DOCUMENT RESUME

ED 209 105 SE 035 860

AUTHOR	Blosser, Patricia E., Ed.; Mayer, Victor J., Ed.
TITLE	Investigations in Science Education, Vol. 7 No. 4. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION	Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	81
NOTE	67p.
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (subscription $6.00, $1.75 single copy).
JOURNAL CIT	Investigations in Science Education; v7 n4 1981
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Abstracts; Cognitive Ability; *Cognitive Development; College Science; Concept Formation; Elementary School Science; Elementary Secondary Education; Higher Education; *Learning; *Measures (Individuals); Research Methodology; Science Course Improvement Projects; *Science Curriculum; *Science Education; Secondary School Science; Student Attitudes; Student Interests; Teacher Attitudes
IDENTIFIERS	*Science Education Research

ABSTRACT

Presented are analytical abstracts, prepared by science educators, of four research studies classified as relating to science curricula, four studies about learning, and one study on the evaluation of an instrument designed to measure students' orientation toward science. Three of the curriculum studies each relate to a specific curriculum project or course while the fourth placed in this group (Israeli pupils' understanding of the mole concept) because the abstractor suggested the necessity of knowing more about the curriculum within which the concept of the mole was presented. One of the learning studies investigates Piaget's ideas of cognitive development while the other three relate to a specific learning theorist, namely, Gagne, Guilford, and Ausubel. The final analysis describes an investigation of the empirical validity of Meyer's Test of Interests. Each abstract includes: (1) bibliographic source of the research study; (2) summary of the purpose, rationale, design, and procedure of the research; and (3) the abstractor's analysis. (DS)

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC Clearinghouse for Science, Mathematics
and Environmental Education

Published Quarterly by

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

Subscription Price: $6.00 per year. Single Copy Price: $1.75
Add 50¢ for Canadian mailings and $1.00 for foreign mailings.

Volume 7, Number 4, 1981

NOTES FROM THE EDITOR

The final issue of Volume 7 contains four studies classified as relating to science curricula, four studies about learning, and one report on the evaluation of an instrument designed to measure students' orientation toward science.

Three of the curriculum studies each relate to a specific curriculum project or course. Charles investigated the Nuffield combined science course. Markle and Capie evaluated a course entitled Physics for Elementary Teachers. Whittaker looked at Science 5/13, a British primary school science curriculum. The fourth study, by Novick and Menis, was placed in this group not because of the topic investigated (Israeli pupils' understanding of the mole concept) but because of the abstractor's remarks about the necessity for knowing more about the curriculum within which the concept of the mole was presented.

One of the four learning studies, that by Shyers and Cox, is yet another investigation of Piaget's ideas of cognitive development: the concept of proportionality as held by college students. The other three each relate to a different learning theorist. Beeson worked with Gagne's learning hierachies as these relate to learning about electric circuits. Ukens and Merrifield used Guilford's Structure of the Intellect model with a COPES learning sequence. Murray looked at cognitive structure (Novak-Ausubel) and achievement in college biology.

The final analysis describes an investigation, by Hofstein et al, of the empirical validity of Meyer's Test of Interests, and instrument that has been used with English students and with Israeli students.

Patricia E. Blosser
Editor

Victor J. Mayer
Associate Editor

CURRICULUM

Charles, D. J. "Nuffield Combined Science--An Evaluation." <u>School Science Review</u>, 58(202): 129-134, September, 1976.
Descriptors--Course Evaluation; *Curriculum Evaluation; *Educational Research; Elementary Secondary Education; Elementary School Science; *Evaluation; Science Education; *Secondary School Science

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

Purpose

The aim of this study is to address two questions dealing with the Nuffield combined science curriculum at the Exmouth School in England:

1) How suitable is the Nuffield combined science course for students aged 11-13 of widely differing ability?

2) Is it really excellent education in and through science or just expensive enjoyment?

Rationale

The Nuffield combined science curriculum, according to Charles, seems more appropriate for students of above-average ability. This study is intended to check the claim that it can be adapted to students "with the whole range of ability."

Evaluation Design and Procedures

Charles presents an informal evaluation model--identify ideal science skills and attitudes, relate skill and attitude emphases of the curriculum to the ideals, and assess skill and attitude change resulting from the curriculum. Four activities represent the evaluation procedures:

1) staff discussion to assess desired content, identify difficult concepts and skills, structure parts of the course, and reduce the variety of instructional strategies;

2) a staff questionnaire to appraise resource materials and skills and attitude training;

3) a student questionnaire to determine interest, difficulty, and enjoyment of each unit in the course; and

4) coded response tests to assess student skill and attitude attainment.

Findings

Several findings are mentioned in the report:

- The course had the desired content; difficult concepts and skills were identified; and parts of the course were found to need structuring.
- The Teacher's Guides were rated as good; and a wide variety of instructional approaches were reduced.
- Growth and reproduction units were most popular; energy was least popular; and student activity books were not as helpful as expected.
- A large proportion of course content was judged suitable for a wide ability range, but students at the extremes of the ability range became increasingly frustrated during the second year of the course.

Interpretations

Charles concludes that most of the Nuffield combined science course is "both suitable and good for approximately the upper 75 percentile of the ability range."

ABSTRACTOR'S ANALYSIS

The adequacy of evaluation can be assessed by answering the following question, a modification of the Evaluator's Question originally posed by Lindvall (1966):

> Does this innovation, in our situation and as determined by our means of appraisal, do what is desired better than alternatives?

1. Innovation refers to what is evaluated. A description of the implementation, content, resources, and presentation modes, for example, provides the basis for an understanding of what is evaluated. The only descriptor provided by Charles is "Nuffield combined science." To many science educators, this identifies the general content but to many others, it does not. Furthermore, nothing is said about the adaptations required by the school during their adoption of the program. Seldom is a general curriculum program successfully adopted without changes made to satisfy the unique needs and resources of individual schools. Thus, we have a general idea of what was evaluated, but we don't know its specific operational parameters.

2. The situation refers to generalizability. A description of the environment in which the activities take place is needed to form the basis for comparisons to other schools interested in or using the program. Staff disciplines, lack of "creaming," and proportion of attendance at direct grant schools in Exmouth's catchment area are mentioned, but nothing is offered to describe student characteristics, funding, teacher relationships, school leadership, facilities, training, or parent/community collaboration. Unless we have a personal knowledge of the Exmouth School, we can't generalize the findings of this evaluation to other schools without knowledge of the environmental situation.

3. The means of appraisal refers to the evaluation model, its objectives and process. Although not specifically identified, the apparent evaluation model could be classified as objectives-referenced (the curriculum was assessed for skill and attitude changes in participating students). A more systematic and powerful evaluation model, such

as a control group design or a special regression design, however, would be more appropriate. It isn't clear what independent variables are included or what the dependent variables are in the report.

4. The what of the Evaluation Question is an indication of the success of the innovation. Is the program suitable? Is it excellent? Suitability and excellence, assumed to be the two primary dependent variables, are not defined in the report. Specific data on student achievement stratified by age levels and ability levels are lacking. Teacher and student ratings are not standardized to any normative response base (perhaps satisfactory or good ratings represent low percentile rankings in a normative distribution). Statistical analyses are missing--only subjective data from discussions and questionnaires are provided. The report includes little information upon which to determine what actually happened in the program.

5. The better part of the question also refers to success, but is specific to the level of success. It deals with the relative degree of accomplishment. How suitable is the program? How excellent is it? Since actual success hasn't been adequately documented, we can't deal with level or degree.

6. Alternatives are other activities, programs, or improvements in existing curricula. Charles does not deal directly with this. There seems to be more interest in confirming or validating the use of Nuffield combined science in the Exmouth School than in comparing its use to alternatives.

Summary of Analysis

Overall, I'm disappointed with Charles' report. An evaluator should be objective, yet Charles indicates a favorable bias toward the curriculum in the introductory paragraphs (..."in my view is well worth it"). Little information is provided to adequately describe the evaluation and its implementation, and, therefore, the Evaluator's Question cannot be answered. The evaluation process may have been adequate to deal with specific local concerns, but I reluctantly suggest the report is inadequate as a contribution to the professional literature.

REFERENCE

Lindvall, C. M. "The Task of Evaluation in Curriculum Development Projects: A Rationale and Case Study." The School Science Review, 74: 159-167, 1966.

Markle, Glenn and William Capie. "Assessing A Competency-Based Physics Course: A Model for Evaluating Science Courses Servicing Elementary Teachers." Journal of Research in Science Teaching 14(2): 151-156, 1977.

Descriptors--College Science; *Curriculum Evaluation; *Educational Research; *Elementary School Teachers; Higher Education; Performance Based Education; *Physics; *Preservice Education; Science Education; *Teacher Education

Expanded abstract and analysis prepared especially for I.S.E. by Leon Ukens, Towson State University.

Purpose

The purposes of this report were to describe and report on the evaluation of a course, Physics for Elementary Teachers, in the areas of 1) understanding selected physical concepts, 2) understanding the nature of science and science processes, and 3) attitudinal changes toward science as a result of the course. Physics for Elementary Teachers is described only as a self-paced, activity-oriented course. It is, however, footnoted in another journal.

Rationale

The Commission of Science Education of the American Association for the Advancement of Science in a preliminary report identified 31 competencies needed by elementary science teachers. These were classified into five categories: scientific inquiry, attitude toward science, processes of science, scientific knowledge, and continuous learning.

Previous science education research has shown relationships between teachers' knowledge and attitude and 1) effort exerted to teach science, and 2) teaching style. This study states that course evaluation should include three components: 1) a content one, 2) a process one, and 3) an attitudinal one. This fits nicely into what previously cited reseatch has reported.

Research Design and Procedure

Three components of the course, Physics for Elementary Teachers, were evaluated utilizing the one group pretest-posttest design. The component "understanding selected physics concepts" was evaluated by using a 44-item multiple choice test developed by the authors. Content validity for this test was established by members of the physics department and science education department at the University of Georgia. The KR-20 reliability estimate of this test was 0.73.

The component "understanding the nature of science and science processes" was evaluated using the Wisconsin Inventory of Science Processes (WISP) and the Welch Science Process Inventory (SPI). Validity of these two were determined in earlier studies. KR-20 reliability estimates were 0.82 for the WISP and 0.79 for the SPI.

The component "attitudinal changes toward science" was evaluated using a subject preference survey. Six science and four nonscience subjects were paired in all possible combinations and the students circled their preference. Validity and reliability statements concerning this measure were also used.

