DOCUMENT RESUME

ED 147 185 SE 023 365

AUTHOR Helgeson, Stanley L., Ed.; Blosser, Patricia E., Ed.
TITLE Investigations in Science Education, Vol. 2, No. 3. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 76
NOTE 67p.
AVAILABLE FROM Information Reference Center (ERIC/IRC), The Ohio State University, 1200 Chambers Rd., 3rd Floor, Columbus, Ohio 43212 (subscription $6.00, $1.75 ea.)

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Abstracts; *Concept Formation; *Educational Research; Fundamental Concepts; *Learning Theories; Research; *Research Methodology; Science Education; Teacher Attitudes; *Teacher Behavior
IDENTIFIERS Research Reports

ABSTRACT

This issue of Investigations in Science Education (ISE) provides analytical abstracts, prepared by science educators, of research reports in the areas of learning theories, concept learning, and teacher behaviors and attitudes. Each abstract includes bibliographical data, research design and procedure, purpose, research rationale, and an abstractor's analysis of the research. Abstracts are clustered by topics investigated. (SL)

ED147189



# INVESTIGATIONS IN SCIENCE EDUCATION



National Association for Research in Science Teaching

ERIC Center for Science, Mathematics,
and Environmental Education

Published Quarterly by

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

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

Volume 2, Number 3, 1976

NOTES . . .

from the Editor

This volume contains research reports grouped into three clusters. The first, LEARNING THEORIES, contains five analyses of research reports related to learning. The second group, CONCEPT LEARNING, includes three studies. Finally, there are four studies related to TEACHER BEHAVIORS AND ATTITUDES. These clusters are an attempt to review research with some common basis representing current trends in science education research.

Your publishable responses are invited.

Stanley L. Helgeson
Editor

Patricia E. Blosser
Associate Editor

LEARNING THEORIES

Koran, John J., Jr., and Mary Lou Koran. "Differential Response to Structure of Advance Organizers in Science Instruction." Journal of Research in Science Teaching, 10(4):347-353, 1973.

Descriptors--Academic Aptitude, *Cognitive Processes, *Conceptual Schemes, Educational Research, *Elementary School Science, *Learning Theories, *Programed Materials, Science Education, *Student Reaction

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Glen S. Aikenhead, University of Saskatchewan.

Purpose

The purpose of Koran and Koran's study was to investigate individual differences in learning when the instructional materials have been preceded by three kinds of introductory passages: (1) higher level generalizations with specific examples (an advance organizer), (2) higher level generalizations without examples (an advance organizer), and (3) a control passage.

> It was anticipated that the greater the degree of structure provided by the advance organizer, the more likely it would be to reduce the burden of semantic processing, and thus benefit those learners with less ability to develop an adequate conceptual scheme of their own for organizing new material. Consequently, in providing differing amounts of structure, performance ... was expected to be differentially related to IQ scores. (p. 349)

Rationale

Ausubel's learning theory predicts that "meaningful learning" (higher-order learning characterized by Bloom's hierarchical scheme) is enhanced by providing students ahead of time with an "advance organizer" (a sequenced piece of instructional material that is more general and more abstract than the ideas it precedes). The function of an advance organizer is to activate relevant concepts (called "subsuming concepts") which students have already learned and which form part of their own cognitive structure. Subsuming concepts are considered relevant if they can be useful in learning the ensuing information. Thus, an advance organizer should facilitate learning.

It follows then that advance organizers will, first of all, differ depending upon the content they precede (for example, familiar versus unfamiliar content). They will also differ depending upon the characteristics of the students interacting with the advance organizers.

The study by Koran and Koran is related to many investigations; those concerned with the effect of advance organizers, and those concerned with the interaction between different student characteristics and varying modes of instruction. The authors drew upon numerous empirical findings that suggest different students learn more, and with greater ease, with different kinds of instructional methods. While Koran and Koran accept the assumptions inherent in Ausubel's theory, there appears to be an important incongruity between the authors' and Ausubel's ideas concerning the function of an advance organizer. A discussion of this apparent difference of opinion appears at the end of the section "Abstractor's Analysis."

## Research Design and Procedure

Design: Students were randomly assigned to three treatment groups. The same achievement test was written on three occasions; pretest, immediate posttest, and delayed posttest (one week delay):

$$R \quad O_1 \quad \left\{ \begin{array}{c} T_A - O_2 - T_A \\ T_B - O_3 - T_B \\ T_C - O_4 - T_C \end{array} \right\} \quad O_5 \quad O_6$$

Treatment: The same programming booklets concerning insects (30 specific concepts dealing with insects) were studied by all three groups. However, the introductory passages differed for each group as follows:

Group A: a passage (advance organizer) describing the study of insects in terms of higher level generalizations along with specific examples using animals familiar to students.

Group B: a passage (advance organizer) describing the study of insects in terms of higher level generalizations without specific examples.

Group C: a control passage containing historical information with neither generalizations nor examples.

All students were given sufficient time to finish their work. The duration of the study was not reported.

Variables: Students' knowledge of the material on insects was assessed by: (1) a 30-item multiple-choice test ($O_1$, $O_5$, $O_6$) which had an internal consistency reliability of .78. Each item presented a name of a concept for which a correct example had to be chosen from three alternatives, and (2) the number of errors made during the study of the programmed material ($O_2$, $O_3$, $O_4$), called "program errors."

In this study, individual differences were defined in terms of student abilities, measured by IQ scores (Otis-Lennon Mental Abilities Test).

Sample: The sample consisted of 89 fourth-grade students enrolled in an upper middle class Houston, Texas, school district which was familiar to the investigators.

Analysis:

(1) "A 3 (treatments) x 3 (trials) repeated measures analysis of variance was used to test instructional treatment main effects for criterion test scores ... The assumption of homogeneity of regression underlying the use of the repeated measures model was tested and accepted." (p. 349) The trials were the pretest, immediate posttest and the delayed posttest.

(2) A one-way ANOVA used program errors made during the programmed instruction to detect differences among the three treatment groups.

(3) A simple linear regression related IQ scores and criterion variables (program errors, immediate posttest, and delayed posttest). "Aptitude x Treatment interactions were evaluated by comparing regression slopes obtained for each aptitude criterion pair in different treatments using F tests for heterogeneity of regression." (p. 351)

Findings:

(1) 3 x 3 ANOVA: "A significant treatment effect was found across trials, from pretest to posttests. However, between-group differences did not attain significance." (p. 349)

(2) One-way ANOVA: No significant differences among the treatment groups were found. "Thus, from the average data alone, the three treatments were about equally effective in terms of promoting retention and errors during instruction." (p. 350)

(3) Simple linear regression (aptitude x treatment interactions): IQ scores failed to interact with instructional treatment for either immediate or delayed posttest scores, but did produce a significant disordinal interaction with the instructional treatments for program errors. IQ scores were negatively related to program errors for groups B and C (those receiving generalizations alone or a control passage) while positively related to program errors for group A (those receiving generalizations and examples).

Interpretations

The three instructional treatments were about equally effective when assessed by program errors and posttest scores (immediate and delayed). This concurs with previous research, which is cited in the article. Koran and Koran thought that the null results might be explained by experimental conditions such as the unexpected familiarity of the content on insects and the generally high ability level of the sample. Also a rehearsal effect of the pretest on posttest scores could have conceivably reduced possible differences in posttest scores. However,

> The significant aptitude X treatment interaction between IQ scores and program errors tends to support the hypothesis that during instruction low ability subjects were more reliant on the organizing features of the program, whereas higher ability subjects were more effectively able to organize material in the absence of highly structured advance organizers. This finding appears to support a growing number of indications that in some cases highly structured treatment may actually be dysfunctional for high-ability subjects. (p. 352)

Such an interaction effect was not observed for the two posttest scores. The investigators suggested that the nature of the programmed materials themselves and the generally high ability level of the sample might possibly explain the null results. (High IQ scoring students may impose their own structure on the materials regardless of any type of advance organizer.)

Koran and Koran, in recognizing their very limited positive findings, mention that the learner's concept organizing operations may require "sharper definition" than is presently found in Ausubel's theory, before further research can be productive. The authors also recognize the need for greater clarity in aptitude measures, instructional procedures, and criterion measures.

## ABSTRACTOR'S ANALYSIS

Criticism of science education research studies usually focuses on three main issues [Hurd (2), Kempa (4), and Shulman and Tamir (7)]:

(1) lack of a theoretical framework or the lack of an extensive research base,

(2) methodological flaws in the design and implementation of the investigation, and

(3) lack of attention to problems in learning scientific knowledge in general, and specifically, lack of attention to individual characteristics of the student.

None of these points can apply to the study by Koran and Koran. In fact, their investigation exemplifies the type of research which critics usually describe as a paradigm to follow.

Koran and Koran's study logically progresses from past research studies. In particular, it relates two areas of great interest and concern: (1) the interaction of different modes of instruction with different characteristics of students, often labelled "aptitude-treatment interaction," and (2) Ausubel's construct "advance organizer" which has been one of the most researched aspects of his theory of learning. This study by Koran and Koran clearly demonstrates how research can take place in a context of established theory and empirical findings.

The reader who is new to science education research reports will benefit from noticing a number of other excellent qualities exemplified in Koran and Koran's report. The article is clear, concise, and precise. The authors remind the reader of the essential aspects of Ausubel's theory which are germane to their study. References are included for those readers who would like to acquaint themselves with Ausubel's theory or with aptitude-treatment interaction. Jargon is used only for precision and not for academic impact. (Some researchers tend to overuse erudite phrases steeped in jargon. However, others attain a clarity of expression with everyday Anglo-Saxon words and expressions.)

I would like to draw the reader's attention to the research design employed in this study. Its randomization of students to treatment groups provides a powerful technique for making valid inferences. The repeated measures analysis of variance is very appropriate and its assumption of homogeneity of regression is discussed. (A writer is well advised to report the major assumptions underlying the statistics he uses and to express the conformity of his data or design to these assumptions.)

The comparison of regression slopes using F tests for heterogeneity of regression is a sophisticated technique for detecting interactions. It has advantages over the more familiar two-way ANOVA. One of these advantages is using continuous data rather than arbitrarily defined discrete groupings of data necessary in most two-way ANOVAs. However, had Koran and Koran included a graph of these regressions, I think it would have clarified their results reported in tabular form.

The investigators use a treatment which is very easy for anyone to replicate because it is clearly defined and is in a written form. In contrast, some research reports are terribly vague about the treatment (for example, "CHEM Study" and "Traditional"). In addition, Koran and Koran minimize spurious treatment effects, such as the effect of different classroom teachers.

The authors are careful not to overgeneralize their findings. Their results are treated in the proper context: That the advance organizers preceded carefully sequenced programmed materials. It is left to further research to resolve what would happen when advance organizers are used with other types of science materials or presentations.

Because of the care with which the authors designed the study, carried it out, and analyzed the results, one can feel confident in the validity of their results.

Let me now turn my attention to some issues which are more related to the current state of research in the area of instruction. In particular, I shall examine three topics with respect to directions for further research: student characteristics, meaningful learning, and advance organizers.

<u>Student Characteristics</u>: It is certainly an understatment to say that the process of instruction and learning is a complex one. The interactive factors that affect student achievement include: Classroom

environment, teacher characteristics, home environment (for example, parental interest and expectations), and student characteristics. Student characteristics themselves encompass a plethora of categories; for example, interests, abilities, attitudes, needs, aptitudes, motivation. This list of interactive forces is far from complete. However, in its shortened form it still points out that the measurement of one or two student characteristics (sex, IQ scores, divergent and convergent thinking, analytic and global or intuitive problem solving, Sigel's cognitive styles, background knowledge of the specific topic, and reading skills, to name but a few) does not hold much promise in accounting for a significant amount of the variance in student achievement. This view is supported by the high instances of null findings reported in the literature, the present study included. [See for example, Herron et al., (1).] Not only does the definition and measurement of student characteristics require a great deal more study, but so does its relative influence compared with other interactive factors. For instance, if it were established that parental expectation accounts for more variance in student achievement than does student learning styles, then the implications for the improvement of instruction would be much different than if student styles were found to be more influential.

Meaningful Learning: Ausubel distinguishes between rote learning and meaningful learning by alluding to Bloom's taxonomy. Meaningful learning is considered to be higher-order learning in Blooms hierarchical scheme (for example, solving novel but relevant problems). Ausubel assumes that advance organizers will only facilitate higher-order learning. Therefore, achievement tests, which are meant to distinguish between control groups and groups having advance organizers, must assess higher-order learning. A study such as Koran and Koran's should ensure that the achievement tests meet this expectation and that lower-order learning, such as recognition, is not a major component of a test. Perhaps Koran and Koran did so, but it is not reported. Kahle and Rastovac (3) have published a procedure which helped them ensure that a full range of questions were used on achievement tests. (They, too, investigated the effect of advance organizers.) Null findings are almost inevitable if criterion measures are not well represented by questions assessing higher-order learning.

Advance Organizers: The concept of advance organizers is perhaps one of the most confused aspects of Ausubel's theory. Novak (5) has attempted to clarify this confusion. In order to emphasize the role played by the individual's previous knowledge (his cognitive structure with its network of subsuming concepts), Novak renamed advance organizers. He called them "cognitive bridges." A cognitive bridge links certain subsuming concepts to the material to be learned. "Cognitive bridges are short segments of learning material that provide guidance to the student as to which concepts in his cognitive structure might best be employed to learn meaningfully." (Novak 5:500)

However, by using Novak's clarification, I see some major problems for researchers. If Koran and Koran were to follow Novak's definition, they would have had to discover the relevant subsuming concepts possessed by their sample of students. Otherwise, one could not expect the advance organizer to have a significant effect on posttest scores. Given today's technology, it seems to be unreasonably difficult to link advance organizers with individual students' subsuming concepts.

