-324, 1968.

2. Willower, Donald J., Terry L. Eidell, and Wayne K. Hoy. <u>The School and Pupil Control Ideology</u>, The Pennsylvania State Studies No. 24, Pennsylvania State University, University Park, 1967.

COGNITIVE DEVELOPMENT

EJ 079 849

Allen, Leslie R., "An Examination of the Ability of Third Grade Children from the Science Curriculum Improvement Study to Identify Experimental Variables and to Recognize Change." Science Education, Vol. 57, No. 2:135-151, April/June, 1973.
Descriptors--*Cognitive Development, Curriculum, Educational Research, *Elementary School Science, *Instruction, *Longitudinal Studies, Science Course Improvement Project, *Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Julian R. Brandou, Michigan State University.

Purpose

The purpose of this study was to compare performances on certain tasks between two groups of third grade children. One group was provided with SCIS science (Physical Science section) and the other group with a non-SCIS curriculum.

Rationale

This study is part of a longitudinal evaluation of the SCIS program. Allen refers to the 1967 Teacher's Guide for a statement of the goals of SCIS. The author has attempted to establish, by means of a factor analysis of the test items, the "cognitive" category for the tasks involved (2). The population of the study likewise is discussed in the other papers in detail (4) and depended on a matching of the students in socio-economic groups in SCIS and non-SCIS Honolulu schools.

Research Design and Procedure

The basic design is a posttest only, randomized within population, matched comparison. The sample size was 176, representing a loss from the original grade one group of 300. Fifty subjects were randomly drawn from the SCIS population for each socio-economic level. The procedure was repeated in matched non-SCIS schools for all first graders. Full details of the distribution remaining by sex, SCIS/non-SCIS and class are given. The children were near the end of their third grade program when tested in May/June, 1971. Protocols for each test item were generated and items were individually administered after pilot testing with grade 3 children at the University of Hawaii Laboratory School. Responses were recorded and grouped by variable, using procedures not described in the article reviewed. The reported numeral represents the number of times a given variable was mentioned by the children in each socio-economic group in the SCIS and non-SCIS schools. Analysis of covariance for the responses to each of five items is reported with interactions using the California Test of Mental Maturity as covariant. T-tests on the difference observed among socio-economic groups were done for the entire sample.

Findings

The five items were on two objectives:

1. Identifying variables that may affect a simple mechanical system, and

2. Recognizing change during an experiment.

Items 1-4 dealt with objective one and question 5 dealt with the second objectives.

TABLE OF ITEMS AND RESULTS OF ANALYSIS OF COVARIANCE WITH SIGNIFICANT DIFFERENCES NOTED

		Programs	Sex	Socio-Eco
Item 1	Eyedropper	X		X
Item 2	Floating syringe	X		X
Item 3	Sling shot	X	X	X
Item 4	Loaded Truck	X	X	
Item 5	Lighting Candle	X	X	X

T-tests show that in all cases the upper socio-economic group out-performed both the average and lower groups. On items 3 and 5 the average group also out-performed the lower group. No interaction effects were found.

Interpretations

"Using (a) the input variables of science program membership, sex, socio-economic status and California Test of Mental Maturity scores and (b) the criterion measures (items) previously described, Honolulu third grade SCIS children appear superior to non-SCIS children both in their ability to identify experimental variables, and to recognize change."

ABSTRACTOR'S ANALYSIS

Haven't the objectives of the third grade SCIS program changed since 1967? Perhaps Honolulu is using the materials as produced by D. C. Heath-Raytheon or perhaps the differences are trivial and teachers using the current Rand-McNally materials can expect the same results. What other objectives are listed by subsequent work as appropriate for this unit? Does the author expect any transfer to these tasks?

The study carefully delineates the protocols for the items but information regarding teacher reaction, pupil response, class size and the effects on other children in the current classes would add much "color" to the paper. What does the author think about the srray of responses given to each item? Of special interest are the observations made by a single pupil. Some of these are particularly astute, others seem so apparent one wonders about why other pupils don't see them. Does the author have any qualitative "feel" for the responses? Were there any expected values? How different are the socio-economic groups? Does a slight variation in parental income, for example, produce a marked difference in response frequency, fluidity, sentence length, use of adjectives or other observable characteristics?

The lack of program X socio-economic interaction seems an interesting area to explore. Is this an artifact based on the covariant adjustment? How tightly are the California Test of Mental Maturity scores correlated with socio-economic class? If the SCIS program can succeed with all socio-economic classes, this would be an important finding.

REFERENCES

1. Allen, Leslie R., "An Evaluation of Certain Cognitive Aspects of the Material Objects Unit of the SCIS Elementary Science Program." Journal of Research in Science Teaching, Vol. 7, No. 4:277-281, 1970.

2. Allen, Leslie R., "An Evaluation of Children's Performance on Certain Cognitive, Affective, and Motivational Aspects of the Systems and Sub-systems Unit of the Science Curriculum Improvement Study Elementary Science Program." Journal of Research in Science Teaching, In Press.

3. Allen, Leslie R., "An Examination of the Ability of First Graders from the Science Curriculum Improvement Study Program to Describe an Object by its Properties." Science Education, Vol. 55, No. 1:61-67, 1971.

4. Interaction. Teacher's Guide. Science Curriculum Improvement Study. Preliminary edition. Boston: D. C. Heath and Company, 1967, p. 2.

