DOCUMENT RESUME

ED 221 360 SE 039 116

AUTHOR	Blosser, Patricia E., Ed.; Mayer, Victor J., Ed.
TITLE	Investigations in Science Education, Volume 8, Number 3.
INSTITUTION	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.; Ohio State Univ., Columbus. Center for Science and Mathematics Education.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	82
NOTE	80p.
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; v8 n3 1982
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Achievement; *Cognitive Processes; College Science; Concept Formation; Educational Research; Elementary School Science; Elementary Secondary Education; Engineering; Females; Higher Education; Science Curriculum; *Science Education; *Science Instruction; Secondary School Science; *Student Characteristics; *Teacher Education
IDENTIFIERS	Cognitive Preference; *Science Education Research

ABSTRACT

Abstractor's analyses of science education research studies are presented in this issue of Investigations in Science Education. The first group of studies relates to research on teacher education (priority of perceived needs common to science teachers and how these might be used to develop inservice programs, determining the level of awareness and use of population education materials by teachers, and behaviors of preservice elementary teachers attempting to acquire science process skills). The second group focuses on achievement (teaching-learning model used with culturally deprived students and predicting success of college biology students using the Cloze test). The third grouping focuses on cognitive variables (mathematical model of concept learning applied to science instruction, cognitive preferences of talented science students, relationship between academic performance and school-related affective characteristics of 12-year-olds, and modification of cognitive style of graduate students). Studies addressing the effect of women students in engineering on their institution's resources (Do women engineering students present different problems for their colleges/universities than do male students?), science interests of junior high school students, and science needs assessment for use with elementary school teachers are analyzed in the final section. (Author/JN)

SCIENCE EDUCATION INFORMATION REPORT

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

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC® Clearinghouse for Science, Mathematics, and Environmental Education

INVESTIGATIONS IN
SCIENCE EDUCATION

Volume 8, Number 3, 1982

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.

INVESTIGATIONS IN SCIENCE EDUCATION

Volume 8, Number 3, 1982

Volume 8, Number 3, of the ISE contains several clusters of articles as well as several individual articles. One grouping relates to research on teacher education. Moore and Blankenship reported on an investigation of the priority of perceived needs common to science teachers and how these might be used to develop inservice programs. O'Brien and her colleagues were interested in determining the level of awareness and use of population education materials by teachers. Sunal looked at the behaviors of preservice elementary teachers as they attempted to acquire science process skills.

A second grouping, achievement, contained two articles. One, by Sabar and Kaplan, looked at a teaching-learning model as it was used with culturally deprived students. Cohen and Poppino used the Cloze test with college students enrolled in an introductory biology class as a method of prediciting success in the course.

The third grouping contains studies focused on one or more cognitive variables. Treagust and Lunetta examined a mathematical model of concept learning as this was applied to science instruction. Tamir and Lunetta studied cognitive preferences of talented science students. Marjoribanks examined the relationship between academic performance and school-related affective characteristics of 12-year olds. Pringle and Morgan attempted to modify the cognitive style of graduate students.

In the final section of this issue a variety of research is reported. Gardner studied the effect of women students in engineering on their institution's resources. In other words, do women engineering students present different problems for colleges and universities than do male students? Palmer investigated the science interests of junior high school students. Moore reported on a science needs assessment for use with elementary school teachers.

Certainly, if one is looking for indications of variety in science education research published in professional journals, this issue provides such evidence.

Patricia E. Blosser
Editor

Victor J. Mayer
Associate Editor

TEACHER

EDUCATION

Moore, Kenneth D. and Jacob W. Blankenship. "Relationships Between Science Teacher Needs and Selected Teacher Variables." Journal of Research in Science Teaching, 15(6): 513-518, 1978.

Descriptors--Educational Research; General Science; Inservice Education; *Needs; Physical Science; *Qualifications; Science Education; *Science Teachers; Teachers

Expanded abstract and analysis prepared especially for I.S.E. by Herbert A. Smith, Colorado State University.

Purpose

The study was intended to identify the priority of perceived needs common to science teachers and to identify relationships among such needs and selected teacher variables.

Rationale

The study is an outgrowth of the widespread criticism of inservice education programs and the alleged failure of such programs to identify teacher needs and the differences in the needs of teachers in different classification categories. It is implicitly assumed that programs more specifically addressed to the perceived needs of teachers will be more effective and secure the desired increase in the quality of science instruction.