Koran and Koran's notion of advance organizers appears to be quite different from Novak's. To Koran and Koran, "the function of the advance organizer is to provide structure of 'ideational scaffolding' (a quotation from Ausubel) for the incorporation and retention of material ... and to increase its discriminability." (p. 348) Therefore, Koran and Koran believe that advance organizers can differ with respect to their degree of structure. They argue that the more structure inherent in an advance organizer, the less the student need take responsibility for organizing the material in his own way. This notion of advance organizers related more to the act of information processing than to Novak's link between specific subsuming concepts already in the student's cognitive structure and concepts to be learned.* To Koran and Koran, the learner constructs his own relevant subsuming concepts, guided by the structure offered by an advance organizer.

> Because of some general familiarity with the content and the high ability level of the sample, many subjects would be expected to already possess relevant subsuming concepts to some extent, thus reducing the potential learning advantage of the advance organizers. (p. 350)

On the other hand, Novak seems to be saying that the relevant subsuming concepts are already constructed in the learner's cognitive structure, and they only need to be cued into action by the advance organizer. The reader is presented with two distinct incongruous views. Unfortunately, these differences cannot be clarified by Ausubel's original work. Therefore, before further research can be very illuminating, investigators must form a concensus concerning the function of advance organizers. Koran and Koran came to a similar conclusion: "As Cronbach and Snow have suggested, the whole concept of organizing operations by the learner may require sharper definition before theorizing and research can proceed fruitfully." (p. 352)

There could be another problem. If advance organizers do not work effectively when students possess the relevant subsuming concepts, and if students usually do possess the relevant subsuming concepts to some degree (as Koran and Koran's study suggests), then the effectiveness of an advance organizer in normal classroom situations would seem to be minimal. Consequently, one might not expect useful results to emerge from research into the effectiveness of advance organizers.

## REFERENCES


1. Herron, J. Dudley, et al. "A Summary of Research in Science Education." Science Education (Special Issue), 1974.


---

*Thus, according to Koran and Koran, different styles in processing information should favor different types of advance organizers. Shmurak (1974) assessed the cognitive styles of some students and constructed different advance organizers to match their different cognitive styles. However, she found that the "non-organizer" was as effective as the advance organizer.

2. Hurd, Paul D. "Research in Science Education: Planning for the Future." Journal of Research in Science Teaching, 8:243-249, 1971.

3. Kahle, Jane B., and John J. Rustovac. "The Effect of a Series of Advanced Organizers in Increasing Meaningful Learning." Science Education, 60:365-371, 1976.

4. Kempa, R. E. "Science Education Research: Some Thoughts and Observations." Studies in Science Education, 3:97-105, 1976.

5. Novak, Joseph D. "Understanding the Learning Process and Effectiveness of Teaching Methods in the Classroom, Laboratory, and Field." Science Education, 60:493-512, 1976.

6. Shmurak, Carole B. "The Effect of Varying the Cognitive Style of Advance Organizers on Learning of Expository Science Material by Eighth Graders." Dissertation Abstracts, 35:2075A, 1974.

7. Shulman, L. S., and P. Tamir. "Research on Teaching in the Natural Sciences." In Second Handbook of Research in Teaching, R. M. W. Travers (ed). Chicago: Rand McNally Publishing Co. 1973.

Noús, Albert and Ronald Raven. "The Effects of a Structured Learning Sequence on Children's Correlative Thinking about Biological Phenomena." Journal of Research in Science Teaching, 10(3):251-255, 1973.

Descriptors--*Academic Achievement, *Biology, Cognitive Processes, *Instructional Materials, Intermediate Grades, *Learning Theories, Science Education, Secondary Grades, *Thought Processes

Expanded Abstract and Analysis Prepared Especially for I.S.E. by William R. Brown, Old Dominion University, Norfolk, Virginia.

## Purpose

The purpose of this investigation was to determine if correlative thinking in fifth, seventh, and ninth grade children could be facilitated by a Piaget-based structured learning sequence. If subjects who have acquired formal operations are presented with the operations of logical multiplication, class inclusion, equivalence, and reciprocal exclusion, then they should be able to solve correlation tasks.

## Rationale

The findings of Inhelder and Piaget suggest that the acquisition of correlation operations is dependent upon the development of the following operations: logical multiplication, class inclusion, equivalence, and reciprocal exclusion. A study by Smedslund concerning the concept of correlation in adults along with the work of Inhelder and Piaget suggest that the use of a training program whose design implemented these logical structures would enhance the acquisition of the correlation operations. Biological principles such as structure and function, predator and prey, camouflage and predation, and mimicry and survival are examples of correlative-based rules found in elementary school and junior high school science programs.

## Research Design and Procedure

The independent variable was a set of structured training exercises in four parts. (1) The logical multiplication exercises required the students to construct four sets of object combinations from 2x2 tables. (2) The class inclusion operations involved the construction of the marginal elements in the 2x2 tables that had completed cells. (3) The equivalence and reciprocal exclusion operations used tasks whose solutions demanded the construction of binary relationships of the diagonal cells. (4) The correlation tasks used problems in which the equivalence and reciprocal exclusion cases were summed and compared to determine the event diagonals of greatest frequency. This enabled students to state the rule.

The three-hour training program used a deductive-generalization mode of instruction. The program and test were used in two pilot studies.

The dependent variable was achievement on correlational thinking tasks. A rationale was presented for a treatment-posttest only design ( x o).

The posttest consisted of six correlation problems. Each of the problems was presented in 4x5 cell pictorial format. KR-20 coefficients ranged from .80 to .86.

The subjects were 119 fifth grade, 90 seventh grade, and 156 ninth grade students from three parochial schools. The samples were all the students at these grade levels.

All students received the same training and testing sequence. The relative efficiency of the training program was determined by comparing the achievements of fifth graders (concrete operations level), seventh graders (transition level), and ninth graders (formal operations level). Multivariate analysis of variance was used to indicate significant difference in achievement among the grades.

## Findings

An F of 4.21, $p < .001$, indicates a significant difference in achievement among the grades.

An analysis of the mastery of the correlation problem in the training sequence showed the following percentages of achievement: Grade 5 - 30 percent, Grade 7 - 29 percent, and Grade 9 - 48 percent. All of the groups had difficulty in those tasks that exhibited no correlation. Ninth grade students were relatively successful in forming rules, but less successful in forming operations on the pair combinations using equivalence and reciprocal exclusions.

## Interpretations

The results of this study support those of Smedslund and those of Inhelder and Piaget. Most students in the early formal operations stage are not able to use correlation operations to construct rules from data. Students in the early formal operations stage did benefit from the training. The biological principles involving correlative-based rules should be presented to young children in such a manner that correlative operations are not required.

## ABSTRACTOR'S ANALYSIS

Beard refers to "relations between relations" as one aspect of formal operations (1:125-133). Several studies are cited dealing with the operations which are part of this study. Beard states that "many adults do not attain the level of formal operations except in some limited areas; if they are neither well educated nor of good intelligence they may hardly reach it at all." "The problem for the teacher is to use teaching methods in such a way as to maximize development of formal thinking whenever it is possible."

Beard's comments have several implications in connection with this study. First of all, even though the investigations reported the mean age of the students per grade level, it was assumed that grade level was an accurate descriptor of Piagetian level of development. Were all ninth grade students at the formal operations level, specifically for the relations between relations set of variables? Perhaps assessment of stage of development should have been considered prior to inclusion of all students in the population sample.

A second factor to be considered is the validity of the criterion test and the training program. No validity indices are reported.

A third issue is the time sequence of the study. The criterion tests were administered the day after the students completed the training program. It would be interesting to assess the long-term effects of the training program. Does the program produce long-term results with wide applications to numerous correlative problems?

If concept development is considered as a developmental phenomena dependent on both Piagetian stages of development and experiences, then several implications can be suggested. From a research viewpoint, task analysis type dissections of curricula should be conducted in order to assess what types of mental operations are necessary in order to successfully handle specific bits of the curriculum. This "need" is also indicated as a top priority for research in science education by NARST members (2:163). The instructional bits could then be arranged in a sequential order that would complement levels of cognitive development. Teachers and administrators at all levels must be encouraged to become involved and stay involved in this type of fundamental research.

From a curriculum and instruction viewpoint, educators might approach curriculum decision-making from a goal-oriented attack rather than from a subject matter orientation. For example, a curriculum decision might be to develop correlative thinking in adolescents. The selection of instructional strategies and materials would be based on this stated goal and the implications of research rather than on the selection of topics from specific disciplines such as science. Implications from the Nous and Raven study indicate that certain adolescents can benefit from instructional strategies that develop correlative thinking. Topics such as the complementarity of structure and function might be included for older adolescents. The placement of this topic would be based on appropriateness in attaining the curriculum goal of developing correlative thinking. Numerous activities should be provided for youngsters to develop correlative thinking such as the use of a two dimensional matrix board where numerous attributes may be placed and interrelated (3:224-225).

Relations between relations is a critical factor in a comprehensive understanding of science as field of inquiry. The development of appropriate teaching strategies, the selection of content, and the assessment of both these factors must be encouraged.

## REFERENCES


1. Beard, Ruth. An Outline of Piaget's Developmental Psychology for Students and Teachers. New York: Basic Books, Inc., 1969.

2. Butts, David, et al. "Priorities for Research in Science Education," paper presented at NARST-1977. Abstract in "Abstracts of Presented Papers, NARST-1977," NARST, ERIC/SMEAC, Columbus, Ohio, 1977.

3. Furth, Hans and Harry Wachs. Thinking Goes to School Piaget's Theory in Practice. New York: Oxford University Press, 1975.

Raven, Ronald and Herbert Strubing. "Intrafactor Transfer in Second Grade Children." Science Education, 55(1):31-38, 1971.

Descriptors--*Achievement, *Elementary School Science, *Instruction, Learning, Physical Sciences, Scientific Concepts, *Transfer of Training, Visual Perception

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

## Purpose

The general purpose of this study is to determine if second grade children's learning from the Science Curriculum Improvement Study's unit Relativity (sic) of Position and Motion can be significantly improved by prior training with one of two units selected from the Frostig Program for the Development of Visual Perception.

## Rationale

There has been an increasing interest in the use of intrafactor transfer in classroom learning. Most educators have discarded the viewpoint that learning is very specific and is limited to the kind of task that is exercised. Contrary to a position taken by Thorndike, Guilford suggests that there is growing evidence of intrafactor transfer which shows that learning need not be narrow or completely specific.

Certain science curriculum projects, such as A.A.A.S. Science--A Process Approach, are designed on the premise that transfer of training can take place in the context of the classroom environment. This position is supported by the work of Gagne and Paradise and Anderson.

One of the assumptions upon which the strategy of curriculum development rests is that there will be transfer of knowledge and skills within and among learning sequences. The purpose of the SCIS unit Relativity (sic) of Position and Motion is to develop frames of reference for describing the position of a system. Piaget's studies of operational coordination suggest that the majority of second grade students would have difficulty achieving on the science relativity unit.

Cronbach has differentiated between "substantive transfer" and "aptitudinal transfer," in that substantive transfer refers to common elements and aptitudinal transfer refers to the prepareation to learn from an environment that may have no substance in common with that which is taught.

## Research Design and Procedure

The sample subjects utilized in this study comprised the students of all nine second-grade classes in all of the public schools in the city of Dunkirk, New York. The subjects were heterogenous with regard to race, nationality and socioeconomic group. The mean chronological age was eight years and two months, and the mean IQ (Kuhlmann-Anderson, 7th edition) was 106.5.

Each class was divided into three randomly assigned experimental groups. The first treatment for each of the three groups is found below:

Group I: Frostig Spatial Relationship unit
Group II: Frostig Visual-Motor Coordination unit
Group III (control): "Outlined pictures to color"

The treatments were provided for fifteen days. The teachers of the subjects monitored the activities and restructured the room arrangment to prevent contamination of the learning situations in the various groups. A second treatment, the science relativity unit, was undertaken by all groups (I-III) for twelve days. The teachers were given a set of lesson plans that described what was to be covered, how it was to be covered, and how much time was to be allowed for each activity.

All students were administered the science relativity unit achievement test three days after the completion of the instruction on the science relativity unit. A pilot study was undertaken for the purpose of establishing the reliability for the test, and it was determined to be 0.80 (Kuder-Richardson Formula #20). The Metropolitan Achievement Test scores and the Kuhlmann-Anderson Intelligence test scores were also obtained for each pupil from the schools' records.

The basic experimental design is diagrammed below:

$R_{I}$ $X_1$ $X_4$ $0_1$

$R_{II}$ $X_2$ $X_4$ $0_1$

$R_{III}$ $X_3$ $X_4$ $0_1$

The smallest number of subjects in any one group was seven. The computer program for the statistical test used in this study demanded an orthogonal design which requires that the number of observations in each possible combination of treatments is the same. The size of the comparison groups was equalized at seven by random procedures. This resulted in a total sample of 189 subjects (three groups x nine classes x seven subjects = 189). A two-way analysis of variance (mixed effects model) for treatment and class was performed. Although subjects were assigned to treatment groups by random procedures, IQ and achievement test scores were compared to determine if a systematic bias did exist.