Each of the above evaluations was given as a pretest and again as a posttest. Correlated t-values for differences in individual scores were the analysis used with the first variable, correlated t-values for matched cases were used for the second variable and t-values for matched scores were used for the third variable.

The evaluation of the course was done for three consecutive quarters at the University of Georgia. Instructors for the courses were not mentioned.

In addition to the above three variables, the authors reported some additional data involving the students' perceptions of the course, Physics for Elementary Teachers. An adjective checklist was administered at the end of each quarter. This checklist was developed by the authors and was well grounded in research literature.

Findings

For the variable, understanding selected physics concepts, the average gain from pretest to posttest was approximately 12 points and the correlated t-value for differences in individual scores was significant beyond the 0.001 level.

For the variable, understanding the nature of science and science processes, the mean score increased by 6 points on the WISP and 5 points on the SPI and the correlated t-value for matched cases was significant beyond the 0.05 level for both tests.

For the variable, attitudinal changes toward science, the t value for matched scores was significant beyond the 0.01 level. Also, student reactions to the course showed the course to be favorable, beneficial, worthwhile, discovery-oriented, understandable, effective, and useful.

Interpretations

Results of the evaluation showed that this self-paced, activity-oriented course increased content knowledge, increased the understanding of the nature of science and science processes, and improved the attitudes of students toward science. Evaluations of this type could serve as a model to help identify strong and weak points of other science courses for preservice elementary teachers.

ABSTRACTOR'S ANALYSIS

Frequently courses are given with little regard to their effect on variables such as understanding of science processes and attitudes toward science. Such courses are usually evaluated concerning only the knowledge of concepts developed. This research report provided a model by which other such courses for preservice elementary teachers could be evaluated regarding not only the variables of knowledge, but also the variables of processes as well as attitudes. Enough information was given to allow other science educators to do likewise. One gets the feeling while reading the report that the evaluation process was not overly time-consuming. This lends credence to having a rather comprehensive type course evaluation of this sort.

There are some data which would have been helpful if reported, especially for researchers interested in the "high marks" the course received. For example, there are other physics courses at the University of Georgia for preservice elementary teachers. Were they evaluated using the same methods? Were the students who enrolled in the evaluated course selected in any way? Was the course elective? How many credits was it? Granted, these questions are probably of greater concern to the science educator attempting to put theory into practice rather than the science educator interested mainly in research, that is those interested more in the course than in the evaluation. Very little information is given concerning the course, although it has been written up in another journal and referenced in this report. This is understandable in this type of journal since the focus of the article is on the evaluation process and not on what was being evaluated. I would imagine there is much more to the course than being simply self-paced and activity-oriented. How rigorous was the course? What types of grades were received by the students? Would any self-paced, activity-oriented course fare as well?

It is easy to question the number of students involved with the research. The content exam was administered to 34 students, which, spread over three quarters, averages less than 12 students per class. The WISP was given to only 9 students, while the SPI was given to 24. No mention is made of this discrepancy in numbers or of the significance of the small class sizes. Questions would also have to be raised regarding the design of the study if the authors were advocating the course being evaluated. These questions are of little significance when one realizes that the function of the report is to share a model of course evaluation and not to dwell upon the course being evaluated.

One apparent typographical error exists when the authors describe disadvantages (apparently meaning advantages) of the attitudinal changes toward science since all items listed are advantages.

The authors do an excellent job in developing the research rationale and are apparently "research wise" in reporting their study.

How helpful this type of research would be to others depends upon whether researchers want to use the evaluation model developed for evaluating their courses (very helpful), or if they want to compare courses (not very helpful), or if they want to develop a similar type course (not very helpful).

Novick, S. and J. Menis. "A Study of Student Perceptions of the Mole Concept." Journal of Chemical Education, 53(11): 720-722, 1976.
Descriptors--*Chemistry; *Cognitive Processes; Cognitive Tests; *Educational Research; *Scientific Concepts; Science Education; Secondary Education; *Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by Elizabeth Kean, University of Wisconsin.

Purpose

The stated purpose of this study was to learn the nature of some Israeli high school pupils' understanding of the mole concept, as a result of their study of this concept in courses produced by the Israeli Science Teaching Center.

Rationale

Since the mole concept is central to comprehension by students of the interaction between microscopic and macroscopic interpretations of the world, it is important that they learn this concept well. Previous analysis of this concept and some experimental studies, as well as the authors' own personal experiences, have suggested that the mole concept and its application are inherently too difficult for the "average" 15-year old. The assumption apparently is made that if students do not succeed in learning the mole concept after it is introduced in their courses, that the failure to learn is a result of the inherent difficulty of the concept.

Research Design and Procedure

This investigation sought to uncover the nature of students' learning of the mole concept after their exposure to a particular learning situation(s). No pretesting was apparently done; results were based on a single, verbally administered post-test.

Instrument. An interview instrument was designed to test students' knowledge of 16 (unidentified) cognitive understandings and skills subsumed under the heading of the mole concept. In the article, the 23 questions of the interview instrument were sorted as belonging to six major categories. Only three of the 23 questions were included in the article; the others were not further described.

Questions for the instrument were generated after analysis of an undescribed multiple choice achievement test administered to an undescribed pilot group of 150 students. This test apparently covered the mole as well as other content in the Israeli Science Teaching Center curriculum. In addition, the authors developed a "Gagné-type" analysis of concepts and skills needed to solve a "typical" stoichiometry problem. The procedures by which questions of the interview instrument were generated from these preliminary activities were not specified. The instrument was pilot-tested with five (undescribed) students and subsequently revised.

Interview Procedures. A single interviewer (one of the authors) apparently interviewed 29 students (the number of students had to be ascertained by counting responses to questions as reported in the article). Not all questions were asked of all pupils. The basis for deciding which questions were to be asked of which students, and the mean number of questions asked were not reported. The interviewer rephrased questions to students to ensure their understanding of the intent of the questions when this seemed appropriate. All students were given an IQ test the day before they were individually interviewed.

Interviews were taped; tapes were analyzed by assignment of student responses to categories (expected, i.e., correct answer; specific wrong answers; miscellaneous incorrect answers; "don't know"). It was not specified whether the categories were determined *a priori* or were generated during the analysis process.

The description and location of the school classes from which students were interviewed, the precise age of students, their previous science backgrounds, selection criteria for students, the time of year, duration of the interviewing process for all students, the mean length of the interviews, the time between interviewing and student exposure to the material, the nature of the teaching process by which students were expected to learn the mole concept--were not specified in the article.

Findings

A 0.69 correlation was obtained between IQ test results and percent score on the interview instrument. However, since not all students were tested on every question, it is unclear how the mean (percent) was obtained. Some possibilities include: number of correct items per total number of questions or total number of questions asked; number of correct cognitive understandings and skills per total number of cognitive understandings and skills, etc.

Students were broken out into three groups on IQ levels; 71-85; 86-100; 122-123. The latter group may represent a typographical error or may represent a very small group with a limited IQ range. The number of students in each group was not reported. Mean IQ for the entire sample and standard deviation were not reported.

Three of the nine questions dealing with "mole definition" and student responses to them were reported in detail. The authors speculated upon the sources of student misconceptions.

The authors' analysis of all interview responses revealed three main misconceptions:

1) the mole is a certain mass and not a number;
2) the mole is a certain number of particles of gas;
3) the mole is a property of a molecule.

Interpretations

The authors speculated briefly on the sources of the first two misconceptions that arose from the curriculum. They then concluded that the results tended to support the theses that most 15 year olds in Israel do not achieve a "coherent" understanding of the mole concept, and cannot effectively use it to solve problems. They tentatively concluded that results indicated that many students do not function at the cognitive level appropriate for such concepts. They also stated that a simpler and less involved treatment of this complex subject might result in less misconceptions.

ABSTRACTOR'S ANALYSIS

The procedure which the authors devised apparently enabled them to define adequately the domain of the mole. The interview was actually a verbal achievement test designed to test whether students had acquired the various parts of this concept.

By using an interview technique to probe students' acquisition of the mole concept, the authors avoided confounding results because of differences of student reading levels, and a host of other semantic factors. When students did not understand a question, the interviewer could rephrase until the student perceived the intent of the question. On the other hand, this process did erode the standardization in testing. Given the paucity of information on how student responses were graded as correct or incorrect, it is impossible to judge the effect of this lack of standardization.

This is a study which measured the degree of learning of a concept. The questions should be asked: What are the reasons why students failed to learn this concept? The authors imply in the introduction and conclusions sections that there are three possible causes: inherent level of difficulty of the concept, lack of appropriate student developmental level, or method of presentation of the concept.

The criteria for characterizing levels of difficulty of subject matter are ill defined. What makes a concept difficult: abstractness, lack of perceptible examples, number of interrelated subordinate concepts that make up the concept? There is a large body of laboratory-based research on concept formation and concept learning, but, in general, there is a lack of theoretical constructs which would enable one to say unambiguously and with ease which concepts are inherently more difficult than others, and why.

The invocation of Piagetian terms to explain why students did not learn the mole concept implies that the lack of learning was related to the developmental level of students. The formal operational stage of development is characterized by the ability to engage in hypothetical-deductive reasoning. The ability to do this type of reasoning is evidence that this stage of development has been reached. However, the converse is not necessarily true. It is not certain that if students do not exhibit this type of thought process or do not acquire formal concepts that the problem lies with their developmental level. It may be that they lack certain background information that prevents them from learning such material, or that they are being required to learn material so rapidly that they do not have an opportunity to internalize it, or simply that they have never before been asked to use this type of thought process in certain types of academic work. This study did not measure developmental levels. It measured only acquisition of concepts.

The correlation between IQ and acquisition of the mole concept likewise does not help the reader to distinguish between student developmental effects and level of subject matter difficulty effects. What precisely does an IQ test measure anyway?

The third factor which could affect student learning of the mole concept is the curriculum used to teach the concept. Without having a copy of the curriculum to view, it is impossible to state precisely why difficulties might have arisen. The reader does not know the time frame allocated to students for mastery of the mole concept.

However, a listing of the contents of Chapter 2 (in which the mole concept is presented) is appallingly long. The mole concept may have gotten lost in the mass of information presented.

Could it also be that the difficulty in learning is due to problems of speed of processing and internalizing information? The level of misconception may be related to the method of presentation and to an inordinately high information density of abstract concepts that prevent students from learning those concepts precisely.