EJ 026 496

Nicodemus, Robert B., "Content and Skill Hierarchies in Elementary Science: An Analysis of ESS Small Things." Journal of Research in Science Teaching, Vol. 7, No. 3:173-177, 1970.
Descriptors--Behavior, *Behavioral Objectives, Complexity Level, *Content Analysis, *Elementary School Science, *Science Course Improvement Project, Scienrific Concepts, *Skill Development

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

Purpose

The basic problems investigated are assumed (by the abstractor) to be related to the following questions:

1. How do science content and the subtle and elusive behaviors of students and teachers relate to the inquiry philosophy of the Elementary Science Study (ESS) in the unit titled "Small Things"?

2. Is there any structure in "Small Things" which would assist a teacher to effectively plan a lesson so as to be able to assist the learner in ways consistent with ESS philosophy?

Rationale

The author views the main activity of science to be inquiry experienced through content. He is concerned, therefore, that important outcomes of some of the new science programs (assumed by the abstractor to mean ESS specifically) are frequently described as attitudes. He implies that such an approach is inadequare and that in order for teachers to communicate inquiry to students there is need for some means of measuring and describing inquiry through its relationships with content and behaviors.

This study is related to two other publications by the same author. The most pertinent is titled An Evaluation of Elementary Science Study as Science--A Process Approach (ERIC REPORT ED 027 217, September, 1968).

Research Design and Procedure

The "Small Things" unit was analyzed using a scheme adapted from the Science--A Process Approach (S-APA) program. The abstractor's summary of procedural steps follows:

1. The Teacher's Guide for Small Things was examined to determine expected student behaviors (objectives).

2. Student behaviors were rewritten using S-APA action words (i.e., identifying, distinguishing, constructing, describing, etc.).

3. Student behaviors were arranged in a vertical hierarchy analogous to the sequential ordering of objectives in the S-APA exercises.

4. Student behaviors were analyzed to determine the major concept (content) area under which each behavior could be categorized.

5. Student behaviors were grouped into levels to establish a general process hierarchy. The relationships among the levels are similar to those among the 13 different processes in S-APA (i.e., observing, measuring, classifying, etc.). Thus, while there is a general hierarchical sequence among the levels, the individual student behaviors are related in more of a non-sequential or horizontal arrangement than resulted from the vertical sequencing in step 3, above.

Findings

1. Forty-five student behaviors were identified from the Teacher's Guide for Small Things. Those 45 behaviors do not include the first ten which represent skills for using a microscope. (All 55 behaviors are listed in the text of the article.)

2. Three content concept areas were identified and each of the 45 behaviors was categorized under the appropriate area--(1) pond life, (2) cells, and (3) living.

3. Seven levels of a general process hierarchy were established and each of the 45 behaviors was listed under the proper level. The seven levels were (1) Identifying, Naming, (2) Describing properties, (3) Describing similarities and differences, (4) Ordering, (5) Applying a rule, (6) Interpreting relationships, and (7) Demonstrating validity.

Interpretations

The author states that the analysis of the ESS unit using the S-APA model provides a structure which may be put to a number of uses. Among them he includes:

1. The sequential development of content may be examined.

2. Behaviors may be reorganized into a General Hierarchy through which the development of skills may be studied.

3. The greatest value may be in its use by teachers in planning and teaching a unit.

    a. The child's knowledge and skills may be evaluated in relation to a hierarchy.

    b. The instructional program may be planned to avoid unnecessary repetition but at the same time assure the possession of behaviors necessary for more complex tasks.

    c. The relation of content behaviors and skills may be readily ascertained and organized in ways useful to the immediate and numerous decisions which constitute part of a strategy of teaching.

ABSTRACTOR'S ANALYSIS

Relationships to other studies. This investigation's relationship to the literature on the "cluster of reviews" topic, cognitive development, is limited primarily to the psychology of Robert Gagné. Gagné's theories on the sequential development of intelligence and the need for structuring pathways to the student's "terminal capabilities" have been popularized in his books, including The Conditions of Learning, and numerous articles such as "The Learning Requirements for Enquiry" (Journal of Research in Science Teaching, 1:144-153, 1962).

The present author makes reference to several other published works which are assumed to be related to the reviewed article. The relationships are vague, however, and tend to give the impression that the current report is an abridged version of a somewhat longer and more complete manuscript.

In terms of cognitive theory, this investigation holds that children must be able to recognize essential features of an objective or solution to a problem before it is presented and that the resulting recognition is an important source of reinforcement. Thus, the study supports the establishment of hierarchies of objectives, content, skills, and processes as a means of providing logical planning, teaching and learning sequences.

The science processes and objectives in S-APA are organized according to an elaborate hierarchial scheme (a reflection of Gagné's strong influence on the program's development). It is understandable, therefore, that the author of this study selected the organizational structure of S-APA as a model for analysis of ESS "Small Things."

Contributions of the study. This investigation's methodological approach represents its major contribution. It is a good example of the application of a specific model (S-APA) derived from general psychological theory (Gagné's) to the task of analyzing a teaching-learning system. Such a technique is welcome since there is a great need in education for specific analytical models.

Of course, precise knowledge on teaching and learning is very limited. Thus, the contributions of this specific model (or any other for that matter) are largely matters of value judgments. Examination of the present study, for example, by one person who supports the S-APA "structured" approach and by another who supports the ESS "messing about" philosophy would surely result in two very different reactions. It is most critical, therefore, that a researcher or consumer of research examine the underlying assumptions of any specific model used in analyzing a teaching-learning approach.