## Findings

Based on the analyses of the data, it was found that the two treatment groups (Frostig Spatial Relationship group and Frostig Visual-Motor Coordination group) did significantly better (at the .01 level) than the control group ("outlined pictures to color"). No difference was found between the two experimental groups (I and II). These results show that a proactive facilitation occurred between the Frostig units and the science unit.

Interpretations

Substantive transfer of the specific content taught by the Frostig Perception of Spatial Relationships unit enhanced learning on the science unit. Aptitudinal transfer of the general content taught by the Visual-Motor unit also enhanced learning on the science unit. One of the reasons for this may be that there are some skills that are treated in the Motor Coordination unit which are distantly related to the science unit.

There were significant differences (.05) found between the classes, but the authors administered the science achievement test to minimize class differences. The results of this study must be tempered to some extent by the fact that the mean IQ and Metropolitan Achievement Test scores of the control group were slightly less than the corresponding scores of the experimental groups.

The results of this study seem to indicate that the type of perceptual training provided may not be the most significant factor in achieving on the relativity science unit. Perhaps the most important aspect of prior perceptual training is its ability to bring into focus for the child the elements of perception in general.

## ABSTRACTOR'S ANALYSIS

The content and nature of this study obviously address concerns basic to the total education process, such as the potential for the transfer of learning and the lack of necessity for narrow and specific learning tasks. Also, as a secondary purpose, it helps to establish validity for a specific program said to develop visual perception. The authors of this research report utilize the SCIS materials as a part of the project, which probably attracts a particular audience of readers that may not normally seek out studies on learning and intellectual development.

The review of the literature was found to be central to the study, and a reader would have difficulty in understanding the results and implications of this research unless it had been thoroughly conceptualized. This abstractor had some minor concerns about the title of the SCIS unit as listed in the report, "Relativity of Position and Motion." The current SCIS unit of this nature is "Relative Position and Motion," and one might presume this to be a preliminary version, although no mention is made of this specific title even as early as 1968 in the SCIS Elementary Science Sourcebook. This problem seems very tangential to this study and warrants no further comments.

Since Campbell and Stanley's works are often used in developing research designs, the references made to this material by the authors is commendable. The paper itself was written in a format that is easy to follow and is free from unrelated information.

The random assignment of students within a class to one of the three treatment groups should be highly regarded. A more usual practice is the assignment of a class to one treatment, thus in this case there would be a total sample size of nine, and a cell size of three for each treatment group.

In reviewing the data which were in the original article but not in this abstract, one cannot help but note that the control group had the lowest mean score of all three groups on the measures of IQ and achievement. The authors set this aside by quoting Campbell and Stanley on randomization as an "acceptable guarantee against one group having more or less ability than another group." One wonders if the determination of the correlation of age, IQ and general achievement with the dependent variable in this study ("science achievement") might not provide some useful insights. Otherwise, the research design and statistical treatment of the data are appropriate within the limitations imposed by the nature of the sample.

In the description of the subjects it was stated, "the students represented all races, nationalities, and socioeconomic groups commonly called 'culturally deprived'." One must find great difficulty in the interpretation of this statement, particularly when attempting to generalize to another locality. Does this statement mean that all races are in the city of Dunkirk, New York, or does it mean that all races that are commonly called culturally deprived are in the city of Dunkirk, New York?

Within more recent years, the findings of Piaget's works are an important part of research in science education. This study predates much of this work, but certainly should be noted and built upon in the future. It gives credence to both current methodology and curricula in science education. The researchers, Raven and Strubing, have recognized important findings in their study and have attempted to cite appropriate limitations. Of equal importance is the merging of the works of Thorndike, Guilford, Gagne, and Piaget, among others, into a study that attempts to provide a basis for teaching and curriculum development in science education.

Raven, R. J., and H. Polanski. "Relationships Among Piaget's Logical Operations, Science Content Comprehension, Critical Thinking, and Creativity." Science Education, 58(4):531-544, 1974.
Descriptors--*Cognitive Tests, *Comprehension, *Concept Formation, Creative Ability, Critical Thinking, *Educational Research, Learning Theories, *Measurement, Models, Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Anton E. Lawson, Arizona State University at Tempe.

Purpose

The purpose of this study was to describe some relationships between a Piaget-based model of content comprehension and other types of cognitive processes that have been related to comprehension. Two questions were asked:

1. Is there a positive relationship between science content comprehension, creativity, critical thinking, and Piaget's logical operations?

2. Do children's science content comprehension, creativity, critical thinking, and logical operations differ between fourth grade and sixth grade?

Rationale

The major thesis of this study is that the student himself must impose some sort of restructuring of the science content before he can meaningfully comprehend it. This notion stems from Piaget's epistemological position that to know something is to act upon it. Presumably, in order to restructure science content--that is to comprehend it-- the Piagetian logical operations involved in classification, seriation, logical multiplication, compensation, proportion, probability, and correlation are used. If these operations are lacking or poorly developed, then the ability to restructure science content will also be lacking or poorly developed. Therefore, science content will be poorly comprehended.

The present study is, in effect, an attempt to gain some empirical support for this thesis. Raven and Polanski call this thesis the "logical operations science content comprehension development model."

Research Design and Procedure

Support is sought for this thesis through a study of the intercorrelations of five criterion measures. The five measures were:

1. Science Content Comprehension Test - developed by Raven and Polanski;

2. Torrence Tests of Creative Thinking - verbal and figural tests;

3. Cornell Critical Thinking Tests - conditional reasoning and class reasoning;

4. Iowa Test of Basic Skills - vocabulary and reading comprehension;

5. Raven's Test of Logical Operations - developed by Raven.

The tests were administered to 111 fourth grade students and 109 sixth grade students of slightly above average IQ. The socio-economic level of the sample was described as ranging from low middle-class to high middle-class. Five testing situations were set up for each grade. Each testing situation lasted approximately one and a half hours with a break midway through the session. One testing session was held per week for each grade. The investigators administered the tests.

## Findings

1. Test reliabilities determined by Hoyt's analysis of variance were 0.85 for the Raven's Test of Logical Operations; 0.82 for the Science Content Comprehension Test; and 0.89 for the Iowa Test of Basic Skills (comprehension). The scoring validity of the Torrence Tests of Creative Thinking were determined by factor analysis. Since the analysis yielded results consistent with the tests' constructs, it was concluded that the tests had adequate reliability.

2. The Science Content Comprehension Test correlated: at 0.62 with the Raven's Test of Logical Operations; at 0.69 with the Iowa reading comprehension test; at 0.34 with the class reasoning test; at 0.42 with the conditional reasoning test; at 0.32 with the Torrence verbal test; and at 0.14 with the Torrence figural test.

3. The means for the criterion tests (and subtests) generally increased from fourth grade to sixth grade. Univariate F ratios comparing mean scores reached significance ($p < .001$) for 14 of the 19 subtests.

4. Multiple correlation coefficients between the subtests of the Raven's Test of Logical Operations and several of the other measures were reported. The reported coefficients ranged from 0.63 for the Science Content Comprehension Test (40 percent of variance accounted for), to 0.23 for the Torrence figural test (18 percent of variance accounted for).

5. A factor analysis of several of the measures was reported. A factor identified as a "comprehension" factor accounted for 13 percent of the variance. The classification and logical multiplication subtests of the Raven's Test of Logical Operations showed moderate loadings on this factor.

Interpretations

The authors interpret the high correlation (0.69) between the Science Content Comprehension Test and the Iowa reading comprehension test to indicate that the student is using related operations in answering questions on both tests. The high correlation (0.59) between the Science Content Comprehension Test and the Raven's Test of Logical Operations suggests that it is important to consider the role of these logical operations in science content comprehension.

The finding that the classification subtest of the Raven's Test of Logical Operations accounted for a substantial portion of the variance of the class reasoning test was expected since the class reasoning process involves the grouping of objects or events and the construction of new group relationships. The positive relationship between the probability subtest of the Raven's Test of Logical Operations and conditional reasoning test was explained by the assertion that the conditional reasoning process uses words such as "if" which have a probabilistic characteristic.

Low intercorrelations between tests, such as the figural creativity test and the Science Content Comprehension Test (0.14), and the figural creativity test and the Iowas reading comprehension test (0.15) were accounted for by the difference in test formats.

The results of the factor analysis, which showed two subtests of the Raven's Test of Logical Operations loading on a factor identified as a "comprehension," were interpreted as support for the argument that the process of restructuring of given information is common to a variety of comprehension operations.

The finding that the sixth graders performed better than the fourth graders on most of the criterion measure subtests was consistent with the authors' expectations based upon their interpretation of Piaget's theory.

The authors conclude that their findings strongly support the validity and usefulness of the logical operations science content comprehension developmental model. This conclusion was drawn since the model was used to design the Science Comprehension Content Test and this test was found to correlate significantly with the other measures used in the study.

## ABSTRACTOR'S ANALYSIS

Raven and Polanski offer an interesting thesis. Following Piaget, they are hypothesizing that an adequate comprehension of science content requires the learner to impose some sort of restructuring upon that content. This restructuring presumably involves mental operations such as classification, seriation, logical multiplication, proportions, correlations, and so on.

Unfortunately, one cannot be sure just what Raven and Polanski mean by "restructuring" since no examples of such test items were given. The research report would have been more informative if examples had been given. Nevertheless, I suspect that they mean that the student must in some way be able to "operate" with the content for that content to be adequately comprehended. In a sense, adequate comprehension is being defined as "operative knowldege" as opposed to static or "figurative" knowledge in the Piagetian sense. If the Piagetian operations required for operative knowledge are lacking or poorly developed, comprehension will suffer. For example, to understand the concept of biological succession, the mental operations involved in serial ordering are needed to place plant and animal types into a serial arrangement through time. Or, to understand the quantitative relationship between the strength of attraction of molecular particles and their inter-particle distance, the mental operations involved in proportional reasoning are needed.

This general thesis seems to me to be of great significance. We have long suffered from the problem of teaching content that, for some reason or another, simply was not adequately comprehended. Piagetian theory provides an hypothesis to explain this lack of comprehension (cf. Lawson and Renner, 5). Further, necessary steps are suggested to correct this difficulty. Namely, design instruction to explicitly teach these operations [e.g., Raven (9); Lawson and Wollman (6)]. The result of such instruction would be students armed with the mental operations necessary for science content comprehension.

Notice that this model of science content comprehension necessarily precludes the teaching of certain theoretical concepts to young elementary school students. These students, who generally are just beginning to develop proficiency with concrete operations such as classification and seriation, are a long way from developing proficiency with formal operations such as proportions, correlations, combinations, probability, and so on. According to this view of content comprehension, these formal operations would be needed to meaningfully comprehend certain theoretical concepts, presumably because our reasons for believing in the validity of such concepts depends upon analyses of data using these operations. If students have no facility with these formal operations, they would have no way of comprehending the nature of such concepts and our basis for belief in such concepts. Their knowledge would have to be based upon faith, rather than upon evidence and upon reason. Also, their knowledge would be static (figurative), rather than operative. Further, the premature teaching of such concepts most likely would not result in the development of formal operations.

It should be noted that not all science educators would agree with this restriction upon the teaching of theoretical concepts. Novak (7, 8), for instance, uses Ausubel's theory as a basis for his argument that children can "acquire" theoretical concepts well before they reach Piaget's stage of formal operations. Novak, however, fails to explain just how Ausubel's theory can be used to justify such teaching and why Ausubel himself acknowledges the necessity for formal operations in such learning situations (e.g., Ausubel, 2:149; Ausubel, 3:279; Ausubel, 4:261; Ausubel, 1:219-220).

Having said this, let us return to the present investigation. Just how much support do the data give to the Raven-Polanski thesis? In my judgment, they lend some support, but not much. As the authors suggest, the strong relationship (0.62) between the science Content Comprehension Test and the Raven's Test of Logical Operations could be accounted for simply by the fact that both tests were designed to require the use of the same logical operations and not because the operations are necessary for science comprehension.

The authors go on to state that the strong association (0.58) between the Iowa reading comprehension test and the Raven's Test of Logical Operations negates this argument since the construction of the Iowa reading comprehension test did not use the logical operations model in its development. This argument, however, seems in error since the authors themselves report that 20 percent of the items on the Iowa reading comprehension test did involve the logical operations found on the Science Content Comprehension Test. The correlation between the Iowa test and the Science Content Comprehension Test could then have been due largely to that 20 percent of items involving the same operations, rather than the need for those operations in comprehension.

Further, it should be pointed out that, although the Raven's Test of Logical Operations and the Science Content Comprehension Test did correlate moderately with a number of other tests, these moderate correlations need not be attributed to the necessity for restructuring operations as the authors argue. They could be attributed to the necessity for restructuring operations as the authors argue. They could be attributed to a general test taking ability, verbal intelligence, or some sort of "g" factor. These general factors normally account for a substandard portion variance in many such studies and must be acknowledged. The computation of partial correlation coefficients with, say, verbal intelligence partialled out, would have been more informative. The computation of correlations as a hypothesis testing tool is useful, however, significant limitations of the method exist.

In conclusion, Raven and Polanski have suggested an extremely interesting thesis about the relationship between Piagetian logical operations and science content comprehension. They have obtained some support for their thesis, however it must be tested with more complex research procedures before its validity can be convincingly established.