The description of the structure of the presentation of the mole concept in Chapter 2 showed that a historical development was apparently used to develop the topic. Such a historical approach may actually obscure the critical features of the concept, as the authors of the curriculum attempt to repeat (but incompletely) the logical process that led to the conceptualization of the mole.

Thus, in a study such as this, it is impossible to separate subject matter difficulty from methods of presentation from developmental level of students as being the cause of student misconceptions. If this type of study is inappropriate for distinguishing causes of student misconceptions, what can it tell us? Is it worth doing?

As the authors state in their Results section, the real value of the interview instrument is in exposing the nature of students' misconceptions about the mole. It seems likely that part of that difficulty is due to the method of presentation of the concept. In exposing the nature of the misconceptions, the authors are in a position to begin a process of curriculum reform, or the preparation of supplementary materials to overcome the deficiencies of the curriculum in use. To be used in this way, the article should have been more explicit about the specifics of the existing curriculum. What exactly did the authors expect students to learn about the mole? What are the boundaries of the mole concept? They mention 16 cognitive skills and understandings belonging to the concept of the mole. What are these? Is there a hierarchy among them? How exactly was this content presented to students? Were they just turned loose with the text? With media? Were teachers involved? What did they say or do to supplement the text? We are not told.

This is not an empirical study which implies generalizability. "Can it be replicated with other students?" is not an appropriate question to ask. Were it to be used as empirical research, additional types of information (such as those listed in the Procedures section) would have had to have been included in the article. However, teaching is, in many ways, an art. The ability to move inside a student's mind, to look at the relationship of student learning to presentation of that material, to explore ways of developing more effective learning materials or procedures--these are still not scientifically established procedures and skills. They explore the question of "What have kids learned or not learned from our curriculum?" This article had the potential for showing almost in a case study format a process by which the gap between what teachers want students to learn and what students do learn could have been narrowed.

Whittaker, Muriel. "An Investigation into Teacher Attitudes to Objectives for Primary Science Teaching." School Science Review, 58(203): 360-366, December 1976.

Descriptors--*Curriculum Development; *Education Objectives; *Education Research; *Elementary School Science; Elementary School Curriculum; *Primary Education; *Science Curriculum; Surveys; *Teacher Attitudes

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Marvin Bratt, The Ohio State University.

Purpose

The purpose of this study was to determine attitudes toward science objectives found in Science 5/13, a set of British materials designed for primary school. Attitudes were inferred on the basis of responses to a 10-item survey distributed and completed by 273 teachers. The following hypotheses are inferred from the study, although not so stated:

1. Data in this survey should correlate with data in a previous study completed by Tabbron (1969).
2. Primary school teachers are more prone to choose objectives related to children's activities than to the child's ability to evaluate results.
3. Primary teachers consider children to develop reasoning ability in a hierarchical fashion rather than teaching as they were taught.
4. Primary teachers are more concerned with successful completion of science activities than with having a child withhold judgment in light of alternative explanations.
5. Primary teachers view science as discrete activities rather than a chain of thought and investigation.
6. Teachers at lower (1-3) primary levels will choose different clusters of objectives than will those teaching higher primary levels (4-6).
7. The relative importance of "interpretation of data" correlates with whether the teacher had a science curriculum course. This correlation will be positive for teachers who considered the course valuable.

Rationale

In England, there is apparently considerable evaluation of science education curriculum in the schools. The model chosen was identical to Tabbron's (1969) with the exception that mailed surveys were used in lieu of personal interviews. It is a straightforward interpretation-analysis of survey data. There was no further mention of a model for data collection. The investigators (Whittaker and Tabbron) both assumed that "...if a teacher chose an objective as important, they expect a child to be able to achieve the objective." Underlying assumptions include: (1) that rank-order correlation techniques can be used to suggest the direction of a teacher's attitude, (2) cluster analysis is a method of interpreting differences among ranks, and (3) that teachers' attitudes toward science may be related to Piagetian levels of development as perceived by the individual teacher.

Previous research (Crossland, Shayer, Tabbron, and Jungwirth) was discussed in light of the child's developmental level with respect to the level of reasoning and/or sophistication in terms of science concepts. Harlen's study utilizing cluster analysis was mentioned.

Research Design and Procedure

The design is simple, a one-shot completion of a survey. The survey was completed by 273 teachers although it was mailed to all teachers in Derbyshire who taught in the primary grades. The teachers were asked to select the six most important objectives from a list of ten (the same survey used in the Tabbron study). Although not asked to rank the objectives, many teachers did so. Teachers also completed a short questionnaire (five questions) asking how science was taught by them. The data are then tabulated as to the number of choices of each objective. These were compared to the ranks found in Tabbron's study. In order to examine levels of choice among the sample, a post hoc cluster analysis was performed on the choices of the members of the sample. Data were presented in tabular form illustrating comparisons with Tabbron's study and the cluster analysis study.

Findings

1. A clear correlation ($p < 0.01$) was found in the data for primary teachers and the data in the Tabbron study of Nuffield science teachers.

2. A X^2 test of objectives 10 and 4 and 7 and 9 was significant ($p = 0.01$).

3. 256 of the primary teachers were doing some science teaching.

4. Cluster analysis provided two clusters of respondents, one with 142 members and one with 118 members. These clusters were significantly different from one another (polar), $p = 0.01$. Sixty-six percent of infant teachers were in Cluster 1. Cluster 1 teachers ranked observation and active exploration (observe, experiment) as most important, Cluster 2 teachers ranked orderly and critical handling of information and materials (i.e., organize and classify, communicate) as most important.

5. Experimentation and communication ranked high, interpretation and withholding judgment ranked low.

6. Infant-level teachers chose objectives differently than junior level teachers, according to analysis of cluster data during cluster formation.

Interpretations

1. An inference concerning the training of primary teachers was made in relation to science courses. The author comments, "...it was clear that a substantial number [of subjects] had not had a science course."

2. Teachers regard low-scoring objectives as "less likely to be attained," rather than as unimportant.

3. Primary teachers attach significance to basic skills in science education.

4. Some "increased understanding" of science has occurred in 1973 as compared to 1968.

5. Primary teachers attach more importance to children doing activities rather than their evaluating results.

6. Children and teachers find it hard to accept the concept of approximation in terms of right and wrong.

7. Primary teachers do not expect children to follow chains of thought or investigation, rather they view science for this age child as a series of discrete activities and/or ideas.

8. Cluster 1 teachers believe that children are in late intuitive-early concrete operations stage, whereas teachers in Cluster 2 teach children at the late concrete-early formal operations stage. Therefore, the dichotomy between "activity" and "organizing" objectives suggests the consideration of developmental level.

9. Teachers who had curriculum courses view "interpretation of data" as important.

ABSTRACTOR'S ANALYSIS

The author remarks in the closing statements of the article that the discussion of results is speculative and subjective. The reader is told the number of respondents (273) but not the total number of surveys mailed out and therefore has no idea of what percent of the population this sample represents. No demographic data were included and the reader is left to sort out how various groups of teachers may have responded. Since none of the hypotheses were stated, the reader must dig them out of discussion throughout the article. It appears as though the hypotheses were generated after the data were analyzed without much prior thought. Certainly such prior thought could have resulted in a much stronger study.

How does this study fit into the matrix of other attitude studies in science education? It seems clear that investigators are interested in variables associated with teacher attitudes (Moore and Piper, 1977). If it were possible to generalize at all from data provided here, and if the study had reported the strength of the correlations, one might be able to draw more significant conclusions. But there are other problems associated with this study. For example, this study purports to investigate teachers' attitudes toward science objectives, but does so by inference from survey data and not actual measurement. No attempt has been made to ascertain the validity or reliability of the instrument, although the agreement (positive correlation) with Tabbron's study suggests some reliability even though the coefficient is not given. Speculation is not considered valid. Interpreting the results obtained by asking subjects to choose objectives from a list does not seem to be the most efficient method to use to infer the directionality of attitude. A question should be raised as to the nature of self reporting of science teaching by elementary teachers since data from studies in the USA (Krockover, et al., 1972) indicate that an overwhelming number of elementary teachers do not teach science.

Are there new conceptual contributions in this study? Perhaps more questions are raised by the speculation on the author than are answered. The lack of science training of elementary (primary) teachers squares with data available in this country. There is little prior research to suggest how teachers perceive their students' attainment of objectives. This would be a good question to study further. It would be interesting to ask teachers in the USA to complete the survey.

Rank order correlation techniques have been used quite often to suggest the direction of attitudes toward science. Cluster analysis of choices is not new either. The assumption that teachers' attitudes toward science may be related to Piagetian levels of intellectual development is supported in a study by Lawson, et al. (1975) who found that Separation and Exclusion tasks loaded highly with teacher attitudes as measured by the Bratt Attitude Test (BAT). They suggested that as persons develop abilities to successfully respond to the Piagetian tasks, their attitude toward science and science teaching improves. The interpretation that primary and intermediate teachers tend to choose different clusters of objectives is supported by Lazarowitz, et al (1978) who found that elementary teachers are more child-centered while secondary teachers are more subject-matter oriented. These researchers also suggest that science methods courses can be productive in changing teacher attitudes toward science.

In summary, the study is difficult to interpret since it does not fit the conventional style used in this country. The problem and related hypotheses are buried in discussion and analysis. The sampling method is not adequately discussed. Although a large number of subjects were polled, it is not possible to identify the characteristics of the sample, where the teachers taught, etc. Underlying assumptions are not stated, with one exception. The design is a "one-shot" approach and the data probably nominal or ordinal at best. There is no report on reliability or validity of the survey. The literature review is very abbreviated and restricted. Therefore, the design, although simple, is confined to post hoc analysis. The findings should be limited to the sample studied at this point in time. Some interpretations are supported in other studies but the majority are purely speculative. There are a number of interesting problems suggested by the study for further research.

REFERENCES

Krockover, G. H. et al. "Teaching Science in the Elementary School: At This Point in Time, 1971-72." Paper presented at the annual meeting of the National Science Teachers Association, New York, 1972.

Lawson, Anton E. et al. "Relationship of Formal Reasoning to Achievement, Aptitudes, and Attitudes in Pre-Service Teachers." Journal of Research in Science Teaching, 12(4), 1975.

Lazarowitz, R. et al. "Student Teachers' Characteristics and Favorable Attitudes Toward Inquiry." Journal of Research in Science Teaching, 15(6), 1978.

Piper, Martha K. and Kenneth D. Moore (eds.). Attitudes Toward Science: Investigations. Columbus, Ohio: SMEAC Information Reference Center, The Ohio State University, 1977.