Weaknesses of the study. The major weakness of the investigation lies in the written report itself. Most critical is rhe arricle's failure to communicate in specific terms exactly what problems or questions were investigated. The questions stated under Purpose above represent the abstractor's attempt, after careful study of the article, to isolate and define the research questions. It was chiefly through the description of procedure and the summary, that assumptions were made about the study's purpose. Furthermore, since no specific problems or questions were explicitly defined by the author, it is difficult to judge the validity of the remainder of the paper--procedures, findings, and interpretations.

Other communication problems appear to result from editorial or typographical errors. There is reference to charts (2 and 3), for example, which were apparently omitted from the final copy. Only their captions are found in the text.

This report could have been improved considerably by:

1. An explicit statement of purpose (problems, questions, or hypotheses) early in the paper.

2. Restricting procedural steps to those required to achieve the investigation's purpose.

3. Stating explicit findings relative to the investigation's purpose.

4. Making interpretations based on rhe findings of the study.

5. Editing the manuscript to assure its ability to communicate (i.e., complete thoughts, references that have meaning, appropriateness and inclusion of charts, etc.).

Current and future research in the area. A great deal of research has been done on ESS and S-APA as separate programs. However, the analysis of one program (in this case, ESS) in terms of the psychological bases of another (S-APA) is a relatively rare occurrence.

Current social and economic conditions have dampened interest in and support for science education. It would seem that there is merit, however, in examining present science curricula in terms of different theories of cognitive development. Such an approach might reveal a great deal about the science programs and about cognitive development as well.

EJ 065 313

Raven, Ronald J., "The Development of the Concept of Acceleration in Elementary School Children." Journal of Research in Science Teaching, Vol. 9, No. 3:201-206, 1972.

Descriptors--*Cognitive Development, *Concept Formation, *Conservation (Concept), Elementary School Science, *Kinetics, *Learning Theories, Physical Sciences, Scientific Concepts

Expanded Abstract and Analysis Prepared Especially for I.S.E. by David H. Ost, California State College, Bakersfield.

Purpose

To ascertain the developmental relationships among the speed and the acceleration concepts described by Piaget, the investigators tested the hypotheses that:

1a. the probability of success was the same for each of seven tasks; or,

1b. at least two of the tasks were different; and,

2a. the true probability of success on the tasks was the same for each grade level; or,

2b. at least two of the grade groups were different.

Rationale

Elementary school science curricula frequently include learning experiences which deal with acceleration. Raven's position is that insufficient attention is given the logical relationships of velocity, time, and distance. It is pointed out that Piaget used four stages to describe the development of these relationships.

Stage I, the child sees acceleration only as a rapid change without consideration to a continual change in speed.

Stage II, the child may have an intuitive feel for acceleration but may not perceive the gradation from "slowest" to "fastest" (six years).

Stage III, the child begins to correlate the distance-time relationships (7-8 years) and subsequently (9-10 years) can compare distances or times traveled between measured intervals.

Stage IV, the child is able to compare the relationships between the increases in the distance between equal intervals of time and the increase in speed over equal intervals of time (11 years).

Raven concludes that the child must have attained reversibility of centration of successive speed relationships as well as decentration of distance-time relationships in order to deal appropriately with acceleration. In addition the child must have acquired conservation of uniform speeds. This suggests that there are several aspects related to the acquisition of the concept of acceleration, not the least of which is the attainment of measurement operations and the achievement of conservation of time and speed.

Research Design and Procedure

Eight different concept tasks are reported to have been administered to a sample of 96 third, fourth, fifth, and sixth grade children. The questions and responses were oral. An incorrect answer or an inadequate explanation of the answer were considered as not "passing" responses. The tasks were administered in random order to each child to control for learning. The titles of the tasks were:

1. Speed task: simultaneous-synchronous time with different distances;

2. Speed task: successive speeds travelling equal distances in unequal times;

3. Speed task: conservation of uniform speed;

4. Acceleration task: successive independent comparisons;

5. Acceleration task: seriate speeds, distance intervals equal;

6. Acceleration task: seriate time, distance intervals equal; and,

7. Acceleration task: seriate distance, time intervals equal.

Appropriate descriptions of the protocol utilized for each task are given in detail.

The null hypothesis (1a) was tested by applying the Cochran Q test. Twelve children at each grade level were selected from the 96 children participating in the study. The hypothesis of no difference (1b) was tested by the application of the post hoc comparison method described by Marascuila and McSweeney.[1]

To ascertain whether or not the probability for success was the same for each grade (hypothesis 2a) for the remaining 48 children in the sample, the Friedman two-way analysis of variance was used.

---

[1]Marasculia, Leonard and Maryellen McSweeney, "Nonparametric Post Hoc Comparison for Trend," Psychological Bulletin, 67:401-412, 1967.

The post hoc comparisons was again employed to determine the degree of difference between age groups (hypothesis 2b).

A scalogram[2] analysis was performed on the seven tasks.

Findings

The data are easily summarized in Table I. All four null hypotheses are rejected.

A reproducibility coefficient of .86 for the seven tasks was found using the scalogram analysis. No other results of the analysis were teported.