## REFERENCES


1. Ausubel, D. P. Educational Psychology: A Cognitive View. New York: Holt, Rinehart and Winston, Inc. 1968.

2. Ausubel, D. P. "Some Psychological Considerations in the Objectives and Design of an Elementary School Program." Science Education, 47(3):278-284, 1963.

3. Ausubel, D. P. The Psychology of Meaningful Verbal Learning. New York: Grune and Stratton. 1963.

4. Ausubel, D. P. The Transition from Concrete to Abstract Cognitive Functioning: Theoretical Issues and Implications for Education." Journal of Research in Science Teaching, 2(3):261-266, 1964.

5. Lawson, Anton E., and John W. Renner. "Relationships of Science Subject Matter and Developmental Levels of Learning." Journal of Research in Science Teaching, 12(4):347-358, 1975.

6. Lawson, A. E., and W. T. Wollman. "Encouraging the Transition from Concrete to Formal Cognitive Functioning: An Experiment." Journal of Research in Science Teaching, 13(5):413-430, 1976.

7. Novak, J. D. "Editorial Comment on the Effect of Visual Devices Based on Bruner's Modes of Representation on Teaching Concepts of Electrostatics to Elementary School Children." Science Education, 60(1):85, 1976.

8. Novak, J. D. "Understanding the Learning Process and Effectiveness of Teaching Methods in the Classroom, Laboratory, and Field." Science Education, 60(4):493-512, 1976.

9. Raven, R. J. "Programming Piaget's Logical Operations for Science Inquiry and Concept Attainment." Journal of Research in Science Teaching, 11(3):251-261, 1974.

Nordland, Floyd H., Anton E. Lawson, and Jane B. Kahle, "A Study of Levels of Concrete and Formal Reasoning Ability in Disadvantaged Junior and Senior High School Students." Science Education, 58(4): 569-576, 1974.

Descriptors--*Concept Formation, Disadvantaged Youth, Intellectual Development, Instruction, *Learning Theories, *Measurement, Science Education, Secondary School Science, *Task Analysis

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Gene Craven, Oregon State University.

## Purpose

The stated purpose of this study was "to evaluate the extent to which disadvantaged junior and senior high school students have developed selected concrete and formal conservation concepts and to evaluate the extent to which they have acquired formal operational reasoning ability."

## Rationale

Four sequential stages of cognitive development are postulated in Piagetian theory, namely: sensorimotor, pre-operational, concrete, and formal operational stages. Development within each of the stages follows a fairly predictable path; (a) an initial experimenting phase during which strategies of experimental interpretations are acquired, (b) the progressive accumulation and elaboration of advanced techniques affirmed by experimentation, and (c) a restructuring and consequent extension of existing cognitive structures to incorporate the newly acquired strategies.

At eleven or twelve years of age there is a transformation in a child's thinking from concrete to formal thinking. Up to this age the operations of intelligence are "concrete," i.e., they are concerned with tangible objects that can be manipulated and subjected to real action. As of eleven or twelve years the logical operations begin to be transposed from the concrete to the ideal or formal plane. Two cognitive skills underlie formal operations--the ability to subordinate the real to the possible and the ability to reflect on one's thought.

While the ages at which children's cognitive development is transformed from one stage to the next vary from culture to culture, the order is invariant. The investigators cite studies which show that "rates of attainment of concrete reasoning ability (by children) . . . vary significantly with socioeconomic level as well."

## Research Design and Procedure

Disadvantaged students randomly selected from two separate populations attending urban schools were administered Piagetian tasks in individual interviews to determine the extent to which they had acquired formal operational reasoning abilities. One sample consisted of 96 students (ages 11.4 years to 14.4 years: mean age = 12.6 years) enrolled in seventh grade science classes at a predominantly black and Spanish-American junior high school. The other sample consisted of 506 science students from a predominantly black senior high school.

The tasks, all of which were administered in individual interviews, had been employed by previous Piagetian investigators. Thus, only brief identifying descriptions of the tasks and materials were included.

If a student was found to be a nonconserver of weight, he was given tasks on conservation of volume-clay, conservation of volume-metal cylinders, separation of variables, and equilibrium in the balance (7th graders) or exclusion of irrelevant variables (high school students). The formal separation of variables, equilibrium in the balance, and the formal exclusion of variables tasks were designed to measure formal reasoning abilities.

Criteria used in classifying a student's responses to each of the tasks and the corresponding point scores are described. Two points were awarded for successful completion of a conservation task--one point for a correct conservation response and one point for a correct explanation. A total of three points each was possible on the formal reasoning (exclusion of variables, separation of variables, equilibrium in the balance) tasks. Criteria for classifying and awarding points for responses to the formal reasoning tasks are described adequately for replication by persons familiar with the tasks. Subject responses to each task were categorized and points awarded as follows:

| | | |
|---|---|---|
| II A | Early Concrete Operational | 0 points |
| II B | Fully Concrete Operational | 1 point |
| III A | Early Formal Operational | 2 points |
| III B | Fully Formal Operational | 3 points |

The total interview score for a subject was calculated by summing the individual task scores. Due to the number and nature of the tasks administered, the Piagetian operational levels of thought and the range of scores were as follows:

| | | |
|---|---|---|
| 0-1 points | | Preoperational Thought |
| 2-6 points | II A | Early Concrete Operational (Three conservation tasks--area, length, weight) |
| 7-14 points | II B | Fully Concrete Operational (II A tasks and two formal tasks awarded 1 point each) |
| 15-20 points | III A | Early Formal Operational (Two conservation of volume tasks and two formal tasks) |

21-22 points — III B Fully Formal Operational
(III A and successful completion of the separation and exclusion or equilibrium tasks)

The highest possible score for the nonconserver of weight was 10 points earned by conservation responses and correct explanations for the 5 remaining conservation tasks on which from 60 to 98 percent of the subjects were successful.

## Findings

In general, the two groups of students performed similarly. Between 91.7 and 97.9 percent of the subjects were successful on the conservation tasks; number, continuous quantity, and substance. A second group of tasks of similar difficulty (between 47.9 and 72.9 percent) included conservation of area, length, and weight. Only 3.1 to 11.0 percent of the subjects demonstrated conservation responses on the volume tasks. Data were not reported for success on the formal reasoning tasks.

| Piagetian Level | Percentage of TOTAL SAMPLE: 96 Seventh Grade Students | 506 Senior High Students |
|---|---|---|
| I Preoperational | 1.0 | 1.0 |
| II A Early Concrete Operational | 16.6 | 16.8 |
| II B Concrete Operational | 66.8 | 69.0 |
| III A Early Formal Operational | 15.6 | 13.1 |
| III B Formal Operational | 0.0 | 0.1 |

According to these data, the majority (84.4 and 85.8 percent) of these 11.7 to 20.0 year old subjects were concrete thinkers. Only about 13 to 16 percent of these subjects demonstrated evidence of formal reasoning ability. Correlations between subject ages and the total Piagetian task scores were reported to be near zero for both the junior high school sample ($r = -0.03$) and the senior high school subjects ($r = 0.00$).

## Interpretations

"These results suggest that the lag in acquisition of conservation concepts in disadvantaged primary school children continues and probably becomes greater in disadvantaged adolescents. The finding that only about 13 to 15 percent of these subjects demonstrated any evidence of formal reasoning ability coupled with the lack of correlation with age suggests that, for the majority of persons in this segment of society, formal operational skills will probably never develop under existing conditions" (p. 574).

"Questions concerning appropriate kinds of instructional methods and materials for this type of student seem indeed crucial. While a person is in the concrete operational stage his thinking and understanding is restricted to what he can physically see, feel, and experience in first-hand situations. He is unable to deal with abstract concepts or processes. Therefore it is imperative that classroom activities, in such subjects as science, involve concrete materials of the discipline. For concrete thinkers to develop meaningful understandings, the laboratory, not the textbook, must become the major source of information and provide the basis for discussion" (p. 574).

## ABSTRACTOR'S ANALYSIS

This study adds a new dimension to a large and rapidly growing body of research on the Piagetian model of cognitive development. It extends to adolescents the findings of studies which report "significant socioeconomic differences in young children's acquisition of conservation concepts" (p. 569). The finding that a majority of the disadvantaged adolescents in the secondary school science classes in this study demonstrated no evidence of formal reasoning contributes to a theoretical basis which should be of importance to curriculum developers and classroom teachers.

In his comprehensive summary of Piagetian research, Modgil states, "Perhaps in no other area of psychology is there so much cross-cultural and cross-social-class empirical research data available as on the Piagetian tasks". He concludes, "however, with the evidence available so far, it is difficult to make any sweeping statements about cross-cultural replication of Piaget's findings. There are problems in interpreting the results from cross-cultural studies, partly because of differences in language and partly due to experience and cultural values" (5:226).

While the cross-cultural and cross-social-class research cited by Modgil is extensive, it is quite diverse. Only a small number of studies have included U.S. blacks and other groups identified as disadvantaged. Except for the present study, the subjects have been young children.

An investigation cited in support of this study was by Waisk and Waisk who reported that "for culturally disadvantaged primary children, acquisition of a variety of concrete conservation tasks lagged 1 to 2 years behind the ages at which such concepts are mastered by middle class children" (6:1587). No similar comparisons can be made from the present study since it was limited to disadvantaged (predominantly black) science students. According to the Campbell and Stanley (1:176) nomenclature, it is a One-Shot Case Study which provides considerable information about a single population but for which caution must be exercised in drawing causal or comparative inferences.

Space limitations that are imposed on the authors of any journal article often leave unanswered many details and questions regarding the research design and investigative procedure which, while appearing to be relatively trivial, could be important to persons conducting replication studies.

Two unstated assumptions appear to have been made by the investigators. The first is that all students attending the two schools and participating in the study were disadvantaged. If this fact was established empirically, the operational definition of "disadvantaged" and confirmation that the subjects were indeed disadvantaged would add to confidence in the findings.

A second assumption is implied in the title of the journal article in which research data from two populations of science students attending predominantly black and Spanish-American urban schools have been generalized to "disadvantaged junior and senior high school students." Predominantly black and Spanish-American implies that students from other racial or cultural groups attended the schools from which the subjects were selected and could have been participants in the study. Since race and cultural group references were used in the journal article to define disadvantaged, an analysis of the school population and the subjects of the study by race and cultural group would be helpful in interpreting the findings.

A study by Gaudia (2) suggests that there is a difference among lower-class environments of racial groups that affects performance on Piagetian tasks. Using subjects from several schools in Western New York state, he found that "Negro children performed at a lower level of conservation than Indian and white children" (2:163). He argued that "this increasing difference between racial groups with increasing chronological age suggests that environments may be entirely different among races" (2:163). Piaget claims that environment is a factor in cognitive development, but a relatively minor one and then only when the cultures are widely divergent. Until empirical data lead to a rejection of racial differences due to differing environments, caution should be exercised in generalizing from a sample representing one or two disadvantaged cultural groups to all disadvantaged students. Also, the possibility of racial differences in conservation due to differing lower-class environments raises questions regarding the validity of combining urban black and Spanish-American students into one group of disadvantaged students.

Inhelder, Sinclair, and Bovet conclude that "differences in the ages of acquisition of various concepts have been frequently noted in cross-cultural research and seem to be governed by the amount of cognitive stimulation the child receives in his everyday life" (4:128). The present study provides little information about the everyday life of the subjects or data on their psychological, cultural, academic, or cognitive attributes other than performance on the Piagetian tasks. Could such data provide clues for the finding that the senior high school subjects performed less well on the conservation tasks than did the junior high school subjects?

For persons familiar with Piagetian researth, the journal article adequately defines the conservation and formal reasoning tasks via a paragraph devoted to each. These tasks have been employed and described by many investigators. While the tasks are adequately defined, the conditions of the interview and the interview technique necessarily vary from study to study due to different physical facilities, different persons conducting the interviews, and possibly different ways in which the tasks are administered. A detailed description of the interview procedure followed in the present study would permit comparisons to other studies and make replication possible. It would be useful to know how the interviewers were trained, the degree of experimenter variance, whether or not a given interviewer interviewed the same proportion of subjects from each of the two populations, how the students were introduced to the tasks, and the procedures used in recording the data.

An effect of interviewer technique in performance of the subjects is reported by Inhelder, Sinclair, and Bovet who state that "in the experimental situation it is often necessary to repeat the question several times in a variety of different ways, so as to elicit the use of unfamiliar types of reasoning" (4:128). Greenfield and Bruner (3) report that many unschooled subjects were led from nonconservation to conservation when the child himself carried out all the actions with the liquid conservation rather than observe the experimenter do so. In the present study the "disadvantaged" subjects are likely to have been unfamiliar with the types of reasoning required by the tasks and may have been led from nonconservation to conservation had there been an opportunity to carry out the actions with the materials.

A basic criterion of statistical analysis appears to have been met by the fact that the subjects "were randomly selected from seventh grade science classes" and were "randomly selected science students from a . . . senior high school" (p. 569). It would be helpful to know if the sample was randomly selected from the set of all of the students in all of the seventh grade science classes and from all of the high school science students. Assuming that this was the case, it would be interesting to know the rationale for the numbers of students (96 and 506) who were selected as subjects.

In summary, the study is described succinctly and concisely, giving the reader a reasonably clear description of the study and its findings within the space limitations imposed on the authors of a journal article. The findings are clearly reported in graphical and tabular form and the conclusions are consistent with the data presented. Several important implications for science teaching are considered.