Tabbron, G. and J. F. Kerr. "Teacher Opinion on the Objectives of Primary School Science." Educational Research, 1(14), 1971.

LEARNING

Beeson, G. W. "Hierarchical Learning in Electrical Science." Journal of Research in Science Teaching, 14(2): 117-127, 1977.
Descriptors--*Educational Research; *Electric Circuits; Evaluation; *Instruction; Physics; *Science Education; *Secondary Education; Secondary School Science; *Task Analysis

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Richard J. Bady, Mount Senario College.

Purpose

The purpose of the study was to apply Gagné's concept of learning hierarchies to a particular area of learning not previously researched: electrical circuits. This is useful because most previous research has been limited to learning of mathematics. Further, Gagné's distinction between intellectual skills and verbal information was to be investigated. Gagné has claimed that only intellectual skills and not verbal information form the elements of learning hierarchies.

Research Design and Procedure

A possible learning hierarchy was constructed by the author. The terminal capability was to find the potential difference, resistance, and current for a circuit with two resistance components. The hierarchy consisted of 21 intellectual skills and 7 information units.

Modifications were made in the hierarchy as a result of: (1) criticism by teachers and subject matter experts, (2) Trial tests of the test questions to determine they actually represented different capabilities, and (3) Trial tests of the learning materials used in the final experiment.

This learning program was constructed using the hierarchy with test questions for each element embedded in it. Two questions were used for each intellectual skill, but since verbal information units could only be tested by the selection of the correct term, only one question was used for these items.

Subjects in the preliminary and final testing were 166 tenth-grade students from five different schools.

The relationships between the various possible elements were investigated by constructing 2x2 contingency tables of success/failure for each pair of items, and applying a statistical test of dependence. This was done for two of the three types of item pairs: those consisting of an intellectual skill and a verbal information unit. No tests were done on pairs of verbal information units.

Findings

Of the 34 contingency tables for pairs of intellectual skills tested, 27 were found to embody a prerequisite relationship as predicted by the hierarchy. Most of the other seven could not be adequately tested in this sample. The author concludes that when intellectual skills are the desired outcome of learning, the hierarchy method should be employed.

The relation between intellectual skills and verbal information units is more complex. It was found that a prerequisite relationship did not always exist for these pairs, but that "...whenever elements of verbal information were included in the final hierarchy, they were subordinate to intellectual skills."

Interpretations

The author concludes that (1) "The study provides support for the distinction between intellectual skills and verbal information units," (2) "Where material to be learned contains intellectual skills...those skills should be ordered into a hierarchical sequence for most efficient learning," (3) "Other outcomes to be learned (for example, verbal information, motor skills, attitudes) can be taught in any order which is considered appropriate...." and (4) "...where students come to a new course deficient in certain skills needed to begin the course, learning hierarchies can be used to identify these missing skills and guide the teaching of them."

ABSTRACTOR'S ANALYSIS

The report is useful in that it presents a paradigm for constructing learning hierarchies in a nonmathematical field. It shows that a valid set of prerequisite relationships can be constructed and tested. Further, the study does show that there is some distinction to be made between intellectual skills and verbal information units. However, the author chooses not to explain why they form different sorts of prerequisite relationships. Further research in this area could be useful in understanding the nature of these different kinds of learning.

Conclusion (4) seems quite reasonable, although it is not clear that it follows from the data in this study.

Conclusions (2) and (3) however must be criticized, for while perhaps they are true, they in no way can be concluded from this study. Conclusion (2), that the use of valid learning hierarchies leads to efficient learning, is not tested in this study for the efficiency of the learning was neither tested nor compared to some other method. Conclusion (3) not only does not follow from the data, it seems unreasonable. The fact that verbal information units did not form a nice hierarchy does not imply that they could be taught in any order. The whole issue of the connection between a prerequisite relationship that tends to exist and the implication of that relationship for teaching are not addressed in this study. The author, and indeed many others, assume that because they find no students knowing B who do not also know A, that it is necessary to teach A before teaching B. A learning hierarchy is descriptive but we lack evidence that it is prescriptive as well. It describes the static structure of knowledge, but taking the next step requires studies in which concepts are taught in several possible orders. Howe (Journal of Research in Science Teaching, 11(2): 105-112, 1974) for example, showed that a hierarchy that seems logically necessary may not be most appropriate for teaching the terminal skill. The implications of static learning hierarchies as now used need further investigation.

Murray, Darrel L. "A Visual Recall Probe of Cognitive Structure." Science Education, 62(1): 39-46, 1978.
Descriptors--*Academic Achievement; Biology; *College Science; *Educational Research; Higher Education; *Perception; Science Education; Visual Discrimination; *Visual Perception

Expanded abstract and analysis prepared especially for I.S.E. by David L. Dunlop, University of Pittsburgh at Johnstown.

Purpose

The purpose of Murray's study was to investigate the visual perceptual dimension of cognitive structure as it relates to achievement in college biology. This included the following: (a) a description of a visual recall probe developed from Haber's work, (b) an examination of the relationship between performance on a visual recall test and performance of three achievement tests, (c) an evaluation of a hypothesis which states that the predictive validity of the visual recall probe will decrease with subsequent measures of achievement on less related learning materials, and (d) a consideration of the implications of using visual recall probes for the early detection of learning differences.

Rationale

The rationale for this study is related to the findings of several researchers (Novak, Ring, and Tamir, 1971; Ausubel, 1968; Ring and Novak, 1971); however, it appears that the focus of this research is based upon Novak's (1971) suggestion that the best predictors of achievement are specific measures of cognitive structure (as defined by Ausubel, 1969).

Murray then reasoned that if visual recall performance shows a significant relationship to achievement on closely related learning materials, then one would hypothesize that the predictive validity of the same visual recall probe will decrease when applied as a predictor of achievement on less related learning materials.

Research Design and Procedures

Two hundred twenty-five introductory biology students at the University of Illinois at Chicago Circle were the subjects of this study. This course, The Biology of Populations, utilized a wide variety of learning opportunities (slide-based lectures, laboratory experiences, text materials, discussion sessions, computer-assisted instruction, etc.) to convey the major facts, concepts, and principles important to the course content.

Haber's (1970) visual recall task was modified and used as a method for probing cognitive structure. During a particular lecture, approximately 80 color slides were presented to the students, and near the end of the lecture the students were told that a visual recall test was to be given. The recall test consisted of 30 randomly arranged slides, 15 of which were part of the previous lecture and an additional 15 were similar slides never seen by the students. As the slides were individually projected, students were asked to indicate on a true-false IBM answer sheet if they recalled having seen the slide during the lecture. The students were informed that their course grades would not be influenced by the results of the recall test. The Kuder-Richardson formula 20 technique was used to analyze student responses, and the results showed a reliability coefficient of 0.70, which when projected to a test of 100 items is predicted to yield a reliability coefficient of 0.89.

Student achievement was evaluated through the use of 30-item multiple choice tests. Test construction followed a carefully developed procedure used for the past decade of testing in the Biology of Populations course, and the tests were designed to evaluate achievement in a delimited area of subject matter presented in lecture, laboratory, and related readings. Altogether, three different achievement tests were administered to the subjects as follows: test I, approximately 9 days after the visual recall test; test II, approximately 30 days after the visual recall test; test III, approximately 45 days after the visual recall test. Reliability (Kuder-Richardson technique) of each test was as follows: 0.83 and 0.81 on forms A and B of achievement test I; 0.84

and 0.85 on forms A and B of achievement test II; 0.76 on the single form of test III.

The general treatment sequence included (1) observation of a slide-based lecture, (2) completion of a visual recall test, and (3) completion of three multiple choice achievement tests.

Analysis of variance (equal cell size) was used to test the relationship of visual recall performance by the students to their achievement. Quintile groupings based on visual recall test scores were formed for the purposes of analysis. An independent analysis was performed for each of the three achievement tests with respective scores serving as the dependent variable. Where F ratios were significant ($p < 0.05$), correlation coefficients were used to detect differences in the predictive validity of the visual recall test with respect to the three achievement tests. All data were computed using the Plato Basic Statistics Package.

Findings

An examination of the relationship of performance on the visual recall test by the subjects to their performance on achievement test I indicates that the amount of visual information gained in lecture is positively related to the amount of new verbal material learned during a following period of instruction. The relationship between visual recall performance and scores on achievement tests II and III was weaker than the relationship between visual recall scores and achievement test I.

Interpretations

Although visual and linguistic information may be considered as different aspects of cognitive structure, these results suggest that a visual recall probe can serve as a predictor of achievement; however, the predictive validity of the visual recall probe decreases as learning shifts to less related content structure areas. The predictive validity of the visual recall probe is highest when applied as a predictor of achievement on closely related learning materials.

Since learning materials presented during the respective achievement testing intervals showed content differences, it appears that the visual perceptual data aspect of cognitive structure is subject to the constraints of content specificity.

ABSTRACTOR'S ANALYSIS

Haber (1970), from his work with visual mental capacity, has indicated that among the most significant of his findings is the suggestion that there is one kind of memory for pictorial material and another for linguistic. This finding is the basis for Murray's interest in the relationship between visual recall performance (pictorial) and performance on related achievement tests (linguistic).

Murray's model of a visual recall probe for the early detection of learning differences related to achievement is easy to administer, and its unobtrusive nature will be an asset to researchers interested in securing the subject's cooperation. Further, this specific technique appears to be unique in the study of cognitive structure and visual recall. As Lutz (1977) indicates, the effect of external imagery on learning has been investigated primarily within the paired associate learning paradigm.

For example, in an article relating to paired associate recall Emmerich and Ackerman (1976) state that a number of recent studies have focused on the effects of the elaboration of stimulus materials on memory for those materials. Generally, it has been found that elaborations, visual or verbal, aid retention of the target item pairs relative to control conditions in which the word or picture pairs are presented without any associating mechanism. As Murray stated, additional work on techniques of attaching words to visual images will be required before the potential of using visual images to facilitate verbal learning can be more fully realized. This also appears to be true for the effects of elaboration upon the results of this and similar studies.

Hoffman (1971), utilizing a recognition-memory process similar to Haber's, obtained only limited developmental differences in performance for 3-, 5-, 7-, and 9-year old and adult subjects following a single presentation of 100 picture stimuli. This could suggest that Murray's visual probe would be equally appropriate for several age groups.