Table I

Percent of Children Attaining a Positive Score (N = 96)

	Age (Years)			
Task	8	9	10	11
1	80	93	87	93
2	7	26	67	87
3	0	20	53	80
4	67	80	87	93
5	53	73	93	93
6	7	13	53	73
7	0	7	40	67

Interpretations

Raven provides a detailed discussion of the data and develops several interpretations and conclusions. The results are interpreted with great care to show that the speed and acceleration tasks that require intensive logical operations (comparison of the whole with a part) could be handled by third and fourth grade children. Tasks which employ extensive logical operations are appropriate for fifth and sixth grade children. Such operations involve the comparison of conceptual parts with one another, thus allowing the child to compare

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

speeds of movements in succession. Apparently the conservation of speed task 3 occurs prior to the acquisition of the acceleration tasks 6 and 7.

ABSTRACTOR'S ANALYSIS

In general, the reported study appears to be tightly designed and carefully thought through. The interpretations drawn appear to be consistent with the data resulting from the investigation as well as with related studies. The position, held by Piaget, that conservation of time and speed is a prerequisite to being able to compare units of speed and time as associated with acceleration, seems to be solidly supported.

The implications for science curriculum developers seem straight forward. Learning activities related to growth rates of plants or populations, evolution, changes in the speeds of wagons, cars or space missiles, as well as other physical phenomena related to acceleration, should be developed only after careful consideration is given to the development of the logical relationships of velocity, time, and distance. This notion supports in part the hierarchal nature of learning ã la Gagné but more specifically gives credibility to Piaget's assertion that conservation is a necessary condition for all rational activity. The acquisition of extensive logical operations is preceded by, if not dependent on, the ability to carry out intensive operations.

The report seems to lack in details in two areas. First, although the data presented are based upon the results of seven concept tasks, the investigator writes that "eight different concept tasks were administered to all of the children." Nowhere in the discussion is it explained what happened to the eighth task and associated results. The reader can speculate that it was a fourth task related to speed and was found to be redundant. (Or perhaps it was an acceleration task that produced inconsistent data.)

A second area where completeness is lacking is in the absence of information in the results section related to the scalogram analysis. Conclusions are discussed which are based upon this analysis but no data are provided other than the reproducibility coefficient of .86 provided in the discussion section. Interpretations seem to be drawn from this one figure without elaboration of the reasoning.

It is not completely clear to this abstractor how one demonstrates the developmental relationships of subconcepts to a more sophisticated concept without including significant biases. The development and selection of tasks seem to be obvious locations of biases. The assumption among the majority of Piagetian studies, such as this one, seems to be that there is intrinsic objectivity in the tasks. Is it not possible for researchers to select and employ tasks which by their very nature support the premise under investigation? Such research might provide evidence of relationship but does not negate the possibility of independence. Thus, although it could be concluded that the tasks which require logical operations are related, this does not provide any more than indirect evidence that the concepts are interdependent. It would seem that the results of relationship type studies can be determined a priori. Could it not simply be that the number of attributes a task incorporates is the basis for a difference of responses and not the relationship of those attributes?

These comments are not directed specifically at the study reported by Raven. In fact, his interpretations are very conservative and couched in noncommital terms. Nevertheless it would seem that some mechanism for randomly selecting tasks that reportedly evaluate specific cognitive stages or concepts could be employed to correct for problems associated with internal validity.

EJ 023 090

Raven, Ronald J., "The Effects of a Structured Learning Sequence on Second and Third Grade Children's Classification Achievement." Journal of Research in Science Teaching, Vol. 7, No. 2:153-160, 1970.

*Academic Achievement, *Classification, *Elementary School Science, Evaluation, *Instruction, Task Performance

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Ann C. Howe, Syracuse University.

Purpose

The investigation was undertaken to determine whether the use of a set of instructional materials based on classification rules would facilitate the acquisition of classification abilities by second and third grade children.

Rationale

The conceptual framework for the study was Piagetian theory as it has been extended by the search for training methods which will promote the development of logical thinking. When knowledge of and interest in Piaget's work developed in this country many attempts were made to train children to perform tasks of logical thinking. Many of the early attempts used direct reinforced training methods and were notably unsuccessful. At the time this investigation was carried out, a few successful training experiments had been reported and investigators were turning to training procedures based on the presu ed logical prerequisites for the task in question. Thus, in the present study, logical operations described by Piaget and Inhelder were rewritten as 12 classification rules and were then used as the basis for instruction and testing. Pictorial representation and verbal rules were used as the method of instruction because of promising results obtained by other investigators who had used similar techniques. The study was grounded in theory and took into account recent work in the field.

Design and Procedure

1. Instructional Materials. Instructional materials consisted of exercise booklets which contained teaching frames followed by transfer frames. The upper part of each page (frame) contained a pictorial representation and a verbal expression of a classification rule; the lower half contained instructions and a problem to be solved. Teaching frames contained shaded and unshaded geometric shapes in two sizes; the transfer frames contained drawings of such familiar objeccs as plants and animals.

The 12 classification rules were divided into five rules for simple grouping (rules 1-5) and seven rules for hierarchical classification operations (rules 6-12).

2. Testing Materials. Three items were written for each of the 12 rules and presented in a format similar to that used for the instructional materials. The upper half of the page contained a solved problem and the lower half presented the problem to be solved. No words were used on the test pages; instructions were given orally. Presumably, one point was scored for each item. A reliability coefficient of .86 is reported.

3. Sample. All the chlddren enrolled in the second and third grades of the public schools in a small city participated in the study. There were 210 second grade children and 175 third grade children from seven classes at each level.