A relatively small amount of research has been conducted to determine the Piagetian level of cognitive development of junior and senior high school students who are commonly taught science and mathematics as if they are at the level of formal reasoning. Further comparative studies appear to be warranted to determine the extent to which cultural and class environments affect formal reasoning ability. Support for such studies is provided by Inhelder, Sinclair, and Bovet who conclude that "investigations in the field of the role of cultural environment on the process of operatory development is only in its infancy and much more research is needed" (4:270).

## REFERENCES

1. Campbell, Donald T. and Julian C. Stanley. "Experimental and Quasi-Experimental Designs for Research on Teaching," Ch. 2 in the Handbook of Research on Teaching, N. L. Gage, Editor. Rand McNally & Company. 1963.

2. Gaudia, G. "Race, Social Class, and Age of Achievement of Conservation on Piaget's Tasks." Developmental Psychology, 6(1):158-165, 1972.

3. Greenfield, P. M. and J. S. Bruner. "Culture and Cognitive Growth," International Journal of Psychology, 1:80-107, 1966.

4. Inhelder, Barbel, Hermine Sinclair, and Magali Bovet. Learning and Development of Cognition. Harvard University Press. 1974.

5. Modgil, Sohan. Piagetian Research: A Handbook of Recent Studies. NFER Publishing Company, Ltd., Windsor, Berks (Great Britain). 1974.

6. Waisk, B. H. and J. L. Waisk. "Performance of Culturally Deprived Children on the Concept Assessment Kit-Conservation." Child Development, 42:1586-1590, 1971.

CONCEPT LEARNING

McIntyre, Patrick J. "Students Use of Models in Their Explanations of Electrostatic Phenomena." Science Education, 58(4):577-580, 1974.
Descriptors--*Concept Formation, *Educational Research, Electricity, *Elementary School Science, Instruction, *Learning Theories, *Models, Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Lowell J. Bethel, University of Texas at Austin.

## Purpose

The stated purpose of this research was to investigate children's understanding of specific types of electrical phenomena (i.e., electrostatics) and the extent to which children used models or analogies in their explanation.

## Rationale

Research into young children's understanding of scientific concepts and phenomena has been conducted over a number of years. Studies have been conducted which indicate that children use models to explain various types of phenomena. Thus, children can be classified as modelers or nonmodelers based on their responses to questions about specific phenomena. This investigation is an extension of these studies relative to electrostatic phenomena.

## Research Design and Procedure

A total sample of 57 pupils were randomly selected from a suburban elementary school population of 405 in grades two through six. Pupils were shown three demonstrations: (1) a charged comb picking up small pieces of paper, (2) a charged balloon sticking to a wall, and (3) a simple circuit being opened and closed by a switch. After each demonstration five questions were asked:

1. What did you see happen?

2. How would you explain what happened to another student?

3. What do you think the comb (balloon, switch) might be like for it to work the way it does?

4. Do you think your explanation would help another student understand what happened?

5. Is there anything else you would like to tell me about the demonstration or how you would explain it?

All pupils were interviewed individually and their responses were recorded on audio-tape. The tapes were analyzed to determine if the pupil used a model to explain the observed phenomena. Accuracy of responses was not considered. The pupils were interviewed until there were at least eight pupils in each of the categories (modeler and/or nonmodeler). A pupil was classified as a modeler if he used a model in at least two explanations.

## Findings

No traditional statistical treatment was used to analze the data. The data were summarized in terms of the number of pupils interviewed and classified as modelers. It was found that no pupils in grades two through four used models or analogies in the explanation of the phenomena. Approximately 50 percent of the fifth and sixth graders used models in their explanations of the demonstrations.

## Interpretations

The findings of this investigation confirm the results of previous studies in that elementary school pupils can and do use analogies or models in their explanation of selected phenomena. It is suggested that the reason that children below grade level five do not use models is because of instruction. This was concluded because some modelers stated that they received instruction in atomic structure while those in grades two through four did not make similar admissions. Thus, it can be conluded that the use of models is probably due to instruction.

## ABSTRACTOR'S ANALYSIS

While this is a nice little investigation, it does not really tell us anything that we did not know already. It appears to be a repetition of the work of Zeigler and Anderson who are referenced in the introduction of McIntyre's article. It would have been good if the investigator had discussed his findings in relation to the studies cited and the contributions this study makes to this area under investigation.

One shortcoming of the study is the description of the sample used. A major objective in reporting research is to provide sufficient information so that the study may be replicated if indeed this is desired. However, in light of this description, this could not be done. Thus, it is also difficult to generalize to other pupils because of this lack of information.

The investigator goes on to state that substantial numbers of pupils in grades five and six use models in their explanations of selected electrostatic phenomena. While this appears to be overstated (50 percent and 53 percent respectively), it is difficult to form any conclusions about children's use of models in explaining phenomena.

Next, the investigator suggests that modeling and the use of models is probably due to instruction. How does he arrive at this conclusion? By stating that the pupils made mention of bits and pieces of information about the topic of previous instruction. However, this was not pursued in the study. Yet it was used to make a conclusion. This is a bit shoddy and should not have been included since this was not really an important part of the study. Nor was this information actively sought as evidenced by the questions asked during the interviews.

Reference is made to a "study of the effectiveness of different types of models in classroom instruction." However, this has no direct bearing on the investigation reported. Why this is mentioned is not clear from the body of the paper.

Finally, it is concluded that the study not only confirms but extends the idea that elementary pupils use models in their explanation of phenomena, but this is never really explained beyond the research cited. It would have been helpful to the reader to expand on this in the "Discussion" section. This was not the case in this article.

An interesting question to this abstractor is: Why did the investigator choose the number of eight pupils for each occupied category? This is never really explained. Since the investigator did not use traditional statistical procedures for data analysis, why the choice of eight pupils as a minimum? This should have been explained since it is not evident in the article.

In conclusion, the questions raised above need to be explained if the reader is to receive a good idea of the sense of this investigation. The article does have merit but has not been reported properly. Finally, the "Discussion" section should be used to discuss the implications of the study. This would have clarified some of the questions and points raised about the article.

Robertson, W. W., and E. Richardson. "The Development of Some Physical Science Concepts in Secondary School Students," Journal of Research in Science Teaching, 12(4):319-329, 1975.

Descriptors--*Conservation (Concept), Educational Research, Learning Theories, *Physics, Science Education, Secondary Education, *Secondary School Science, *Sequential Learning

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

## Purpose

To replicate and extend a number of tests of "conservation" of some physics concepts; to measure the conservation of some concepts not previously tested; to administer with standardized procedures the tests on a group basis, checking reliability with standard clinical testing; and to investigate predictions based on hypotheses of the hierarchical attainment of concepts in physics.

Specific hypotheses are:

A. If the conservation of derived quantity in physics is dependent upon the prior conservation of constituent quantities, then students will conserve:
   1. mass before weight
   2. length before area
   3. length and time before speed.

B. If the conservation of a derived quantity in physics is dependent upon the prior conservation of its elements, then students will conserve:
   1. length and area before volume
   2. mass and volume before density
   3. area and force before pressure
   4. mass and acceleration before the force relationship (F = ma)
   5. force and distance before the work releationship (W = Fs).

## Rationale

While much is now being made of the hierarchical structure within science in curriculum projects and the stages or levels of cognitive development in learning theory, little research evidence exists in relation to such basic questions as: (a) are science concepts attained in particular hierarchical sequences, and (b) is the conservation of a derived quantity in physics dependent upon the prior conservation of the fundamental quantities--mass, length, and time?

The present research is related to Piaget's genetic approach to cognition and work done by Elkind; Lovell and Slaters; Lovell, Healey and Rowland; Lovell, Kellet and Moorehouse; and others concerned with developmental stages.

## Research Design and Procedure

A sample of 25 boys and 25 girls was randomly drawn from each of grades 7 to 10 in one school giving a stratified sample of 200. The mean IQ was 107.3 (S.D. 11.9) with no significant difference between samples for each grade.

The concepts - mass, length, time, weight, area, speed, volume, density, force, pressure, acceleration, force, and work - were selected for study as a result of analysis of the concept "energy." A group of tests for each concept sequence was developed. The pupils were tested in groups of 25, utilizing procedures to minimize subject-subject interaction and learning through tests. Some of the tests were modifications of tests used by other investigators--Elkind (mass, weight and volume), Lovell and Slaters (time), Lovell, Healey and Rowland (distance and length), and Lovell, Kellet and Moorehouse (speed).

The results were based upon children's predictions, judgments and explanations as signs of conservation. Each subject was classified as exhibiting conservation according to the following criteria:

1. Conservation: The subject responded correctly to all the questions related to the task.

2. Nonconservation: The subject makes one or more errors in prediction, judgment or explanation.

The classifications were subjected to scalogram analysis. The criterion used for assigning the conservation of a quantity to an age level is the age or grade at which 75 percent conserve.

## Findings

1. The 75 percent criterion was exceeded at grade 7 by boys and girls for the concepts: mass, weight, weight-force, length, distance, speed (straight tunnels) and speed (concentric circle). In addition, at grade 8 vertical height was conserved by both boys and girls. At grade 9 time was conserved by boys and girls, but volume by boys only. At grade 10 boys conserved area.

2. In relation to the general hypotheses the results of the concept attainment provide evidence concerning possible hierarchical structures for learning physics.

## Interpretations

Clearly the conservation of derived quantity in physics is not necessarily dependent upon the prior conservation of its constituent fundamental quantities.

Depending on the choice of operational definitions describing basic phenomena and the logical manipulations of these definitions, different patterns of the organization or structure of the discipline can be built up. The research indicates that time is not a first order concept. Time is more removed from reality than speed.

The tests of the conservation of pressure, force, acceleration, work and potential energy provide dramatic evidence of the lack of understanding of these concepts. If a student is unaware of the invarient aspects of a concept in the face of transformations, his understanding is very limited. There is much evidence of verbal learning.

## ABSTRACTOR'S ANALYSIS

This study compares favorably with others of the type that are concerned with selected ideas from Piaget. The testing has been carried on with care and the data have been subjected to reasonable analysis. It may be assumed that the results of any research are credible if the measurement devices are reliable and valid. It is also obligatory to assume that the results may lack credibility if the data are curious.

This study, though nicely reported, except for the confounding of the results in the discussion and the extreme statement, "clearly, the conservation of a derived quantity in physics is not necessarily dependent upon the prior conservation of its constituent fundamental quantities", presents the reader with many problems. Some of these are:

1. There are no reports of instructional programs provided for any grade level. If there were no programs of instruction the entire study is without merit. It is not possible for pupils in grades 7-10 to intuitively develop derived definitions and units as density, force, pressure, weight, acceleration, etc., and also not possible for them to intuitively develop a system in which length, mass, and time are fundamental to the units derived.

2. There are no definitions of the terms <u>predict</u>, <u>judgment</u>, and <u>explain</u>; the fundamentals that make up the data. Was the term <u>predict</u> to mean the forecasting of the future based upon use of a given scientific law in which noncapriciousness of nature is accepted, or something else? Did the term <u>judgment</u> mean guess, intuition, or discursive reasoning? Did the term <u>explain</u> mean the application of a scientific law or teleological purposefulness? The frequent use of the term "why" in the questions leaves the reader confused--is this the way of the educational psychologist, or is it an attempt to get to something else?

3. Probably the most serious problem is the use of the term <u>conservation</u>. As one reads the report it becomes more and more impossible to give meaning to the results because the term is not used as in science in the statement of the laws of conservation of energy, matter, momentum, charge, etc. According to this study the notion of conservation was "the ability of the student to be aware of the invariant aspects of a concept in

the face of transformations." With this use it must be recognized that in physical theories and laws, invariance occurs relative to a frame of reference; conservation occurs within certain physical systems, namely closed systems.

Within this paper and in the test questions there is no concern for the requirement of a system. It seems that the meaning attributed to conservation in this paper was that attributed to the term <u>equivalent</u>. There are significant differences between the two concepts, hence the results lack credibility. Note the use of the concept of equivalence in the test items. (a) The test for density involves two plasticine objects, one of which is formed into a cylinder. (b) The test for time involves two cars moving through tunnels of unequal length. (It is also impossible to have a concept of the conservation of time in a non-relativistic system, however, time equivalence is possible.) (c) The use of force as a conserved quantity is not reasonable. Note that force is a vector quantity and thus it has magnitude and direction. Reflect upon the common experience of using a simple machine; unequal forces may be acting. The equivalence idea is again present, but now the idea must include moments; two trolleys are used and comparisons made. (d) Acceleration is another nonconserved quantity and two trolleys are used. (e) Vertical height is not a conservable quantity. As in all other instances the right or wrong answers depended upon knowledge of a general physical law; in this case "the work done is independent of the path of the force." (f) Conservation of work relates to conservation of energy since energy is potential work. Again two trolleys were lifted.

In all examples the use of knowledge of laws in physics would, and did, enable the subject to score properly; the subject could respond as equivalent or not equivalent concerning the two instances.

At no time was the criterion of a system mentioned, the subjects were merely asked to identify equivalents. The subjects were never asked to apply the concept of conservation to explain observations; they were merely asked to demonstrate a functional knowledge of some physical law.

The best that may be said from the results would be that the pupils did not have mastery of the physical laws applied in the problems:

$$\text{density} = \frac{m}{v} \quad d = vt, \quad v = \frac{d}{t}, \quad p = \frac{f}{a}, \quad f = ma, \quad w = f \times d.$$

It is definite that no statement may properly be made about conservation of energy, conservation of matter, conservation of momentum, conservation charge, conservation of mass number, etc.