However, another study by Hoffman (1976) deserves mention at this point as it relates to the number of slides to be utilized during the visual probe. In this study Hoffman found clear developmental differences in performance, with the results extending the finding of poorer performance with an increased number of recognition alternatives for adults to three- and seven-year olds. Additional research in this area would also be of value, especially if the number of slides in a given probe approaches 100.

Murray stated that, "these results suggest that a visual recall probe can serve as a predictor of achievement." Although future research may support this claim, it is important to emphasize the author's use of the word, "suggest." One area of caution is the relatively low (but statistically significant) correlation coefficients which exist between visual recall scores and scores on achievement tests I, II, and III (0.385, 0.276, and 0.302, respectively).

Murray also stated that, "Based upon the present study, it appears that visual recall probes exhibit the highest predictive validity when applied during the period of instruction in which achievement is to be measured." This indicates that a "time variable" is involved. If this is true, it is difficult to understand how Murray can also state, "...the learning materials presented during the respective achievement testing intervals showed content differences. Thus, it appears that the visual perceptual data aspect of cognitive structure is subject to the constraints of content specificity." In the first instance the relationship between visual recall and achievement is dependent upon the time interval between the probe and the achievement test while in the second instance this relationship is dependent upon the content. Additional research controlling for these multiple variables should be conducted to clarify this point.

Future researchers should also examine the content specificity of each achievement test and relate this finding to the content of the visual probe (slide-lecture). Careful analysis of this situation is needed to advance and refine our understanding of Murray's hypothesis which states that the predictive validity of the visual recall probe will decrease with subsequent measures of achievement on less related learning materials.

Another intriguing variable related to this study is that of student "attentiveness." Does this variable help to explain why visual recall probes exhibit the highest predictive validity when applied during the period of instruction in which achievement is to be measured? Additional research in this area would be interesting, but possibly difficult.

Although this study has raised many questions, it has also demonstrated a new and potentially useful methodology for examining the relationship between cognitive structure and visual recall. The conceptual contribution is also significant and warrants further study.

REFERENCES

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

Ausubel, D. P, and F. G. Robinson. School Learning. New York: Holt, Rinehart, and Winston, 1969.

Emmerich, H. J. and B. P. Ackerman. "The Effect of Pictorial Detail and Elaboration on Children's Retention." Journal of Experimental Child Psychology, 21: 241-248, 1976.

Haber, R. "How We Remember That We See." Scientific American, 222(5): 104-115, 1970.

Hoffman, C. D. "Recognition Memory for Picture: A Developmental Study." Paper presented at the meeting of the Eastern Psychological Association, New York, April 1971.

Hoffman, C. D. and S. Dick. "A Developmental investigation of Recognition Memory." Child Development, 47: 794-799, 1976.

Lutz, K. and R. Lutz. "Effects of Interactive Imagery on Learning: Application to Advertising." Journal of Applied Psychology, 62(4): 493-98, 1977.

Novak, J. D.; D. G. Ring; and P. Tamir. "Interpretation of Research Findings in Terms of Ausubel's Theory and Implications for Science Education." Science Education, 55(4): 483-526, 1971.

Ring, D. G. and J. D. Novak. "The effects of Cognitive Structure Variables on Achievement in College Chemistry." Journal of Research in Science Teaching, 8(4): 325-333, 1971.

Shyers, J. and Cox, D. "Training for the Acquisition and Transfer of the Concept of Proportionality in Remedial College Students." Journal of Research in Science Teaching, 15(1): 25-36, 1978.

Descriptors--*Abstract Reasoning; *College Students; Educational Research; Higher Education; *Learning Theories; *Learning Processes; *Remedial Programs; Science Education; Transfer of Training

Expanded Abstract and Analysis Prepared Especially for I.S.E. by J. Dudley Herron, Purdue University.

Purpose

In this study of acquisition and transfer of the concept of proportionality, the authors attempt to answer several questions. First, will training on the operations of the INRC group which Piaget sees as inherent in proportions lead to improved ability in solving the proportions? Second, will training transfer from one task to another? Third, will the training transfer to some new task requiring similar logic? Fourth, if transfer does occur, is it due to similarities between the structure of the training and transfer tasks (e.g., both tasks involving direct proportion or both tasks involving inverse proportion) or due to the presence of identical elements in the two settings (e.g., both tasks involving the manipulation of weights and lengths of a lever arm)? Finally, the authors tested the "Deep-end Hypothesis"--an hypothesis that more learning will occur when a complex task precedes a simpler task than when the order is reversed.

Rationale

The research is clearly based on Piaget's notions concerning the logical structure of formal operations in general and proportional reasoning in particular.

The question concerning the "Deep-end Hypothesis" is an attempt to check the generalizability of a result obtained by Dienes and Jeeves in 1965.

Research Design and Procedure

The study was performed in a remedial mathematics class at the college level. The group was very diverse. Ages ranged from 18 to 55 with a median age of 22. Mathematics scores on the California Achievement Test ranged from the sixth- to the thirteenth-grade level with a median of 9.1.

The study was conducted during the third week of an eight-week summer session. Delayed posttests were given during the seventh week.

Students were assigned alphabetically to two groups of 24 students each. Both groups were pretested on proportions and then trained on three tasks: Rings and Shadows (R&S), Half-Balance (H-B), and Wheelbarrow (WB). However, the order of presentation of these tasks differed between the two groups. Both groups were given a posttest on proportions immediately after training and again four weeks later. A written test of Piaget's Balance Beam Task was given as a transfer task. This test was also given to 21 more able students enrolled in a special program involving computers. This more able group was treated as a control group. The total design is summarized in Table I taken from the original paper.

TABLE I

Design of the Study

Group 1[a]	Group 2[b]	Control
Pretest proportions	Pretest proportions	
Train R&S and H-B	Train WB	
Test R&S and H-B	Test WB	
Train WB	Train R&S and H-B	
Test WB	Test R&S and H-B	
Posttest proportions	Posttest proportions	
Posttest balance	Posttest balance	Posttest balance
Delayed posttest	Delayed posttest	

[a]Initial training on two four-element group structures followed by training on the eight-element group structure.

[b]Initial training on the eight-element group structure followed by training on the two four-element group substructures.

Understanding this study requires understanding of the tasks used in the study. For a detailed description, one should study the original paper. However, focus on a few characteristics of the tasks should clarify the intent of the study.

Let us take the simplest part first--the complexity of the tasks. The Half-Balance Task is designed so that it illustrates only an inverse proportion. The Rings and Shadows Task is designed so that it illustrates only a direct proportion. By contrast, the Wheelbarrow Task is designed to illustrate either a direct or an indirect proportion. Thus, there are simple tasks (H-B and R&S) and a complex task (WB). Consequently, by varying the order in which the tasks are presented, the authors are able to test the "Deep-end Hypothesis."

The Balance Beam Task that was used to check transfer of training may be presented so that it illustrates either inverse or direct proportion. In addition, it involves the manipulation of weights and distances as does the Half-Balance Task used for training. Then if one compares the performance on the total Balance Beam Task with the performance on the Wheelbarrow Task, the comparison is between two tasks that have similar logical structures (both direct and inverse proportions are involved) but have dissimilar elements (the physical arrangements in the two tasks are sufficiently different that students are unlikely to sense that they are "doing the same things"). In similar fashion, if performance on that portion of the Balance Beam Task that involves a direct proportion is compared to performance on the Rings and Shadows Task, the comparison is between two tasks that have similar structures but dissimilar elements. By contrast, comparison of performance on that portion of the Balance Beam Task that involves indirect proportion with performance on the Half-Balance Task involves comparison of tasks with similar structures and similar elements.

Findings

All tests of hypotheses were done by comparing the errors made by students when they performed the various tasks. There were more possibilities for errors in some tasks than in others, necessitating the use of proportions rather than raw scores. Furthermore, the range of scores was very small, making it necessary to normalize the data by the arc sine transformation. It is probably wise to keep these alterations in mind as the results are interpreted.

First, let us deal with the "Deep-end Hypothesis." Essentially, the results provide no support for the hypothesis that more learning will occur when a complex task precedes a simpler task. An analysis of covariance of the data suggested that the reverse may be true, but the evidence isn't strong one way or the other.

The results indicate that the training (regardless of order of task presentation) had a beneficial effect. The scores on the posttest and delayed posttest were higher than the scores on the pretest.

Having determined that there was a training effect, the next question addressed is whether this learning transfers to another task that involves proportions. This was tested by comparing the performance of the two trained groups and the control group on the Balance Beam Task. The mean number of errors made by the three groups was 4.92 for Group 1, 5.52 for Group 2, and 6.22 for the control. Statistical tests revealed that the difference between the mean for Group 1 (4.92) and the mean for the control group (6.22) was significant at the 0.05 level. Thus, it appears that the training which was provided did have an effect on performance of a new task that involves proportions.

The final question addressed is whether this transfer of training is due to identical structures or to identical elements found in the training tasks and the Balance Beam Task. Recall that by identical structure the authors mean that the same kind of proportional relationship is involved. For example, the WB Task used in training and the Balance Beam Task used as the transfer task involved both direct proportions and inverse proportions so they were said to have identical structure. A chi square test was done to see whether success on one of the tasks was related to success on the other. The test indicated no significant relationship. Thus, the transfer of training that was found does not appear to be explained by similarity in structure between the training and transfer tasks. (Other comparisons were made and all such comparisons support the conclusion that transfer was not due to identical structure.)

The Half-Balance Task that was used in training has, according to the authors, elements that are identical to elements found in the Balance Beam Task used to test transfer. In both cases, weights are added to get bending or rotation. In both cases the distance between the weight and a fulcrum is considered. It is this kind of similarity that the authors seem to have in mind when they refer to identical elements. A chi square test of the hypothesis that success on the H-B Task is independent of success on the Balance Beam Task led to rejection of the hypothesis. There is a relationship between performance on the two tasks and the authors conclude that the transfer of training that they observed is due to identical elements found in the training and transfer tasks.

Interpretations

The authors conclude that training that takes into consideration the logical structure of proportions and that builds on the existing cognitive structure of the learner can result in learning of proportions. They also conclude that the transfer of this training is dependent on the presence of identical elements in the training and transfer tasks rather than on identical structures in the training and transfer tasks.

ABSTRACTOR'S ANALYSIS

I found this study to be both interesting and difficult to evaluate.