4. Design. Children in each class were assigned at random to Group 1, Group 2, or the Control Group. This resulted in the formation of six groups, three at each grade level. Achievement and I.Q. test scores are presented to show the equivalence of the three groups at each level.

After assignment to groups, all children worked in exercise booklets as part of the regular classroom schedule for 15 minutes per day for 24 days. Children in Group 1 received booklets containing instruction on all 12 rules; those in Group 2 received booklets containing instruction only on rules 6-12 but the total number of frames was the same. Children in the Control Group were given drawing and coloring activities. Classroom teachers supervised the work of the children.

One or two days after the completion of the exercises the test was administered to all the children by persons other than the classroom teachers.

The design of the experiment included random assignment of subjects at each grade level to two treatment groups and one control group, followed by instruction and a posttest carried out as part of the regular classroom routine.

5. Analysis of Data. The dependent variable was the score on the 36-item posttest. These data are presented as total mean score for each group and as subtest mean scores for each group. (The three test items for one classification rule make up a subtest.) The group mean scores were subjected to multivariate analysis for treatment, grade, and class and were used to make planned comparisons of control vs. experimental groups and of Group 1 vs. Group 2. The simple grouping subtest scores (items 1-5) and hierarchial classification subtest scores (items 6-12) for each group were subjected to multivariate analysis of covariance using first one (subtest scores 1-5) and then the other (subtest scores 6-12) as the covariate.

Findings

The principal findings were as follows:

1. Children in the experimental groups scored higher than those in the control groups on the posttest ($p < .01$).

2. Third grade children scored higher than second grade children on the posttest ($p < .01$).

3. Within grade levels, no significant class differences were found.

4. The two experimental groups scored at the same level on the posttest.

5. The two experimental groups scored at the same level on subtest items 6-12 (hierarchial classification).

6. Children in Group 2 scored higher than children in the Control Group on subtest item 1-5 (simple grouping) ($p < .01$).

Interpretations

The main conclusion was that the program of instruction did enhance the children's classification test performance. The two procedures used in the study were equally effective, both with respect to total test performance and performance on subtests. Third grade children performed at a higher level than did second grade children, particularly on the hierarchical classification tasks. The results of the analysis of covariance, in which items 1-5 (simple grouping) and items 6-12 (hierarchical classification) were separated, were taken to mean that practice on the simple grouping tasks did not facilitate performance on the hierarchical tasks, but that practice on the hierarchical tasks retroactively facilitated performance on the simple grouping tasks.

ABSTRACTOR'S ANALYSIS

This investigation contributes to the continuing effort to translate Piagetian theory into classroom practice. Piaget described the course of the development of logical thinking but left it to others to make what practical use they could of his work. When psychologists and educators learned that children did not understand the concept of conservation, they set about to teach conservation. When their efforts did not have the desired results, conservation and the mental processes leading up to it took on the aura of a mystery; the idea that conservation and, by implication, all other Piaget tasks, could not be taught became accepted as revealed truth. In the study which is being considered here the focus was not on a specific task but, instead, on the underlying logical operations. Growth in classification ability (i.e., logical thinking) was measured by an increase in test scores rather than by a shift in status from classifier to non-classifier. The approach adopted in this study takes away the mystery surrounding Piagetian theory and puts it into the practical world of the classroom teacher who is familiar with exercise books containing examples and problems to be solved.

In order for the results to be convincing we have to believe that the test measured only growth in classification ability and that the experimental groups did not have an advantage gained through familiarity with the format of the test pages and practice in solving similar problems. There is no way for this to be determined and no mention is made of it in the article, but there is evidence from the subtest scores which is somewhat reassuring on this point. If the Control Group scores were uniformly lower than the Group 1 or Group 2 scores, this difference might be attributed to lack of familiarity with the test format, but, in fact, the difference between Control Group scores and experimental group scores varies from .04 to 1.89 (of a possible 3.0). So, while the possibility of a reactive effect exists, it is unlikely that it was a significant factor.

It is not clear how the multivariate analysis of covariance was used in comparing subtest scores of the experimental groups. Apparently, subtest scores for rules 1-5 were used as covariates for comparing subtest scores of rules 6-12, and subtest scores of rules 6-12 were used as covariates for comparing subtest scores of rules 1-5. But all the subtest scores were presumably affected by the instructional program and, therefore, could not be used as covariates in comparing treatment effects. The results of this analysis were interpreted as showing that practice on the more complex rules enhanced learning of the simpler rules and that practice on all the rules produced no higher test performance than practice on only the complex rules. The results of this analysis would carry more weight if the logic of the statistical procedure were made explicit.

This is one of a series of investigations in which materials programmed to facilitate the acquisition of logical operations have been developed and subsequently tested in classrooms. The implications of this investigation are best understood when viewed as part of a larger research project.

EJ 034 667

Allen, Leslie R., "An Evaluation of Certain Cognitive Aspects of the Material Objects Unit of the Science Curriculum Improvement Study Elementary Science Program." Journal of Research in Stience Teaching, Vol. 7, No. 4:277-281, 1970.

Descriptors--*Achievement, *Curriculum, *Elementary School Science, *Evaluation, Grade 1, Student Evaluation, Student Behavior

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Chester E. Raun, Temple University.