The second concern was for hierarchical structure for learning physics. It seems that the evidence, if existing at all, is weak to indicate a possible structure for learning physics. In order for the research to produce meaningful data on this, a variety of teaching sequences would be necessary. Even this procedure would be risky because each may be based upon erroneous judgments of the increments of a particular learning product.

The conclusions exhibit some attitude of freedom to go beyond the data: "the results of measurement of concept attainment shown in Table II provide evidence concerning possible hierarchical structures for learning physics." On the same page the claim is "clearly, the conservation of derived quantities is not necessarily dependent upon prior conservation of its constitutent fundamental quantities."

There was essentially no experimental design.

Although the data were treated as null hypotheses they were not so stated.

The authors expressed serious concern for the reliability of the instruments used but ignored the important quality of validity. The data are no better than the instruments.

It thus seems that this project, though nicely conducted and reported, has produced little of value. The real problems come from the confounding of the scientific concept of conservation necessitating a closed system concept and the concept of conservation as being equivalence. This confounding is magnified by attributing the quality of conservation to quantities not really conserved in physical manifestations.

The ideas of sequences of increments of learning needs to be researched for all science concepts, empirical laws, and theoretical laws. The first steps must be the identification of the increments and some means of measuring knowledge of the increments.

The consistent use of vague undefined terms will continue to produce confounding findings.

Kass, Heidi. "Structure in Perceived Relations Among Physics Concepts." Journal of Research in Science Teaching, 8(4):339-350, 1971.
Descriptors--*Cognitive Processes, *Factor Analysis, Factor Structure, Learning Theories, *Physics, *Psychology, Research Methodology, Scientific Concepts, Secondary School Students

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

## Purpose

This study was conducted to determine the cognitive structure of the relations among 20 mechanics concepts (e.g., Newton's Laws, centripetal force, centrifugal force, friction and uniform acceleration) as perceived by high school physics students.

## Rationale

Much social science research has been conducted on analyzing the degree of differentiation of the individual's personal structure and the influence which the nature of this structure exerts upon judgmental behavior. The author suggests that consideration of such a conceptual basis for judgmental behavior should not be restricted to social cognition. The students' perceptions of the domain of science concepts may influence their judgmental behavior. Isolating and representing aspects of the cognitive structure related science concepts should be the first step in answering questions such as: How does the structure of perceived relations among the concepts affect performance in the subject?

## Research Design and Procedure

Three hundred fifty-three Grade 12 physics students rated the difference in difficulty between 190 pairwise combinations of 20 mechanics concepts on a nine-point scale from 1, very similar in difficulty, to 9, very different in difficulty. For data analyses, the subjects were randomly assigned to three groups of 67 in order to assess the extent to which one may expect to obtain similar results across samples drawn from the same population. For each group, the difficulty difference ratings were arrayed in a matrix consisting of 190 rows for the concept pairs and 67 columns for the subjects. The matrix of sums of squares for individuals and sums of cross products between individuals was then analyzed through principal component factoring.

## Findings

Analyses of the data indicated that either a four- or a five-dimension representation would be appropriate. The factors in the four-dimension solution were labeled by the author as:

I Motion-Statics
II Vectors
III Gravity-Circular Motion
IV Force-Work-Power

The five-factor solution contained all the above factors and another which was not readily interpretable but involved a cluster of concepts on kinetic and potential energy.

## Interpretations

The author believes that perceived differences in the difficulty of physics concepts resemble distances in Euclidean space. And students construe the concepts along two or more difficulty dimensions. Also, the perceptual space seemed to be relatively stable for different samples from the same population.

## ABSTRACTOR'S ANALYSIS

The application of multidimensional scaling techniques to describe the structure in students' perceptions of science concepts is an important area of research and should be pursued. But results in this area are going to be elusive and costly. This study is only a beginning, with 20 physics concepts, and the results can hardly be considered elucidating. The author does need to be commended for assessing the degree of generalizability, by triple sampling, in an area for which conventional probability statistics and error estimates are not available.

The study is a step along a long road of determining the nature of the learner and the learning process, and instructional strategies which maximize student achievement should be based on current knowledge in this area. But at this point in time we need to channel the majority of our resources into more evaluative research on methods of improving achievement and attitudes of the science learner.

# TEACHER BEHAVIORS AND ATTITUDES

Butzow, John W., and Alan Davis. "The Development of a Semantic Differential Test of Teachers' Attitudes Toward Teaching Elementary School Science." Science Education, 9(2):211-220, 1975.

Descriptors--*Career Choice, Educational Research, *Educational Philosophy, Learning Theories, *Student Attitudes, *Student Teachers, Teacher Behavior, *Teacher Education

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

## Purpose

This study had two objectives: (1) to develop an instrument for measuring teachers' attitudes toward teaching an open-ended type of science program such as Elementary Science Study (ESS), and (2) to administer this instrument to a group of teachers using the ESS program and correlate the results with their actual teaching behavior as viewed and analyzed via video tape, hence building a case for reliability and validity.

## Rationale

The success of an elementary school science program such as ESS depends in large measure on the degree to which the teacher is willing to follow the philosophy embodied in the curriculum. Investigators in this study sought to develop a measure that would predict the degree to which individual teachers would be "student centered." Further, by correlating responses to this attitude instrument with data generated by an already existing observation scale, it was possible to make judgments on the validity of the new instrument. This instrument was based on the semantic differential technique and followed earlier work by the senior author.

## Research Design and Procedures

In the first part of the study, standard procedures described by Osgood et al. (2) were used to develop an attitude instrument called the Semantic Differential Test of Teacher Attitudes (SDTTA). Initially, 104 elementary school majors enrolled in a science methods course responded to three concepts relevant to teaching science in the elementary school: "For me, doing science is ...," "For me, teaching science is ...," For me, science concepts are ....". Using a five-point scale, each concept was subjected to 46 adjectival pairs. Results were factor analyzed using the Varimax procedure. The investigators found four major factors which they categorized as valuing, enjoying, striving, and difficulty. For the SDTTA they selected the adjective pair under each category with the highest loading ("important-trivial" for valuing, "enjoyable-unenjoyable" for enjoying, "powerful-powerless" for striving, and "easy-difficult" for difficulty) and used these four bipolar adjectives to measure feelings toward 21 teacher behaviors associated with the ESS program.

After the SDTTA was developed, the instrument was administered to 29 elementary classroom teachers attending an inservice institute for implementation of ESS. The teachers responded to the instrument after completing formal course instruction but prior to actual classroom implementation. After classroom work with ESS had progressed for several months, each teacher was filmed on videotape while teaching science. The Science Curriculum Assessment System-Teacher (SCAS-T) developed by Matthews and Phillips was used as a scale for judging the videotaped teacher behaviors. The teacher behaviors being studied were independently rated by three professional science educators trained to use this instrument and were divided into two subgroups on the basis of the teacher directed index (TDI) of the SCAS-T. Hence, two groups emerged: one known as the teacher-directed group (TDG) and the other as the student-directed group (SDG). The latter group contained teachers with a lower teacher directed index, indicative of the type of teaching emphasized by the developers of ESS.

A Spearman Rank Correlation was conducted to determine the relationship between teacher scores on the SCAS-T and SDTTA. Further analysis using chi-square was conducted in order to determine if any items on the SDTTA produced significantly different scores across the two subgroups.

## Findings

When scores of the 29 teachers in this study were rank ordered for both the SCAS-T and SDTTA and were correlated, a coefficient of 0.79 was found (significant beyond the 0.01 level of confidence). When the 21 concepts contained in the SDTTA were compared across the two subgroups, TDG and SDG, Chi-square values suggested a significant difference (beyond the 0.1 level of confidence) between the groups on the following six items:

A. For me, allowing children to mess around with water is ...

I. For me, keeping live plants and animals in the classroom for use in experiments is ...

N. For me, being able to correctly answer student questions in science is ...

O. For me, allowing children to work in groups to discuss their point of view and findings is ...

Q. For me, teaching science is ...

S. For me, having a strong background in conceptual and factual science is ...

## Interpretations

In this study an instrument, the Semantic Differential Test of Teacher Attitudes, was developed, administered, and correlated with teacher behaviors and scored by trained observers using another instrument known as the Science Curriculum Assessment System. The strong positive rank order correlation (0.79) between the two measures suggests that teacher attitudes toward a teaching philosophy such as the one embodied in ESS can be used to predict the "student centeredness" of individual teachers. Also, findings in this study suggest there are some concepts associated with teaching a course like ESS which elicit significantly different attitudes depending on where a student "fits" along the continuum of "teacher directedness" versus "student centeredness." In short, this study produced an attitude instrument that appears useful in predicting the tendency of a teacher to exhibit student centeredness in the elementary school science classroom.

## ABSTRACTOR'S ANALYSIS

In conducting research where attitudes are measured by paper and pencil instruments, one can always raise the question, "Do responses on such instruments really reflect the true feelings of people, and are the alleged attitudes really reflected in their behavior?" The investigators in this study addressed themselves to this question, subsequently producing evidence that the manner in which a teacher expresses attitudes toward a philosophy of teaching indeed correlates positively with how they actually behave with students while teaching. A major contribution of this study, then, is that an instrument was developed that predicted for the 29 teachers in the study "student centeredness" as related to teaching science at the elementary school level. Furthermore, the feelings these teachers possessed toward concepts dealing with children, science, classroom management, teaching approaches, plants, animals, etc., corresponded to their observed behavior via videotape analysis.

The import of this study and these findings are significant for at least two reasons. First, an attempt was made to study attitudes and behaviors beyond a paper and pencil approach. In this regard, I believe this study serves as a potential model for other studies. Secondly, this study illuminates further the notion that the success of a given science curriculum may impinge at least in part on the philosophical and attitudinal orientation of the teacher. If a curriculum is designed for maximum student involvement but the teacher does not perceive this teaching style as "good," "important," or "powerful," then an incompatability results that may negate many of the strong features of both the program and the teacher. Many excellent science programs have been developed over the past two decades. In many instances these programs reflected definite assumptions and attitudes on the part of the developers. Many programs destined for success have failed because educators who adopted them did not possess attitudes congruent with the program.

This study can serve as a springboard for further research in several directions. The most obvious direction, perhaps, is that of preservice and inservice preparation to teach elementary school science. Conant (1) found that elementary school teachers in Portland, Oregon, taught science on the average of no more than two to three minutes per day. If this is indicative of a nationwide situation, one can one can conclude simply that elementary school teachers prefer not to spend their time teaching science. As I look back over the SDTTA instrument developed for this study, I can see many attitudes commonly found in our society (and in our classrooms) that are not only incongruent with the spirit of science but serve as serious barriers to the implementation of programs like ESS. The attitude of teachers toward science, children and teaching appears to be a powerful factor in how teaching behaviors are ultimately expressed. By dealing with these attitudes, science educators involved in teacher preparation and in-service programs will surely become more effective in bringing about change in the quantity and quality of elementary school science programs.

I found this study clearly written and easy to follow. The research design and statistical procedures appeared appropriate and were adequately communicated. The method used to develop the SDTTA was sound and indicated a thorough understanding of the semantic differential technique by the investigators. Perhaps more could have been said about the nature of the 29 elementary classroom teachers in this study and about the SCAS-T instrument developed by Matthews and Phillips. In working with a select group of subjects that may be more homogeneous than the average of the population, it is always difficult to arbitrarily group persons as "high" or "low" on a scale because of the potentially skewed nature of the group. Additional normative data on both the SDTTA and SCAS-T will be useful. This study provides an important link in current attitude research. These results suggest that attitude and teaching behavior are correlated and that an instrument measuring the former may be used to predict the latter. Further studies designed to consider cause and effect relationships appear in order.

REFERENCES

1. Conant, Eaton H. "What Do Teachers Do All Day?" Saturday Review World, 1:55, 1974.

2. Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. The Measurement of Meaning. Urbana: University of Illinois Press. 1957.

Pempek, Louise C., and David J. Blick. "An Evaluation of Elementary Teachers' Behavior and Attitudes in the Use of Inquiry-Oriented Science Programs." School Science and Mathematics, 73(5):414-419, 1973.

Descriptors--*Educational Research, *Elementary School Science, Inservice Education, *Program Effectiveness, Program Evaluation, Science Education, *Teacher Attitudes, *Teacher Behavior.

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

## Purpose

The major purpose of this investigation was to study changes in attitudes and classroom behaviors of teachers participating in the Cooperative College-School Science Program conducted at the University of Connecticut. Two specific questions were identified:

1. Did the Experimental Group differ from the Control Group in change of behavior from pretest to posttest?

2. Did the Experimental Group differ from the Control Group in change of attitude from pretest to posttest?

Although not specifically stated, another research question was:

3. What variables are related to changes in teacher behavior?

## Rationale

The authors point to the large amounts of money spent by the National Science Foundation in support of the development of new science curriculum projects. They state that even though it is assumed that teacher acceptance determines curriculum success, there has been little independent evaluation of the effects of the new curricula on the attitudes and behaviors of the teachers using them.

This investigation was designed to study attitudes and behaviors of teachers who were using new elementary science curricula. Further, it is stated that the project was conducted, "In order to initiate and implement the new inquiry-oriented elementary school science programs in Connecticut and to demonstrate the effectiveness of these programs in selected classrooms...."

## Research Design and Procedure

The research design used was a pretest-posttest control group design. There were 68 subjects in the experimental group and 14 in the control group. The treatment for the experimental group was participation in the Cooperative College-School Science Program. The control group subjects did not participate in the Program but were teaching inquiry-oriented science programs. Subjects were not randomly assigned to groups.