The conception of the research seems to be much better than average. The authors used considerable care to design training tasks with the same logical structure found in the Balance Beam Task used by Inhelder and Piaget to investigate proportional reasoning, and they were careful to vary the complexity of the tasks, the nature of the tasks (i.e., the elements that are manipulated to produce the proportional relationship) and the logical structure of the tasks (i.e., whether a direct or inverse proportion is involved). The authors also address important questions: Is there any learning? Does the learning transfer to new situations? If transfer occurs, what accounts for the transfer?

Overall, the design of the study appears to be sound. However, several questionable assumptions were made. Subjects were assigned to treatment groups alphabetically, assuming that no extraneous variables were correlated with alphabetical order. Although I cannot cite the study, I believe previous research has shown that alphabetical assignment often leads to nonequivalent groups. As an example of factors that may operate, Anglo-Saxon and Spanish surnames tend to begin with different letters. Such influences could lead to groups that differ in English language skills when assignment is made alphabetically.

Another questionable assumption is that estimates of test reliability derived from data for a group of seventh-grade students constitute suitable estimates of the reliability of these tests when administered to a very diverse group of college students. In similar fashion, the use of a superior group of students in a special program as a "control" for the treatment groups appears questionable.

I also worry about some assumptions made in the data analysis. I understand why it was necessary to compare proportions rather than raw scores and why it was necessary to normalize the data by the arc sine transformation, but I am uneasy about it. One chance error on a task with three possible errors affects the proportion of errors much more than the same chance error on a task with ten possible errors. With very large numbers of data points in each cell in the analysis, such factors should apply evenly, but I infer from the report of this research that some cells simply didn't contain a very large number of data points. Perhaps I worry needlessly. I still worry.

I have already mentioned that this study was complex. This complexity undoubtedly accounts for some of the difficulty that I had in sorting out what the authors did and what they found. However, I believe that attention to a few details in the writing could have saved readers a great deal of time and effort. I mention a few.

I was never quite clear about the training. Since this is a study of the effectiveness of a training procedure, it seems very important to me that the reader have a clear understanding of just what that training was. A sample protocol, an outline of one lesson, or sample questions asked during instruction would tell a great deal about what was done. Instead, we are simply told that training was done by means of a classroom demonstration of each task "with stress on the reversibilities needed to establish a state of equilibrium after an identity transformation was performed." What do the authors mean by "stress on reversibilities"? Did they simply point out that balance can be restored after adding a weight to one side of a balance beam by removing that weight, by adding weight to the other side or by altering the length of the lever arm, or did they have students do calculations to demonstrate equality of moments?

Other important points are left unclear. Just what is meant by "identical structures," "identical elements" and "identical concepts"? If I am to benefit from this research as it pertains to transfer, I must know what these terms mean. I am not told and am never sure that my inferences about their meaning are correct.

The exact nature of the tests used in this study isn't clear. There is reference to four types of proportions on these tests: Direct-Abstract, Direct-Numerical, Inverse-Abstract, and Inverse-Numerical. I think that I know how the ~~direct~~ and inverse proportions questions would differ, but I'm not sure that I know what the numerical and abstract questions would look like. I want to know what is being tested, and the best way for authors to tell me what is being tested is to provide sample items.

I fault those who reviewed this article for publication as much as the authors for the lack of clarity that I have cited. Authors are always close to their work and statements that appear perfectly clear to them are not clear to others. A conscientious reviewer can point to areas where clarification is needed. Unfortunately, reviewers and editors sometimes stress brevity to the point that clarity is sacrificed. We sometimes act as though our purpose is to publish as many articles as possible; it should be to communicate as much as possible. Fewer articles, very carefully presented, may benefit us more than many articles that leave the reader wondering what was learned from the research.

Now let me turn to what I have learned from this research.

The "Deep-end Hypothesis" seems so unimportant to me, and the results concerning it seem so tenuous that I ignore it altogether. Not so with the other considerations.

In spite of the questionable assumptions that are mentioned above, I am inclined to believe the results of this research. They appear reasonable to me. They are consistent with other things that I know. For example, a great deal of research suggests that students who are taught something are more likely to know it than people who aren't taught that thing. Thus, the finding that students who were given the training did better on the posttest than on the pretest is important for the purposes of this research, but it is not new. (If the authors had provided a description of the training and testing materials, this finding would be important to others who are trying to teach proportions. As reported, nobody but the authors can benefit from the finding.)

The finding that the training procedure leads to better performance on the part of those who were not taught is consistent with other research, but not profound.

The result that is potentially important is the one pertaining to the factors that influence transfer. Let us look more carefully at just what was shown.

On the Half-Balance Task, training apparently focused on the fact that the bending of the metal bar when additional weight is hung on the bar can be nullified (reversed) by either removing the additional weight or shortening the distance between the weight and the support for the bar. Presumably the quantitative relationship that the product of the weight and the length of the lever arm is constant for a particular bend of the rod was also taught.

It was learned that success on a test over this training and a test over a portion of the Balance Beam Task was highly correlated. The portion of the Balance Beam Task in question requires that conditions on one side of the beam is increased, the beam will bend, and that balance can be restored by either removing the added weight or shortening the distance between the weight and the support for the beam. Presumably the quantitative relationship that the product of the weight and the length of the lever arm is constant for balance to be maintained was also covered.

The similarity between these two situations is rather clear to this reader and it was apparently clear to the subjects in this study. Those who did well on one task did well on the other task as well.

Now let us examine a situation in which success on one task does not predict success on the other task.

If the weights on either side of the Balance Beam are fixed, the equilibrium of the balance can be upset by moving one of the weights. To restore the equilibrium, that weight can be moved back to its original position, or the weight on the other side can be moved. It can be shown that the movements on either side must be directly proportional. For example, if the beam is balanced with a weight on the left 2 cm from the support and a weight on the right 5 cm from the support, moving the left weight to a position 4 cm from the support can be compensated by moving the right weight to a position 10 cm from the support. Doubling the length of the lever arm on one side must be compensated by doubling the length of the lever arm on the other side.

One of the training tasks, R&S, also had variables that were related by a direct proportion. Consider a ring 4 cm in diameter that is located 10 cm from a candle. If this ring is replaced by one with a 2 cm diameter, the shadow cast by the ring will be smaller. However, the original shadow size can be restored by moving the 2 cm ring until it is located only 5 cm from the candle. Halving the diameter of the ring is compensated by halving the distance from the candle to the ring.

The authors found that performance on the R&S task was independent of performance on the direct proportion task with the Balance Beam, even though both tasks involved a direct proportion. I don't find this result too surprising but I do consider it important.

I believe that most of us would describe the Half-Balance Task as "very similar" to the inverse proportion Balance Beam Task. I believe that we would describe the Rings and Shadows Task as "rather different" from the direct proportion part of the Balance Beam Task. We are probably reacting to what the authors of this research describe as "identical elements" in the balance tasks and "nonidentical elements" in the R&S and Balance Beam Tasks. They don't seem the same, even though the mathematical relationship inherent in the two tasks is the same. The authors conclude that transfer is influenced by identical elements and not identical structures. Under the conditions of this study, I think that their conclusion is correct. I also think that this result deserves further research along the following lines.

I ask you to think about your own notion of direct proportion. You probably see direct proportions everywhere. In density you see that mass is directly proportional to volume. In kinetics you see that acceleration is directly proportional to the accelerating force. In Hook's law you see that the stretch of a spring is directly proportional to the force applied to the spring. In chemistry, you see that the mass of a product is directly proportional to the mass of the limiting reagent in the reaction. And on and on.

It is likely that this relationship that we call a direct proportion did not become apparent until we had encountered quite a number of specific relationships, all of which were related in this sam way. After a while we began to look for such a relationship because we had found that it was a useful way of making sense out of other data that we wanted to interpret.

My point is that these "structural" relationships are not immediately transparent when we view data. Indeed, it is probably necessary to force ourselves to ignore more obvious information (such as the fact that a particular relationship involves manipulation of weights and lengths whereas another involves manipulation of diameters of circles and rather different lengths), getting beyond this surface information to the logical elements underneath.

This is similar to what I understand Piaget to be saying when he discusses the inability of young children to conserve substance (Piaget, 1964). The young child is likely to focus on the change in length or the change in height when a ball of clay is rolled out. Perhaps the deformation itself commands attention so much that the child is unable to attend to the logical structure in the setting. It seems as though it is not until such events become routine that the mind is able to disengage from the surface phenomena and deal with the logical underpinnings.

Think of your first visit to a foreign country or a strange city in your own country. Although you have the training to enable you to observe differences in economic structure and governmental organization, you are unlikely to make such analyses in the beginning. You are too busy with the new smells, the unusual colors, the strange dress, and the fascinating architecture. Not until our mind has dealt with such surface observations and become saturated with them do we naturally move to such "second order" considerations as why the architecture is different or why our government and economic system has evolved as it has.

The kind of question that I am trying to raise and the one that would appear to be a logical extension of the research reported in this study is this. Is it reasonable to expect students to be able to transfer what they have learned about direct proportions and inverse proportions during a short period of training to a somewhat different task that they encounter several weeks later? Is it perhaps more reasonable to expect the ability to apply such logical operations spontaneously to develop over an extended period of time as the result of encountering this same logical structure in a variety of situations which, on the surface, appear to be totally unrelated? Is it not the discovery that such logical operations enable us to find order in events that appear to lack order what we are talking about when we speak of intellectual development? If so, I should think that the "transfer due to identical structures" discussed in this research would be unlikely to appear in a short-term study such as this one, but might appear in a similar study that extends over a period of a year or longer.

Ukens, Leon, L. and Philip R. Merrifield. "The Structure-Of-Intellect Model Applied to a COPES Learning Sequence." Journal of Research in Science Teaching, 13(3): 221-225, 1976.

Descriptors--*Curriculum; Educational Research; Elementary Education; *Elementary School Science; *Intelligence Factors; *Predictive Measurement; Predictor Variables; Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Willis J. Horak, University of Arizona.

Purpose

This study was designed to investigate relationships among student's mental abilities and student's achievement on a specific unit from a nationally developed elementary science curriculum project. In order to effectively fulfill this purpose, the researchers had to identify relevant mental abilities and develop criterion measures for concept achievement. The findings were then interpreted for their relevance to future curriculum-development projects and to classroom instruction.