Research Design and Procedures

A questionnaire study was conducted in 1977 of teachers in a single Texas county and included teachers from 21 districts. A stratified random sample of 500 teachers was selected in three categories: elementary (200), junior high school (150), and senior high school (150). A total of 283 teachers responded. Data were collected through use of a previously developed instrument designated the Moore Assessment Profile. Factor analytic techniques were used to identify 11 areas as "factor needs." These were denoted as:

I Developing basic scientific reasoning skills.

II. Planning science instruction.

III. Providing for students with differing cultural and social background.

IV. Providing realistic science experiences.

V. Updating in science content.

VI. Improving classroom management skills.

VII. Evaluating and reporting student science progress.

VIII. Developing a greater understanding of human behavior.

IX. Training in science teaching methodology.

X. Updating in educational foundations and technological application to education.

XI. Improving student guidance.

Findings

Perhaps the major finding was that the priority of perceived intensity of needs of teachers to be served is related to selected teacher variables. More specifically, elementary teachers perceive a need for help in providing realistic science experiences, relatively inexperienced teachers perceive a need for training in science teaching methodology, primary teachers have priority needs for help in planning science instruction, and physics teachers identify a priority need for updating science content. Teachers had in common a need for assistance in developing scientific reasoning skills.

Interpretations

The investigators suggest that the study has implications for the developers of inservice programs. The point is made that developers need to be cognizant of differences in the needs of elementary and secondary teachers and of differences related to experience and discipline or levels taught. To the investigators, it is clear that inservice programs should be more clearly targeted to the needs of individual teachers with the result that a general upgrading of science instruction will result.

ABSTRACTOR'S ANALYSIS

It is probable that the "factors" identified represent general areas of need and they are probably very similar to a list which a well-informed science educator would have jotted down on the basis of an hour's reflection. But it is quite possible that the problem is not so much in the content of the inservice program, but rather in its quality. The factors identified are so broad that they provide little in the way of specific guidance to the developers of inservice programs. One might legitimately ask what Factor V, "Updating in Science Content," really means. Evidence clearly shows that the amount of training in the discipline, particularly in such highly structured disciplines as physics and mathematics, is not highly correlated with student achievement. If an inservice trainer took Factor V (or any other) very seriously, he would still be faced with the question of an appropriate translation of such a factor into an instructional program. It is perhaps not unreasonable to view the translation and its quality as the all-important element in gauging the effectiveness of the inservice program. The research reviewed does not address this problem.

A major assumption in the study is that statements about "perceived needs" can be accepted as a valid measure of "real" needs. Although "needs assessments" have often been made on the basis of statements from a variety of audiences relative to their perceptions about such need, the validity of such assessments may well be in question. Psychological research provides ample evidence that what people say they will need or do, and what they actually need or do when confronted in actual situations is not necessarily the same. Individual responses to questions about their attitudes to controversial areas may not square with other more tangible evidence provided by their actions in a variety of settings. Thus, in general, verbal statements about a variety of domains do not necessarily relate highly to other more objective measures of reality and "needs" is probably no exception. In addition, perceived need statements may also be contaminated by community expectations, the current educational fads and such individual eccentricities as ego, fear, pride, pressure to conform, and by a variety of other factors. Certainly it is time to have evidence adduced which shows the degree of relationship between perceived needs statements and other more objective measures of need. In addition, the nature of the questionnaire and its structure reflect the values, concepts, and biases of the designers and responses to the instrument are necessarily constrained to the structure provided.

There remains the question, too, as to whether inservice programs are really less effective than they used to be. Institutions have competed in offering off-campus courses, weekend programs, evening classes, and other adaptive-type adjustments to "accommodate the clientele." Unfortunately, it is very doubtful that quality and standards have generally been maintained. In attempting to be adaptive, institutions may have only cheapened their programs. They have also not adequately distinguished between appropriate inservice training and genuine graduate study. Teachers may have responded, too, by taking the new and easier paths to alleged professional competence but have found that programs of convenience have not led to the higher plane of professional competence which they desired. In summary, it is possible that ineffective inservice programs are a reflection of the general educational malaise of the times: lowered standards, grade-inflation, cheapened degrees, lower morale, and a general deterioration of what has historically been termed "professionalism."