Pretests were administered in the spring of 1970 and posttesting was done in the spring of 1971. Two instruments were employed: The Pempek Teacher Behavior Checklist (teacher classroom behavior as viewed by students) and the Pempek Teacher Attitude Scale (measure of teacher attitudes toward science, science teaching, and scientists).

The attitude tests were completed by the teachers during pretest and posttest sessions. Ratings on teacher behaviors were completed by students of the individual teachers. An average of 30 students rated each teacher on both "pre" and "post" measures.

The experimental group attended four introductory sessions in the spring of 1970, a two-week workshop in August, and nine follow-up sessions during the ensuing year. The experimental treatment involved study of and actual experience with three major science programs: Science--A Process Approach (S--APA), Science Curriculum Improvement Study (SCIS), and Elementary Science Study (ESS).

The assumption was made that the Pempek tests did not necessarily yield interval data. Thus, all data were analyzed by use of non-parametric statistics: data on changes in behaviors were examined by the Mann-Whitney U Test; attitudinal change data were subjected to the Wilcoxon Sign Rank Test and the Spearman Rank Order Coefficient of Correlation; and the Kruskal-Wallis Analysis of Variance Test was used to determine the contribution to the variance by several different variables.

## Findings

The following findings are based on the authors' summary of data analysis:

1. Although the behavior of both the experimental and control group changed, the difference between the two groups was not significant at the 0.05 level.

2. The experimental group exhibited a statistically significant positive change in attitude from pretest to posttest at the 0.05 level.

3. The control group did not show a statistically significant change in attitude.

4. The teachers who had taken the fewest credit hours in science courses showed the greatest change in attitudes. Teachers who came into the program with strong science backgrounds showed the least change in attitude.

5. The attitudes of those teachers who had not taken graduate work in science changed more in terms of the objectives of the program than the attitudes of those teachers with graduate course work in science.

6. The teachers who taught in self-contained classrooms, even though they taught fewer hours per week with science project materials, had the most change in attitudes. Those teaching in departmentalized systems changed least.

7. Change in attitude as measured by the Pempek Teacher Attitude Scale in this study was found not to be statistically significant as related to the following factors:

   a. grade level taught;

   b. type of school district;

   c. science project selected;

   d. years of teaching experience.

Interpretations

On the basis of findings 4, 5, and 6, above, the authors conclude that:

> The teachers in self contained classes were the ones who had weaker backgrounds in science and who taught science only 0-3 hours per week. They were the teachers who at the beginning of the project were afraid to teach science and felt inadequate in science. Their attitudes changed most. However, those teachers in departmentalized upper grades had many courses in science and taught science many hours per week. They had definite attitudes established before participating in this project and they therefore exhibited the least amount of change in attitude.

The authors also concluded that the project was effective in meeting the objective of "introducing and implementing the use of the new science programs in Connecticut." They feel that the project not only improved science teaching and attitudes toward science, but that it also gave teachers a new perspective in which to view themselves, their teaching, and their interaction with students.

## ABSTRACTOR'S ANALYSIS

In general, this investigation fits into the larger group of studies on teacher characteristics which has received a great deal of attention over the past 15 years or so. Classroom interaction, teacher attitudes, etc., became primary research targets as the nation focused on improvement of teaching in the schools. Of course, partially as a result of the 1957 launching of Sputnik I, the science programs themselves also became targets of research and development. Thus, this study-which investigated relationships between teacher characteristics (attitudes and behaviors) and experiences with new science programs seems a logical outgrowth of the times.

In this report, however, the authors do not state clearly what experiences the teachers had with the new science programs. What was the experimental treatment during the introductory sessions, the two-week workshop, and the follow-up sessions, for example? How might this treatment have affected the teachers? Since the project involved three different programs (S--APA, SCIS, and ESS), is it possible that one program might have affected attitudes and behaviors differently than another? Certainly there are different philosophies which underpin S--APA and ESS, for example. And finally, which of the three projects were taught by the teachers of the experimental and control groups during the school year?

The basic research design (pretest-posttest control group) used in this study is sound. Because it is not stated differently, however, it must be assumed that the teachers were not randomly selected nor randomly assigned to groups. This factor is of special concern when one notes that there were great differences in the numbers of subjects in each group (68 experimental; 14 control) and that very little is known about the experiences of either group except that the control group teachers "...were not attending the workshop but...were teaching one of the inquiry-oriented programs for the first time...." But while randomization is ideal, behavioral researchers often find themselves in situations where it is not possible. The nature of this study suggests that was the case here. The authors are commended on their attempt through the use of the Median test to show that the experimental and control groups were drawn from the same population.

It would have been helpful if the authors had given more information on the behavior and attitude instruments and the nature of the data collected. There is no indication of the types of behaviors or attitudes examined nor is there information on the validity or reliability of the instruments themselves. The lack of such information places the validity of the total study in question. The study does reference the behavior and attitude instruments to the doctoral dissertation of one of the authors. However, since the dissertation is not readily available, the reader is left with several questions about instrumentation. Given the nature of the data assumed to have been collected, the use of non-parametric statistics appears to have been a wise choice.

The problem is not stated as succinctly and directly as it might be and the findings reported in the study are not entirely consistent with the research questions posed. One of the research questions, for example, asked, "Did the Experimental Group differ from the Control Group in change of attitude from pretest to posttest?" However, only the analyses of within group (experimental and control) differences are reported; the differences in between group performance on pretests and posttests are not reported. The authors also refer to null hypotheses which are accepted or rejected but there are no null hypotheses stated in this report. It is the abstractor's assumption that this report is derived from a much larger work, possibly a dissertation, and that some of the shortcomings identified above simply emphasize the difficulty of reducing a lengthy report into a much smaller size suitable for journal publication.

As a final comment on the written report itself, it is the abstractor's opinion that a better choice of caption (title) could have been made for Table 1. Generally, a caption should reveal something about the content (kind or nature of data) in the table--not simply the name of the statistic used to analyze the data (Kruskal-Wallis Analysis of Variance in this case).

Studies like this one and others on teacher attitudes and behaviors provide much insight into what teachers think and how teachers act. Great amounts of data have been accumulated. It would seem, however, that the press for performance-based certification, accountability, and the like, will demand that future studies on teacher characteristics be directed toward the relationships between those characteristics and <u>pupil success</u>. To continue to accumulate data on teacher characteristics without the link with pupil success will prove to be of limited value!

Orgren, James. "Using an Interaction Analysis Instrument to Measure the Effect on Teacher Behavior of Adopting a New Science Curriculum." Science Education, 58:431-436, 1974.

Descriptors--Classroom Observation Techniques, *Curriculum, Earth Science, Educational Research, *Instruction, Interaction Process Analysis, *Measurement, Science Education, Secondary School Science, *Teacher Behavior.

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Victor J. Mayer, The Ohio State University.

## Purpose

During the late 1960's and early 1970's, the New York State Department of Education revised the Regent's program in earth science. This study sought to document changes in teacher classroom behavior resulting from the implementation of the revised syllabus during a two-year time period when it first became available on an optional basis in 1970-71 and became mandatory in 1971-72.

## Rationale

No attempt is made by the author to relate this study to others concerned with the classroom behavior of teachers using "new" science curricula.

## Research Design and Procedure

Ten teacher-volunteers from each of the following groups were selected for study: (A) teachers who elected to continue teaching the traditional syllabus during the last year it was offered (1970-71), (B) teachers who elected to begin teaching the revised syllabus during the first year it was generally available (1970-71), (C) teachers who had participated in revising the earth science syllabus, and who were continuing to teach it for the second or more years. By the end of the study, attrition had reduced the number of teachers to about eight per group. More specific figures were not given.

Sequences of videotapes of their classroom behavior were collected by all teachers in Fall, 1970, Spring, 1971, and Spring, 1972, following directions provided them by the author. An additional sequence of tapes was recorded by teachers in Group A in Fall, 1971. Each sequence consisted of five consecutive days of teaching. Five trained analysts each categorized one day's behaviors for each teacher for each of the recording periods. Interobserver reliabilities were in excess of .70 on Scott's scale. Data consisted of the fraction of the week a teacher devoted to activities such as those listed below:

TABLE 1

SELECTED CATEGORIES OF THE TAPE ANALYSIS INSTRUMENT

| Item # | Descriptor |
|---|---|
| 2 | All laboratory related activity |
| 4 | Small group laboratory activity |
| 7 | All lecture-discussion activity |
| 8 | All higher level discussion in large group format |
| 15 | All knowledge and translation in large group format |
| 18 | Discussion of laboratory procedures in large group |
| 20 | All student verbal behavior in large group format |
| 25 | All teacher verbal behavior in large group format |

A General Index Score was also obtained from a weighted combination of items 2, 8, and 20 versus 7, 15, and 25. The complete listing of categories is not included in the report, nor is the procedure used to validate the list of categories.

Differences in teaching behavior were looked for among three groups based upon the Fall, 1971, data. Longitudinal differences within Group A were also examined over the two-year period of the study.

## Findings

The following differences between groups were found to be significant at the .05 level (using the F ratio). Both groups of teachers using the revised syllabus in Fall 1970 (Groups B and C) used laboratory related activities 80 percent of the time and lecture discussion about 20 percent, whereas the group using the traditional syllabus (Group A) used laboratory related activities 30 percent of the time and lecture-discussion, 70 percent. The teacher dominated during periods of large group instruction in all groups (80 to 85 percent).

Differences in teacher behavior were looked for between the Fall, 1970, data for group A and the Fall, 1971, data for the same group. They were teaching the traditional syllabus in 1970 and changed to the revised syllabus the following year. T-values were determined for each of the teaching behaviors, as well as the General Index Score. The General Index Score indicated that teachers employed significantly more of the advocated instructional behaviors after they adopted the new syllabus. Inspection of specific behaviors indicated in most cases a doubling of the time devoted to the desirable behaviors and a one-third reduction in time devoted to lecture-discussion. The fraction of time devoted to higher level discussion tripled.

When a similar analysis was performed on the Spring, 1971, and Spring, 1972, data, however, no differences were observed.

## Interpretations

The differences observed between those teachers using the traditional syllabus and those following the revised syllabus only suggest a modification of teacher behavior as a result of the adoption of a curriculum. Alternative explanations are available. When this information is combined with the longitudinal study, however, it seems quite convincing that a change of teaching behavior occurred among those teachers who began teaching the revised syllabus in Fall, 1971. That they seemed to "revert" to previous teaching practice in the Spring of 1972 can be explained by one or all of the following: (1) the nature of the revised Syllabus which is deficient in laboratory experiences at that time of the year, (2) a possible reversion to more "efficient" means of covering material when teachers became concerned about being behind the pace of the syllabus as spring approached, and (3) the involvement of some teachers in preparing their students for the Regents examination.

Whatever the cause for the apparent discrepancy noted above, it does point up a problem for educational researchers that is often ignored. Teacher behavior is highly unstable. Interaction analysis techniques, then, to be adequate in describing teacher behavior must be applied over long time frames, perhaps as long as a year or more.

## ABSTRACTOR'S ANALYSIS

In critiquing this study, one could take the author to task for failing to randomly select teachers, for small samples, and perhaps for other points of design and analysis. However, such criticism would be superficial and miss the implications of this study for research design in education. Science educators seem to have gotten "hung up" on sophisticated design and analysis procedures developed in the context of the physical and agricultural sciences, attempting to apply them in a field of study, education, which is totally unlike those for which the techniques were developed. And, not surprisingly, we end up with no results or mixed results. Creative effort must be directed to developing designs fitted to the problems of conducting in-school research with very complicated subjects, including teachers and students. This study makes a positive contribution in this regard. It used a sensible design, operable within the confines of the school situation and, as a result, it has obtained logical and sensible results; results that are interpretable and consistent with those of the few other researchers seeking to describe our schools and what goes on in them, a necessary first step in generating theories that will eventually enable us to do theory-guided research.

Unfortunately, however, there are three major deficiencies that reduce the significance of this study and the results obtained. First is the failure to report results from all of the teacher behavior categories included on the tape analysis instrument. It would appear that there were a total of 25 or more such categories, yet results are stated for only eight. Were the results from all of the others non-significant, or perhaps significant but in the "wrong" direction? If so, and if all categories were closely related to the nature of the revised syllabus, then the author's conclusion that teachers did indeed change their behavior is overstated. Changes in only six out of the 25 or so categories are hardly convincing of a significant change in teacher behavior.

The second deficiency is a failure to report Fall 1970 -- Spring 1971 longitudinal analysis for groups B and C. Such an analysis could add light to the interpretations of the discrepancy in the Spring 1971 -- Spring 1972 comparisons for group A. If there were a significant regression to "traditional" behaviors on the part of these two groups, then that would add evidence as to the the effect of curriculum characteristics on teacher behavior and, depending upon the potential extent of such differences, allow some judgement on the relative importance of the other two potential reasons for the Spring "regression effect". Several other comparisons could also be suggested, each with the potential of adding additional light to the interpretations made. Since these data were collected and presumably analyzed, one wonders why they were not reported.

The final major deficiency is a failure to relate the results of this investigation to others that have been conducted on teacher behavior in the context of new curriculum implementation.

Campbell, James Reed. "Cognitive and Affective Process Development and Its Relation to a Teacher's Interaction Ratio." Journal of Research in Science Teaching, 8(4):317-324, 1971.