Rationale

The structure of the intellect model developed and described by Guilford in the book The Nature of Human Intelligence (1967) formed the basis for the mental abilities assessed in this research study. This model views intelligence in terms of 120 specific intellectual abilities. The presence of abilities which interact with specific educational tasks were conjectured to influence student achievement. Thus the researchers chose to study 17 specific intellectual abilities. No ability relating to the structure of intellect mental operation of divergent production was chosen. Similarly, no ability relating to the structure of intellect content area of behavioral was chosen. All other categories of Guilford's model were deemed appropriate for mental abilities utilized in the COPES sixth-grade mechanical energy teaching-learning sequence.

Research Design and Procedures

The research design utilized in this study was a one-group pretest-posttest design (Campbell and Stanley, 1963). For the study six teachers taught the COPES Mechanical Energy Sequence to a total of 158 sixth-grade children. All student scores were originally combined for data analyses. However, subsequent analyses were also carried out on boys' scores and girls' scores separately. The pretest and the posttest were constructed by the researchers special for this study. For these tests, curricular validity was determined by judgment of a panel of science educators. Reliability estimates were computed as 0.63 for the pretest and as 0.77 for the posttest.

The following structure-of-intellect variables were chosen for study: (1) cognition of semantic units, (2) cognition of semantic classes, (3) cognition of semantic implications, (4) cognition of figural classes, (5) convergent production of semantic relations, (6) convergent production of semantic systems, (7) convergent production of symbolic systems, (8) convergent production of symbolic relations, (9) convergent production of symbolic implications, (10) convergent production of figural transformations, (11) divergent production of semantic units, (12) divergent production of semantic transformations, (13) divergent production of semantic relations, (14) divergent production of figural systems, and (15) evaluation of figural classes. All of these abilities have published tests available with reported reliabilities ranging from 0.41 to 0.92.

Findings

The data for this study were analyzed utilizing stepwise multiple regression procedures. The pretest score was the dependent variable. All the other variables were entered into the regression equation in the order of decreasing prediction importance. The analysis revealed that four mental abilities significantly (0.05 level) predicted student achievement on the posttest. This was true for the combined sample, the sample of males, and the sample of females. However, the same four abilities were not evident in each sample. The only one which appeared in all the sample analyses was convergent production of semantic relations.

Interpretations

The statistical analysis of the data revealed that convergent productive thinking is the most important type overall for the COPES Mechanical Energy sequence. Three of the four significant predictors were concerned with convergent production of various kinds of content. The COPES sequence is structured so that all activities converge on one of the main conceptual themes. Thus the curriculum developers succeeded in their efforts to produce a science curriculum that required thinking in line with the overall project objectives. However, the difference in the outcomes for the male and female sample indicates that there are probably specific differences in abilities between the sexes. These findings indicate that it is possible to determine prerequisite mental abilities for success in the COPES curriculum. Guilford's structure of intellect model is also useful when deciding which type of mental abilities are of importance. Classroom teachers as well as future curriculum developers may profit from this type of information.

ABSTRACTOR'S ANALYSIS

This study is one of the type that is necessary in order to more fully analyze the modern elementary science curriculum projects. Teachers today are faced with choosing from diverse projects such as ESS (Elementary Science Study), SCIS (Science Curriculum Improvement Study), S:APA (Science: A Process Approach), COPES (Conceptually Oriented Program in Elementary Science), and other textbook-based programs. All of these programs stress varying amounts of student involvement and student activity. Often one is unsure if the programs actually do develop abilities that the curriculum developers stressed. Thus studies analyzing the mental abilities prerequisite to achievement in the various projects are quite necessary.

The report is presented in a form that is readily understandable. For those not thoroughly familiar with Guilford's structure of intellect model or COPES overall concepts and objectives, there are introductory explanations. However, the brevity of the total report somewhat limits many aspects of this presentation.

The educational implications of the study are limited somewhat by the fact that no attempt was made to examine the overall teaching behavior associated with the instruction. It was stated that six teachers were used to teach the instructional sequence and that the six teachers were provided inservice training by the researcher when needed. It would be useful for others to know how familiar each of these teachers was with the COPES materials. Also it would be useful to know how long each had been using COPES materials in their classrooms. Additionally, it would be of interest to other researchers whether a specific teacher's style determined, in effect, some of the necessities for specific mental abilities on the part of their students. This could have been ascertained by doing the data analysis separately for each teacher. While the resulting sample sizes would have limited the interpretations of these analyses, evidence would have been obtained for considering or not considering all teachers as one group as was done in the reported analysis.

The report also needed to include the total amount of instructional time spent on the COPES Mechanical Energy sequence. While the COPES teacher's guides specify some general time limits, not all teachers follow these suggestions. Did all students in fact receive the same amount of instructional time on the tasks during the sequence?

The results of the study are interesting in many aspects. Some science educators consider the objective of elementary science to develop divergent thinking to be most valuable. In fact, many see a relationship between this divergent thinking ability and educational success. They stress that teachers should be encouraged to most of the time stress divergent thinking in their classrooms. Doing otherwise is viewed as educationally unsound, leading to student conformity and also to uninspiring teaching. The fact that the convergent production operations were most important should therefore raise many questions. Similarly, the fact that divergent production of semantic transformations was the most important predictor for males, while it was not even a significant predictor for females should raise further educational questions. Some of these questions may have been answered by providing a table of means and standard deviations for the total group, the males, and the females on all of the measures. Did the males and females differ in their mental abilities on the structure of intellect variables? One table would have helped others answer this question. If the two groups were indeed similar, then the cognitive functioning during the instruction was most important. If they were not similar, then one group may have had to utilize other mental abilities to be able to successfully achieve with the COPES program.

Lastly, the rationale for the use of Guilford's model on intellect needs to be more fully explained. Was this study designed to further explore the applicability of this model? Or was it designed to study student abilities prerequisite to success in the COPES curriculum? There are other student aptitudes that may be just as useful when answering the second question. Student variables as cognitive styles, cognitive preferences, or locus of control may be more useful. Some of these variables are also easier to measure. The testing time of four hours appears to limit the usefulness in many school settings. Further studies could thus compare other constructs with the structure of intellect variables measured to see which are more important to significantly predict success in science instruction using the COPES curriculum.

The educational task of assessing all the significant variables however may still offer diminishing returns for the classroom teacher. Science is only one area of the total elementary school curriculum. Other subjects may need other abilities. Assessment of pertinent structure of intellect variables and the construction of mental maps for each child could be a most time-consuming task.

Studies of this type are important, however, when trying to explain the relative effectiveness of children in our modern science programs and in our elementary schools.

REFERENCES

Campbell, D. T. and J. C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company, 1966.

Guilford, J. P. The Nature of Human Intelligence. New York: McGraw-Hill Book Company, 1967.

TEST CONSTRUCTION

Hofstein, A.; R. Ben-Zvi; and D. Samuel. "A Factor Analytic Investigation of Meyer's Test of Interest." Journal of Research in Science Teaching, 14(1): 63-68 (1977).

Descriptors--*Educational Research; Factor Analysis; *Measurement Instruments; Science Education; *Scientific Attitudes; Secondary Education; *Secondary School Students; *Student Interests

Expanded abstract and analysis prepared especially for I.S.E. by Rodney L. Doran, State University of New York at Buffalo.

Purpose

The purpose of this study was to investigate the empirical validity of Meyer's Test of Interests (MTI). No hypotheses were stated, although the intent was to explore the quantitative relationships among the major construct variables of the MTI.

Rationale

Meyer (1970) developed and established face validity for the MTI. The MTI has subsequently been used with several English and Israeli populations. The main outcome of the Meyer test is a "science-orientation" score, which is assumed to provide an overall measure of a student's leaning toward science. Science orientation itself is seen as the aggregate of three major construct variables:

1. Interest in science (i) as a leisure activity; (ii) as a problem-solving method; and (iii) as a source of information in relation to physics, biology, chemistry, geology and the history of science.
2. Scientific thinking and attitude toward science in general.
3. Attitude toward school science teaching.

Research Design and Procedure

The research procedure followed was primarily a factor analytic examination of MTI data collected previously in Israel (Ben-Zvi, et al., changed to Hofstein, et al., 1977) and in England (Meyer, 1970). The variables analyzed were sub-scale scores from the MTI for both samples. The Israeli sample was 380 tenth-grade students and the English group was 680 G. C. E. O-level students. The reader was referred to the original studies for detailed descriptions of the samples.

From the Meyer article, it appears that the English sample was composed of students in the "experimental" Nuffield curriculum, "at the end of the fifth form of secondary school (approximately grade 11 or 12 in the United States)." These students were selected from "all schools able to cooperate from a list provided by the Nuffield Science Teaching Project." The sex distribution was not equal with 470 boys and 210 girls. The Israeli sample, as described by Hofstein, et al., "involved 410 tenth-grade students from seven Israeli high schools in different parts of the country." The selection of schools was undertaken in such a way as to ensure an adequate distribution of subjects according to their socio-economic origin. All students had, during their previous educational career, followed a uniform curriculum comprising the conventional curricular disciplines, including science. In addition to the MTI, the Israeli students completed two standard intelligence tests and two academic achievement tests developed for the original Israeli study. No reliability estimates for the MTI when used with the English sample were calculated, but the authors listed the reliabilities for the three MDI subscores as 0.71, 0.76, and 0.73, respectively, with the Israeli sample.

According to Meyer (1970) the MTI was "designed to assess levels of interest in and attitudes toward various aspects of science in their schools. In addition, for purposes of contract, some assessments were made of interests in non-scientific subjects." This test, "an assessment of sophistication in scientific attitude" consisted of three subtests; Interest, Thinking, and Attitude. A main score called "Scientific

Orientation" or "S.O." was derived from the sub-tests. The weightings of the sub-tests in the S.O. score were:

T_1 Interest 44 percent

T_2 Thinking 28 percent

T_3 Attitude 28 percent

In addition, the "Interest in Science" variable was composed of the following three sub-variables:

S_1 1. Interest in a scientific hobby or leisure activity.

S_2 2. Interest in solving problems in science by practical activity in contrast to appeal to authority; that is interest in science as a method of solving problems.

S_3 3. Interest in science as a body of facts.

Further, the last subvariable is further divided into interest in various areas of scientific content: physics, biology, chemistry, astronomy, geology, and history of science.

In addition, the MTI gathered student responses as to their "Favored Sources of Scientific Information" (teachers, expert or book) and "Interest in Non-Science Leisure Activities" (literature or art).

The authors of this article submitted the intervariable correlation data from the two samples separately to factor analysis, "using the customary procedure: principal component analysis is followed by a Varimax rotation. Only factors with eigenvalues equal to or greater than one were retained." Data for variables S.O., S_3, and T_1 (described above) "were not included in the factor analysis since they are obtained by the addition of subscores on the Meyer test and hence do not represent 'individual scores'."