Purpose

The investigator identifies the purpose of this study as an effort to determine whether those first grade students who participated in the SCIS unit "Material Objects" demonstrated a performance, on certain SCIS objectives, at a level superior to first grade students who were non-participants in the SCIS program. No specific testable hypotheses related to cognitive aspects are stated.

Rationale

"Material Objects" is described as a unit which teaches the first grade child the fundamental concepts of (a) objects and their properties and (b) material. It is indicated that the study is part of a larger study involving two other elementary science programs and extending, eventually, through grade six. The investigator makes the statement that "this and subsequent studies will produce information that can be used in making curricular decisions." Previous research, identified by the investigator, to which this study may be related was carried out on the "Material Objects" unit by Hagen (9), and Haan (8), with Allen (1) investigating classification skills of SCIS children at grades two, three, and six.

Research Design and Procedure

The investigator indicated that it was impossible to carry out randomization of subjects into groups, treatments to groups, and teachers to treatments because classes to be selected for study were already in existence. Therefore, the Static Group Comparison Design, as described by Campbell and Stanley (7), was selected for use. Reading level, socioeconomic status, sex, and type of science program were identified as the independent variables while six selected objectives from the "Material Objects" unit became the dependent variables. The six objective selected and the procedures for measurement were:

1. To describe an object by its properties: The child was given a soda cracker and asked to state as many properties of the object as possible without any imposed time limit.

2. To group objects by material: A collection of objects each made of a single material (rubber, glass, plastic, metal, wood, paper, cloth) was placed before the child. The child was asked to group the objects by material and state why.

3. To order objects serially by a stated property: The child was presented with a random arrangement of blocks of wood of different lengths, thicknesses, and widths, to one side of which was glued sandpaper of different textures and colors. The child was asked to arrange the blocks in order in as many ways as possible and to state reasons.

4. Simple inference: The child was presented with a paper clip, a plastic button, a magnet, and an eraser, and asked to play with them. A closed matchbox (containing a paper clip) was added to the set and the child was told that another one of the objects from the set was inside the box. The child was asked to name the hidden object and give his reason.

5. To group objects by other than visual means: This item was designed to measure grouping performance when visual clues were absent. Because the matchboxes used were visually identical, the children were forced to group using criteria such as weight of the boxes and whether or not the contents rattled.

6. To group objects by visual means (Resemblance Sorting): The child was presented with 35 drawings of geometric and non-geometric shapes which could be matched by size, shape, pattern, or by each combination of these attributes.

The first five of the criterion measures were administered individually while the last was a group test. All test items were scored on the basis of one point for each acceptable response.

Among the independent variables only the type of science program is clearly delineated, i.e., "SCIS" or texts such as "Concepts in Science." The variable of sex does not indicate how many girls or boys were involved in the experimental or control groups nor is the variable of reading level delineated as to ability level, source and means of assessing. The variable of socioeconomic status is only indicated as upper-average-lower without specific parameters or definitions of each level.

The study was conducted during the 1968-69 school year at the University of Hawaii utilizing first grade children in the Honolulu schools. From those schools participating in the SCIS program (7 schools, 21 classrooms) a random sample of 50 subjects was drawn for each of the three socioeconomic levels for a total of 150 children in the experimental group. With I.Q.'s unavailable for first grade children, each of the SCIS schools was matched by a non-SCIS school of comparable socioeconomic status. The above selection procedure was repeated for non-SCIS children resulting in 150 children in the control group. It was indicated that a majority of the 300 subjects were of Oriental and Polynesian ancestry with a minority of Caucasians but no racial percentages are reported.

The exact duration of the study or the exact time of conducting the study is not clear though the investigator does indicate that the control group "had spent a year with conventional, hard-cover, commercial elementary school science texts such as 'Concepts in Science'." This would lead the reader to assume that both groups had received equal treatment over time and that the study was conducted either at the end of grade one or at the beginning of grade two.

Data collected by use of the criterion measures were summarized in a separate table for each criterion measure, indicating the total number of acceptable responses and the mean for each group (experimental and control) and for each socioeconomic status level. Each table was then analyzed by the parametric technique of multiple regression. Although not specifically stated in the narrative it appears, from a review of the summary table of multiple regression analysis, that a probability level of less than 0.10 was considered significant.

Findings

For the sample of this study the results for criterion test items 1, 3, 5, and 6 indicate that the SCIS group produced a greater number of acceptable responses than did the non-SCIS group. For criterion test item 2 the non-SCIS group accumulated a greater number of errors in sorting and had a greater number of objects not grouped by material than did the SCIS group.

A summary of multiple regression analysis for the criterion test items 1, 2, 3, 5, and 6 produced several F values with a low probability that the results were attributable to chance. Similar F values were also reported for items 1 and 5 with regard to the variable of science program; for items 2 and 6 with regard to the variables of socioeconomic status and reading; and for item 3 with regard to the variables of socioeconomic status, reading, and science program. The variable of sex produced no significant effect.

A more fundamental question of interest to the investigator, i.e., "What percentage of the total variance is explained by the full equation and for which items is this explained variance statistically significant?" found another table showing the R^2 and R values for the full equation, and the R^2 values for the restricted equations. (R^2 = the amount of predicted variance; R = the relationship between variables). The findings from this analysis indicated the relationship between variables to be statistically significant in items 1, 2, 3, and 6, but only in items 3 (Serial Ordering) and 6 (Resemblance Sorting) did the amount of explained variance exceed 10 percent. Of the latter items, only item 3 involved a significant decrease in predictive power when the SCIS program was omitted.