Descriptors--Achievement, *Affective Objectives, *Cognitive Development, Curiosity, Educational Research, *Instruction, *Interaction Process Analysis, Scientific Enterprise, Secondary School Science, *Teacher Characteristics.

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

## Purpose

The purpose of this study was to investigate the relationship between a teacher's revised i/d ratio and the cognitive and affective process development of junior high school low achievers.

## Rationale

The rationale for the study grew out of the researcher's concern that previous efforts to determine teacher effectiveness using such variables as teacher ratings and/or characteristics were largely unsuccessful. Within the past decade, however, advances have been made in the development and use of systematic observation of classroom behavior and the results of teacher education research involving systematic observation have been reported as being more consistent than the earlier efforts. Indirect teaching has been found to correlate with student achievement and attitude at the junior high level in three short-term investigations by Flanders (6), Amidon and Flanders (1), and LaShier (8). This study examines similar relationships over a longer duration of time.

Assumptions were not presented, but two seem inherent in the study. The first specifically relates to the use of the Flanders System of Interaction Analysis. It assumes that teacher verbal behavior is an adequate sample of a teacher's total classroom behavior. The second assumption is either that the characteristics of the participating teachers and various other variables in the classroom environment are similar or that they do not appreciably influence student attitude and achievement.

## Research Design and Procedure

The research design did not follow one of the paradigms suggested and described by Campbell and Stanley in Gage (7), although it showed some similarity to the nonequivalent control group design. Pretests and posttests were administered to naturally assembled classrooms; however, comparison was made between two treatments, and random assignment was not under the control of the researcher. The independent variable was classroom behavior as measured by the Flanders System of

Interaction Analysis (FSIA) and the dependent variables were student cognitive and affective process development. Cognitive development was measured by the STEP test, and affective process development was defined as scores on DeProspo's Scale of Suspended Judgment, Howard and Robertson's Scale of Cause and Effect, and Campbell's Scientific Curiosity Inventory. The two scales measuring suspended judgment and cause and effect were modified, combined, and called the Scale of Scientific Attitudes.

The verbal classroom behavior of 10 junior high science teachers was taped for 11 to 13 forty-five minute lessons over the course of one academic year. The tapes were analyzed with the FSIA, and the data for each teacher were combined into a grand matrix, consisting of 360 minutes of predominant lecture-discussion activity. A revised i/d ratio was calculated for each teacher by adding the totals of tallies in categories 1 (accepts feelings), 2 (praises or encourages), and 3 (accepts or uses student's idea) and dividing by totals in categories 6 (giving directions) and 7 (criticizing or justifying authority). The teachers were ranked on the basis of their revised i/d ratios. The top five (mean i/d ratio of 3.14) were considered as being indirect, and the lower five (mean i/d ratio of 1.32) were considered as being direct in their classroom influence. The two groups were then shown to be significantly different (at the .01 level) by analyzing the combined matrices of both groups with a modified Darwin Chi Square.

A series of t tests was used to analyze differences between students in the two groups on an IQ test and all pretest scores on the criterion instruments. Since the two groups were essentially equivalent in IQ and on their Scale of Scientific Attitudes pretest, the posttest scores on the scale were subjected to a t test. Group scores on Campbell's Scientific Curiosity Inventory and STEP test were analyzed, using an analysis of covariance with IQ and pretest scores serving as covariates, because significant differences were found between the two groups of students on the pretest scores.

After the results of the statistical analyses were obtained, the grand matrices of both groups of teachers were compared. This analysis was performed in an attempt to identify elements within the teaching methodology which were responsible for the differences found in student performance.

## Findings

The findings reported by the investigator were as follows:

(1) the mean posttest score of students in the indirect group on the scale of Scientific Attitudes was significantly higher at the .05 level than the mean posttest score of students in the direct group;

(2) the adjusted mean difference on the posttest scores of Campbell's Scientific Curiosity Inventory was significant at the 0.25 level, with students in the indirect group being significantly higher than those in the direct group; and

(3) the adjusted posttest mean of students in the indirect group on the STEP test was significantly higher at the .001 level than was the adjusted posttest mean of students in the direct group.

## Interpretations

The conclusion was reached that the indirect teaching methodology was more effective than the direct for cognitive and affective process development of low achievers at the junior high school level. Differences in the posttest means of Scale of Scientific Attitudes were attributed to higher cooperation in the indirect classes as contrasted to hostility, resistance, and a high level of negative emotional feeling on the part of students in the direct classes.

After examining the grand matrices of both groups of teachers, it was concluded that student growth in attitude and achievement was not specifically the result of indirect behavior alone. In fact, the two groups of teachers were not substantially different in their indirect verbal behavior. The percentage column totals for indirect behavior (categories 1, 2 and 3) were 11.43 for the indirect and 11.06 for the direct teachers. They differed in their direct classroom behavior (categories 6 and 7) with the direct group exhibiting twice as much direction activity and five times as much criticism as the indirect teachers. The direct teachers also spent more time giving extended directions and criticism. Thus, it was further concluded that the critical element responsible for differences in student growth in cognitive and affective process development was the excessive, or lack of excessive, use of negative verbal behaviors on the part of classroom teachers.

## ABSTRACTOR'S ANALYSIS

This study represents one of several in which attempts have been made to find relationships between science teacher behavior as measured by FSIA and student achievement and/or attitude. At least 12 were reviewed by Evans in the 1973 Yearbook of the Association for the Education of Teachers in Science (3). Although Campbell reported a positive relationship between indirect teaching methodology and student performance, his results have not been consistently supported by other investigators. Unfortunately, the results of teacher effectiveness research in science education remain contradictory and inconsistent even though category systems have been used to quantify classroom dialogue (3).

Overall the report was well-written. The problem, methodology, analyses, findings, and conclusions were clear and to the point. Unlike many reports involving the FSIA, the category system was only briefly described, but the report spelled out the method for calculating the revised i/d ratio in detail. This latter procedure is particularly important, because the revised i/d ratio can be calculated in one of two ways depending upon which observer manual the researcher was following. In additon to the method described in the report, the revised i/d ratio can be obtained by adding the total of tallies in categories 1, 2, and 3, and dividing by the totals in categories 1, 2, 3, 6, and 7 (2). Simply presented revised i/d ratios without the method for calculation has little meaning and cannot be properly compared with findings of other investigations involving the ratio.

A further merit of the research was the length of the observation sessions. Campbell reported that each taping session was 45 minutes in length. The entire science lesson was apparently analyzed, although the report does not specifically comment on this point. Regardless, a 45-minute observation is more likely to include teacher statements at the beginning and end of a lesson. Such statements frequently set the emotional climate for the entire lesson and for lessons that follow. A 45-minute observation period provides a more accurate description of classroom dialogue, and it represents an improvement over the usual 10- to 20-minute sample taken during the middle of a lesson that seems to be the standard for most researchers who use the FSIA.

The report would have been improved and easier to analyze if it had contained more information about the teacher, student, observation, and instructional variables. How were the teachers selected, and what were their general characteristics? What instructional materials were used in the classrooms? What topics were covered in the science classes? What were the sizes and locations of the schools? Were the observations made at random throughout the school year and day? Did the teachers know in advance when the observations were to be made? What were the student characteristics besides IQ and pretest scores? A report should include this type of information or a statement that the variables were examined and found to be similar for both groups. Such information is especially necessary when randomization has not been under the control of the researcher. Lack of the information makes it impossible to determine whether or not the variables offer plausible hypotheses for explaining differences in student performance that rival the stated effects of the indirect teaching methodology.

Additional information that would have been useful in the analysis of the report includes the revised i/d ratios of the individual teachers, combined matrices of each group, rationale for using the ratios while omitting other available measures of indirect and direct classroom behavior, and a coefficient of inter-observer agreement. The mean revised i/d ratios were presented and differences between the groups had been determined by analyzing the combined matrices of both groups. As a result, we cannot be certain of the exact differences between the groups of teachers. For example, they could have differed significantly as a result of the tallies in the lecturing, silence and confusion, asking questions, or student talk categories. Another possibility was that one or two very indirect and one or two very direct teachers were involved in the study. The extremes could cause the group means to be significantly different, and yet, a majority of the teachers might still have been quite similar in their indirect or direct teaching methodologies. Inclusion of the individual revised i/d ratios and the combined matrices of each group would eliminate the question of whether or not the individuals and groups of teachers were different in their indirect and direct classroom influence.

The reason for requesting a rationale for using only the revised i/d ratio in the analysis is that it does not include lecturing and questioning behaviors. The ratio is associated with control and motivation rather than with the presentation of subject matter. One might suspect that the revised i/d ratio would be related to attitude. but achievement might be more closely related to lecturing and asking questions. The fact that some relationship appears logical does not mean that it actually exists, but a rationale for not examining the relationship seems reasonable, especially when the data were already available.

The establishment of inter-observer agreement is a crucial task, and one that should be accomplished before, and checked periodically during, all research involving systematic observation of classroom behavior. A coefficient of inter-observer agreement of .9 and above is not unusual after a fairly short training period in the use of FSIA. It provides an indication of the objectivity of the observational technique and a measure of how reliable the observer has been in encoding behaviors. Inter-observer agreement and the method of determining a coefficient were not mentioned in the report. No doubt this was an oversight, or perhaps it was omitted by the journal editor. Most researchers using systematic observation of classroom behavior are aware of its importance. Measures of the independent variables with the FSIA are used if a coefficient is not established.

The conclusion regarding the effectiveness of indirect teaching methodology and student attitude was logical based solely on an analysis of posttest or adjusted mean posttest scores, but examination of both pretests and posttests provides a different picture. Significant differences between the groups were the results of losses, not gains, on the criterion instruments. The indirect group mean increased slightly from 5.01 to 5.10 between the pretest and posttest on the attitude scale, but a majority of the difference resulted from a decrease in the direct group mean from 5.01 to 4.83. The indirect group mean went from 26.93 to 26.34 on the curiosity inventory, while the direct group mean went from 27.25 to 22.97. Both groups of students decreased in scientific curiosity with the indirect group decreasing significantly less than the direct group. From these data it is tempting to conclude that no science teaching would be more effective for affective process development than was either an indirect or direct teaching methodology; unfortunately, data from a control group who did not receive science instruction are not available.

Significant differences resulting from losses in positive attitude by treatment and control groups, with the control group losing significantly more than the experimental group, are fairly common. They point out the need for additional research on attitudes and the relationship between student attitude and instruction. Such findings question the use of a posttest only control-group design in student attitude research when the treatment is simply an alternate instructional strategy, and they indicate the need for comparison data on students who have not received instruction in the subject area under consideration (3).

A critique of any research involving the FSIA would be incomplete without mentioning a few of the major limitations of the observational system. One major limitation is that the system is based on a questionable assumption; i.e., a sample of verbal behavior is an adequate sample of a teacher's total classroom behavior. It has been shown that a major portion of a teacher's classroom behavior is nonverbal (5). Nonverbal behaviors accent, illustrate, coincide with, substitute for, augment, and contradict verbal behaviors (4), and evidence exists indicating that they are better communicators of emotion and attitude than verbal behaviors (9, 10). These findings suggest that nonverbal behaviors are too important to be excluded from any accurate description of classroom behavior.

A second limitation is related to the large number of ground rules used in encoding the observations. Ground rules improve inter-observer agreement, but they can provide a distorted description of what actually took place in the classroom. A third limitation is that the FSIA does not account for differences or extremes within each category. A distinction is not made between silence and confusion, and both mild and vehement criticism are simply recorded as criticism. The fact that the system is not appropriate for certain classroom activities is a fourth limitation. Among the activities that must be excluded are student laboratory work, small group projects, individualized work, and teacher use of instructional materials and strategies which do not require teacher talk.

## REFERENCES


1. Amidon, Edmund, and Ned A. Flanders. "The Effect of Direct and Indirect Teacher Influence on Dependent-Prone Students Learning Geometry." Journal of Educational Psychology, 52:286-291, 1961

2. Amidon, Edmund, and Ned A. Flanders. The Role of the Teacher in the Classroom: A Manual for Understanding and Improving Teacher Classroom Behavior. Minneapolis: Association for Productive Teaching, Inc. 1967.

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

4. Eckman, Paul. "Communication Through Nonverbal Behavior: A Source of Information About an Interpersonal Relationship." In Affect, Cognition, and Personality, S. S. Tomkins and C. E. Izard (eds.). New York: Springer Press. 1965.

5. Evans, Thomas P. "Teacher Verbal and Nonverbal Behaviors and Their Relationship to Personality." Journal of Experimental Education, 38:38-47, 1969.

6. Flanders, Ned A. Teacher Influence, Pupil Attitudes and Achievement. Cooperative Research Monograph No. 12 (OE-25040). Ann Arbor: The University of Michigan, School of Education. 1965.

7. Gage, Nathaniel L. (ed.). Handbook of Research on Teaching. Chicago: Rand McNally and Co. 1963.

8. LaShier, William S. "An Analysis of Certain Aspects of the Verbal Behavior of Student Teachers of Eighth Grade Students Participating in a BSCS Laboratory Block." Unpublished doctoral dissertation, The University of Texas. 1965.

9. Mehrabian, Albert, and Susan R. Ferris. "Inference of Attitudes from Nonverbal Communication in Two Channels," Journal of Consulting Psychology, 31:248-252, 1967.

10. Mehrabian, Albert, and Morton Wiener. "Decoding of Inconsistent Communications." Journal of Personality and Social Psychology, 6:109-114, 1967.