Findings

The findings of this study were summarized in a table of factor loading coefficients. Individual variables were judged to be significant contributors to a factor if the loading coefficient was greater than $\pm$ 0.3.

The first noticeable difference between the two sets of data was the number of factors obtained: four for the English sample and five for the Israeli. However, the fifth factor was composed of significant contributions only from the four additional cognitive tests administered to the Israeli students. None of the MTI variables loaded significantly on this factor.

Factor one from both samples "links together four of the main variables of the Meyer instrument: interest in science as a leisure activity (S_1) and as a problem-solving method (S_2), scientific thinking and attitude (T_2), and attitude to school science teaching (S_3)." Additionally, factor one had loadings from the several S_3 subvariables. With the Israeli sample, Factor I included Interest in Science as a set of facts - chemistry, while the English sample had loadings from the biology, chemistry and physics S_3 sub-variables. The percentage of variance explained by Factor I in the Israeli and English sample respectively was 11.3 percent and 25.5 percent.

Factor II for the Israeli sample was composed of loadings from all the S_3 sub-variables: physics, chemistry, biology, astronomy, geology, and history of science. The English sample had significant contributions on this second factor only from S_3 variables - astronomy, geology, and history of science. The variance explained through the Israeli and English Factor II loadings was 18.5 percent and 10.9 percent, respectively.

Factor III for both samples had loadings from the Interest in Nonscience Leisure Activities. In addition with the English sample, a loading from S_3 biology was identified. The proportion of total variance of each set of data explained by Factor III was 10.0 percent and 12.1 percent, respectively, with the Israeli and English samples.

The fourth factor with both samples was composed of significant loadings from the same variables; the favored source of scientific information - teacher, expert and book. This factor contributed 9.0 percent and 12.5 percent, respectively, to the total variance of the data with the Israeli and English samples.

Interpretations

Due to the separate loadings of the cognitive variables used with the Israeli sample, the authors claimed that "the essential independence of the cognitive and the affective variables is demonstrated." They claimed that this independence of the interest and attitude variables with the student's cognitive abilities and achievement is in agreement with other findings.

From the results obtained with both sets of data, general correspondence was observed in relation to the four MTI factors. Factor I, a general science interest/attitude factor linked together four of the main variables of the MTI. "The communality, in terms of the factor analysis, of these four variables suggests that in the psychological sense they are essentially 'unidimensional, i.e. that they may be reduced to a single underlying construct which, on the basis of the present evidence, does not allow for a differentiation between interest components and attitude components'." The authors also concluded that this finding "supports the validity of the scoring procedure for the test which leads to an overall science orientation (S.O.) score through a linear combination of scores on the science interest and attitude variables." This conclusion is enhanced by the similarity of the results with data from two separate samples.

Interest in science as a source of factual information (related to biology, chemistry, etc.) was apparently sensitive to the different curricular patterns in Israel and England. For the Israeli students, all these sub-variables loaded on Factor I, while they were split between Factors I and II for the English group. This corresponds with the separation for the English 0 - level science programs with biology, chemistry, and physics as the traditional disciplines taught, while astronomy, geology and history of science are treated as extra-curricular subjects. In Israel, most students begin studying individual science courses at grade 10 (two years later than England). This is apparently related to the finding of "no strong association of general science interest and attitudes with particular science subjects" for the Israeli students. The authors concluded that

> the independence of at least some of the science subject variables of the general science interest/attitude factor, as revealed by both sets of factor analyses, suggests that interest in science as a source of information cannot be considered a genuine component of the overall science orientation variable, except perhaps where such interest relates to science subjects actually studied by students at school level.

Although these "non-curricular" science variables do not contribute to the S.O. variable directly, they appear fairly "neutral." According to the authors, this would not distort the meaning of the S.O. scores by their incorporation.

The loadings with respect to Factors III and IV were nearly identical with two samples. The Humanities Factor (III) appeared "to be a mild anti-literature/art variable" and "interest in biological and history-of-science information appear as weak concomitants" of this factor. Meyer's assumption underlying the design of his test - "that consulting teachers, experts, and books is a genuine nonscience alternative to experimentation as method of problem solving" was not supported by the findings of this study as the loading of the problem solving sub-variable (S_2) on Factor IV (humanities) was only marginal.

In summary, this study reaffirmed the "general coherence among these science interest and attitude variables." It would appear "that a genuine differentiation between interest traits and attitudinal traits is not possible in terms of the Meyer constructs." Therefore, the recommended scoring procedure - aggregating the interest and attitude sub-scores into an overall science orientation score - appears justified.

It appears that the MTI and its scoring has considerable internal validity. The very similar results from different countries enhances this view. "However, a factor-analytic study of a test cannot provide evidence of its validity as a measure." The question as to what extent the MTI and other science-affective instruments measure common constructs still remains.

ABSTRACTOR'S ANALYSIS

This study was a re-examination of data, some previously gathered. Therefore, the review of literature was very limited. Considering the difficulty in conceptualizing and measuring student interest in science, brief mention of the few other related studies may have been helpful. The Reed Science Activity Inventory (Cooley and Reed, 1961) and its modification, the Science Activities Checklist (Skinner and Barcikowski, 1973) and the Semantic Differential Interest Test (Butzow, 1974) are examples.

This study emphasized an in-depth, empirical analysis of the instruments we use to assess important variables in science education. The use of the analysis procedures and criteria (eigenvalue and factor loading) was correct if one is justified in using data from eight Likert items as a variable in parametric analyses. The number and kind of items was not described in this article. For this information, means of variables and descriptions of the samples, the reader was referred to the earlier reports. It would seem room could have been made for a brief description of the items.

The data analyzed in this study were from two samples: one in England (1970), the other in Israel (1976). While such divergence in years might seem unusual, the fact that the data patterns were so similar seems to minimize one's concerns. The reliability estimates with the Israeli sample of 0.71, 0.76 and 0.73 (for the three subscores) did not agree with their other report (Hofstein, et al., 1977) which only listed the MTI reliability as 0.75. Since the basic article also had a reordering of the authors, the presentation of reliability data may have been changed during revisions. While the numbers are all very consistent, the concern should be to show reliability estimates for the variables used in the analysis.

A typographical error on the last line of p. 65 (S_3 should have been T_3) was not a major distraction. The data in Table I were much too voluminous to use - indeed the factor loading in Table II presented the inherent patterns and relationships. The lack of analysis due to sex differences is puzzling, especially as it was used by Meyer (1970) and Kempa and Dube (1974) on MTI data.

The major comments are about the conclusions related to the uni-dimensionality or multi-dimensionality of science interests. At best, a single study can contribute to the evidence being collected in those issues - not settle the issue. Data collected with the Reed Science Activity Inventory and its modifications have been interpreted as consistent with the multi-dimensionality of science interest. Interestingly, the data they used were also factor pattern coefficients as in this study. In both studies, the responses were loaded on several factors. Cooley and Reed interpreted this as evidence for several dimensions of science interest, while Hofstein, et al., discuss the MTI as "providing measures of several facets of interest in science..." The Hofstein, et al., study found evidence of the internal structure of the MTI - generally consistent with that described by Meyer (1970). The first factor (with both the English and Israeli samples) included loadings from the Interest and the Attitude variables, prompting the conclusion that "in the psychological sense they are essentially uni-dimensional." It appears that one's interpretation is partially determined by one's a priori disposition "to lump or split."

Another area worthy of comment relates to the dependence or independence of cognitive and affective traits. The data from the Israeli sample on four cognitive tests loaded on a separate factor. This prompted the authors to state "hence, the essential independence of the cognitive and affective variables is demonstrated." They claimed that this outcome "is in line with other findings" (Evans, 1971). Kempa and Dube (1974) examined "the dependence of students' behavior on their academic achievement." This was considered important in view of the evidence contained in the research literature (Evans) that interest traits and academic achievement are positively correlated. Kempa and Dube felt that "Although the correlations tend to be low, their existence must be acknowledged and allowed for in the interpretation of affective measures..." Apparently, weak empirical results can be interpreted in different ways. It appears very clear that more and cleaner research in education is necessary before sweeping generalizations are claimed.

As recognized by Hofstein, et al., "a factor-analytical study of a test cannot provide evidence of its validity as a measure." It is in this regard that theoretical as well as empirical emphasis is needed. A careful analysis of the affective domain is essential to determine uniquely what "interests" are. The editor of the Affective Taxonomy suggested that interest related to the first three levels - Receiving, Responding and Valuing. Campbell (1972) used these three levels in constructing a Scientific Curiosity Inventory. Butzow (1974) attempted to construct an Interest Scale utilizing a semantic differential format. These and other sources in psychological and educational literature need to be examined by all interested in assessing student interest in science.

It is the hope of this reviewer that work such as that described by Hofstein, et al., and others continues in an effort to assess those most elusive but important aspects of our science programs.

REFERENCES

Butzow, John W. "A Semantic Differential Science Interest Test." School Science and Mathematics, 74: 189-196, 1974.

Campbell, James Reed. "Is Scientific Curiosity a Viable Outcome in Today's Secondary School Science Program?" School Science and Mathematics, 72: 139-147, 1972.

Cooley, William W. and Horace B. Reed, Jr. "The Measurement of Science Interests: An Operational and Multi-Dimensional Approach." Science Education, 45: 320-6, 1961.

Evans, K. H. Attitudes and Interests in Education. London: Routledge and Kegan Paul, 1971, Chap. X.

Hofstein, Avi; Ruth Ben-Zvi; David Samuel; and Richard F. Kempa. "Some Correlates of the Choice of Educational Streams in Israeli High Schools." Journal of Research in Science Teaching, 14(3): 241-7, 1977.

Kempa, R. F. and G. E. Dube. "Science Interest and Attitude Traits in Students Subsequent to the Study of Chemistry at the Ordinary Level of the General Certificate of Education." Journal of Research in Science Teaching, 11(4): 361-370, 1974.

Meyer, G. R. "Reactions of Pupils to Nuffield Science Teaching Project Trial Materials in England at the Ordinary Level of the General Certificate of Education." Journal of Research in Science Teaching, 7(4): 283-307, 1970.

Skinner, Ray, Jr. and Robert S. Barcikowski. "Measuring Specific Interests in Biological, Physical and Earth Sciences in Intermediate Grade Levels." Journal of Research in Science Teaching, 10(2): 153-158, 1973.