Criterion test item 4 (Simple Inference) produced a negligible number of successful solutions (1 SCIS and 3 non-SCIS) and was not included in the analysis.

Interpretation

It is concluded, as a result of the study, that first grade SCIS children are superior to non-SCIS children only in their ability to serially order objects. As a result of chi-square calculations it was further concluded that in four out of the five test items in which objects were presented for examination and manipulation (Items 1, 2, 3, and 4) the SCIS children exhibited significantly more exploratory behavior than did non-SCIS children. In general it appears that differences are minimal with respect to successful item responses. The investigator suggests that superior exploratory behavior may be all that can realistically be expected after only one year of exposure to this (SCIS) program. He further suggests that should annual increments of change occur, and if these changes are additive, there may be a cumulative effect of the program measureable at grade six.

ABSTRACTOR'S ANALYSIS

There have been, in recent years, numerous comparative studies between the "newer" elementary science curriculums, e.g., ESS, SAPA, and SCIS, and the more conventional, textbook-centered science curriculums. A limited number of these have been research studies concerning the child's cognitive development as a result of the treatment effects of the "newer" curriculums. This number is further reduced if one considers only those related to the SCIS program.

In addition to the studies cited by the investigator, one finds that Thier (17) investigated first graders' understanding of the concept of matter using the "Material Objects" unit for the specific purpose of revising the unit when lack of understanding was identified as being significant. Raven (15) used SCIS as a vehicle to investigate concept development. The purpose of the study was to determine the developmental sequence necessary for understanding momentum. The results supported Piaget's findings that children understand concepts about matter before they understand concepts about speed.

Stafford and Renner (16) studied the effect of instruction, based on the "Material Objects" unit, in accelerating concept skills attainment in the area of conservation. The experimental group of first grade children showed a greater growth in each of the six conservation areas tested (i.e., conservation of number, length, liquid amount, solid amount, weight, and area). Conflicting results are reported by Neuman (14) when exploring the effect of instruction in selected experiences, taken from the "Material Objects" unit, on first grade pupils attainment of conservation of weight and quantity. Following 18 weeks of instruction no differences were observed between the experimental and control groups, on pretest and posttest scores, as measured by adapted Piagetian conservation tasks.

Weber (18) conducted a study to determine the level to which SCIS develops the learner's ability to use selected science processes. Experimental and control groups of fifth grade pupils were compared regarding their performance on a series of science process-oriented tasks. The results showed the SCIS group to be clearly superior in regard to each of the six processes studied (i.e., observation, classification, measurement, experimentation, interpretation, and prediction). It was also noted that SCIS pupils were more aggressive in their approach to problems and were more diverse, persistent, and creative in their experimental designs. The aggressiveness of the fifth grade pupils would appear to be consistent with the superior exploratory behavior for first grade pupils reported by Allen.

Considering the influence of science on other academic areas Kellogg (10) explored the effect of instruction in the "Material Objects" unit on reading readiness of first grade pupils. Thirty-two pupils in the control group experienced a commercial reading readiness program, while thirty-seven pupils in the experimental group used the "Material Objects" unit instead of a reading readiness program. Both groups were pre- and posttested on the Metropolitan Reading Readiness Test with an intervening treatment period of six weeks. Students in the experimental group showed greater gains in five of the six subtest areas. It was concluded that to best teach reading, first teach thinking, as represented by conservation reasoning.

Other studies have been concerned with investigating teacher behavior and the SCIS program. Wilson (19) discovered that SCIS teachers used a significantly higher level of questioning than non-SCIS teachers. This finding was also supported in independent investigations by Bruce (6) and Moon (12). A conflicting report by Moriber (13) indicates that SCIS teachers were consistently asking lower order, stimulus-response type of questions despite summer program preparation. The latter study raises serious questions about what is taught and how it is taught in teacher preparation programs as well as what actually occurs when teachers teach the program in the classroom.

Kondo (11) focused on an analysis of the relationship between questioning behavior of the teacher and types of SCIS lessons (i.e., exploration, invention) as found in the "Material Objects" unit. The results indicated that the way a lesson is approached (i.e., teacher demonstration, children manipulating materials) has a greater influence on the type of question the teacher asks than the type of lesson.

Though there appears to be a dearth of very recent studies reporting on the impact of the SCIS program, all of the above studies have direct or indirect relationship to the present study. However, in all of the studies it is difficult, if not impossible, to synthesize results because extensive and detailed information is not provided about the actual approaches being compared. A comparative study without such information lacks external validity. The present study is no exception.

The reader is left to assume that all of the experimental group received equal treatment. What evidence might have been presented which would insure equal treatment? What evidence exists that insures adequate teacher preparation and presentation for equal treatment? At what specific point in time was testing conducted? What, precisely, was different in treatment for the experimental and control groups? Answers to such questions, though requiring additional input, could be presented in a concise manner and would begin to provide some of the information necessary to establishing external validity.

The investigator refers to "certain cognitive aspects" in the title of the report but does not clearly delineate these cognitive aspects in the stated design of the study. The reader must assume that these are the selected objectives from the "Material Objects" unit. It would appear that these are reducible to the process skills of classifying and inferring since they involve ordering, describing, and comparing. The question however remains. Are these the cognitive aspects the investigator refers to? In a similar vein, how does the investigator establish the parameters of socioeconomic status and what are these parameters? How many children were in each of the socio-economic levels? Were boys and girls equally distributed in each level?

Although it was apparently not the intent of the investigator to analyze data other than that relevant to the avowed purpose of the study, there are considerable data presented, particularly regarding the variables of socioeconomic status and reading, which are not addressed in interpretations. Conversely, there is reference to chi-square calculations in the conclusions which are not addressed in discussion of analysis or findings. It seems appropriate to make reference to all reported findings, if only to indicate interpretations at a later date and certainly to substantiate all conclusions with necessary findings.

Though it is relatively easy to assume the role of the armchair tactician, it is felt that attention to details, such as those raised by the comments and questions above, would strengthen this, and other studies, enhancing in turn their contribution to the science education literature.

Coupled with this is a need for improvement in comparative studies. Most such studies, as indicated earlier, have considerable potential for improvement in terms of descriptive information regarding treatment variables. Anderson (5) has appropriately discussed the need for such information. The need for descriptive information is not for statistical purposes but to fully describe what is being compared. Despite impressive statistical data unless one knows what is being compared, knowledge about which approach is better is of little value.

Aside from the need for descriptive information there is also a need for studies which follow through on a given population over a period of years. It is refreshing to note that Allen has not only indicated a series of studies would be conducted with the same population over a period of time but has actually reported subsequent studies (2, 3, 4). The concept of longitudinal studies is not new but few such studies are executed though they are desperately needed to establish the worth and continued use of any curriculum and/or methodology. It is doubtful, however, that this study and others have produced results which "can be used in making curricular decisions."

REFERENCES

1. Allen, L. R., "An Examination of the Visual Classificatory Ability of Children Who Have Been Exposed to One of the 'New' Elementary Science Programs," Science Education, 52:432-439, 1968.

2. Allen, L. R., "An Examination of the Ability of First Graders From the Science Curriculum Improvement Study Program to Describe an Object by its Properties," Science Education, 55:61-67, 1971.

3. Allen, L. R., "An Evaluation of Children's Performance on Certain Cognitive, Affective, and Motivational Aspects of the Interaction Unit of the Science Curriculum Improvement Study Elementary Science Program," Journal of Research in Science Teaching, 9:167-173, 1972.

4. Allen, L. R., "An Evaluation of Children's Performance on Certain Cognitive, Affective, and Motivational Aspects of the Systems and Subsystems Unit of the Science Curriculum Improvement Study Elementary Science Program," Journal of Research in Science Teaching, 10:125-134, 1973.

5. Anderson, R. D., "So It's Significant, But What Is It?" NARST Newsletter, 17 (4):1-2, December, 1975.

6. Bruce, L. R., "A Determination of the Relationships Among SCIS Teachers' Personality Traits, Attitude Toward Teacher-Pupil Relationship, Understanding of Science Process Skills and Question Types." University Microfilms, Ann Arbor, Michigan, 1969.

7. Campbell, D. T. and J. C. Stanley. "Experimental and Quasi-Experimental Designs for Research." Chapter in Handbook of Research in Teaching, N. L. Gage (ed.). New York: Rand McNally & Company, 1963.

8. Haan, N. An Exploratory Investigation of the Effects of an Initial Experience with SCIS's Material Objects Unit on First Grade Children and Their Teachers. Berkeley, California: Far West Laboratory for Educational Research and Development, 1968.

9. Hagen, E. Science Curriculum Improvement Study Evaluation Development: Material Objects Unit. Berkeley, California: Science Curriculum Improvement Study, University of California, Spring 1966.

10. Kellogg, D. H., "An Investigation of the Effect of the Science Curriculum Improvement Study's First Year Unit, Material Objects, on Gains in Reading Readiness." Unpublished doctoral dissertation, University of Oklahoma, 1971.

11. Kondo, A. K., "The Questioning Behavior of Teachers in the Science Curriculum Improvement Study Teaching." Paper presented at the 42nd annual meeting of the National Association for Research in Science Teaching, Pasadena, California, February, 1969.

12. Moon, T. C., "A Study of Verbal Behavior Patterns in Primary Grade Classrooms During Science Activities." Unpublished doctoral dissertation, Michigan State University, 1969.

13. Moriber, G., "Levels of Questions Asked by Elementary School Teachers Involved in Teaching SCIS Material." Paper presented at the 48th annual meeting of the National Association for Research in Science Teaching, Los Angeles, California, March, 1975.

14. Neuman, D. B., "The Influence of Selected Science Experiences on the Attainment of Concrete Operations by First Grade Children." Paper presented at the 42nd annual meeting of the National Association for Research in Science Teaching, Pasadena, California, February, 1969.

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

16. Stafford, C. G. and J. W. Renner, "SCIS Helps the First Grader to Use Logic in Problem Solving," School Science and Mathematics, 7:159-164, February, 1971.

17. Thier, H. D., "A Look at a First Grader's Understanding of Matter," Journal of Research in Science Teaching, 3:84-89, March, 1965.

18. Weber, M. C., "The Influence of the Science Curriculum Improvement Study on the Learner's Operational Utilization of Science Processes." Unpublished doctoral dissertation, University of Oklahoma, 1971.

19. Wilson, J. H., "Differences Between the Inquiry-Discovery and the Traditional Approaches to Teaching Science in the Elementary School." Unpublished doctoral dissertation, University of Oklahoma, 1967.