If these observations are reasonably correct, one may question whether research designed to evoke opinions about perceived need is likely to contribute much to the improvement of science instruction.

O'Brien, Patricia, Carl Huether, and Susan Philliber. "Teacher Knowledge and Use of Population Education Materials: Report from National Surveys." Science Education, 62(4): 429-442, 1978.

Descriptors--Curriculum; *Demography; *Educational Research; *Instructional Materials; *Population Education; Research; Science Education; *Surveys; Teacher Education; Teachers

Expanded abstract and analysis prepared especially for I.S.E. by Gerald Skoog, Texas Tech University.

Purpose

This study was designed to determine the level of awareness and use of 20 different sets of population education materials among population education teachers.

Rationale

In response to the growing awareness of the importance of population education, funding for the development of population education materials increased in the 1970s. The level of teacher awareness and use of these materials is a measure of the effectiveness of the strategy of diffusing population education through material development and dissemination.

Research Design and Procedures

Data were collected through two surveys using questionnaires designed to measure the level of teacher use and awareness of specific population education materials. In 1974, population education teachers were identified and surveyed on the basis of one of the following criteria: 1) participation in a population workshop or training program; 2) subscription to the Population Reference Bureau publication Interchange; and 3) recognition as a population education teacher by a principal in one of 1000 randomly selected secondary schools. Questionnaires were returned by 593 teachers or 47 percent of those surveyed.

In 1976, a list of secondary school science teachers maintained by the National Science Teachers Association was used to select a sample of teachers stratified on the basis of whether they taught general science, environmental and earth science, social studies, or biology. Questionnaires were returned by 1709 teachers or 46 percent of those surveyed. Data from 803 or 47 percent of the questionnaires were used as 906 of the respondents were not considered population education teachers.

In this survey, two bogus materials were listed to check respondent overstatement of material awareness and use.

Information regarding the cost, distribution, and promotion of the materials was collected from the publishers involved.

Findings

Population education teachers in both samples reported limited awareness and use of the population education materials listed. Fifty percent or more of the teachers were familiar with 3 of the 16 materials listed in 1974 and 2 of the 20 listed in 1976. In 1974, an average of 61 percent of the teachers were not familiar with these materials. This percentage was 70 in the 1976 survey. No source was used by more than 37 percent of the teachers in either survey. The average percent use of the materials was 14.9 and 11.6 in the 1974 and 1976 surveys, respectively.

The data on awareness and usage of materials may have been inflated by at least 10 percent due to the overstatement of knowledge by the respondents. In the 1976 survey, 14.7 percent of the sample reported using one of the bogus materials. This bogus item was familiar to 42.1 percent of the sample. The other bogus item was recognized by 32.4 percent of the teachers and used by 12.7 percent.

Data collected in the 1976 survey indicated the average difference between trained and untrained teachers in material use was 5 to 6 percent. Because this was the same difference that appeared in the bogus material data, seven materials that showed more than 6 percent difference in use by the trained teachers were identified as probably in greater use because of training received by the teachers.

Social studies teachers reported using the materials more than the biology teachers, in both the trained and untrained categories, in the 1976 survey.

Six materials were found to be used most by the teachers. Consideration of factors such as price, nature of the materials, dissemination strategies, and availability did not provide any answer to why some materials were used more than others.

Interpretations

The population education materials identified in this study were not being used. The percentage of teachers using the most popular materials was very low. There was some evidence that training in population education increased the use of the materials. The data suggested that the prevalent policy of concentrating funds on material development and providing little attention to distribution strategies had not been efficient or successful. A need to distribute the materials in greater numbers and to provide teachers help in articulating the materials into the existing curriculum was noted.

ABSTRACTOR'S ANALYSIS

In view of the existing research and literature on school change processes and the implementation of curriculum projects and materials, it was not surprising that the data in this study revealed population education materials were not recognized or used by many teachers. The materials listed in the study surveys were developed by textbooks publishers, professional education organizations, and groups concerned with population education. There is growing evidence that the centralized approach to curriculum change and reform, where there is a top-down flow of products, strategies, and advice has not been effective. This "professionalization of reform," where professionals in specific fields attempt to foster change, is not congruent with the realities of schools, how they are changed, and the role of public demand in institutional reform (Boyd, 1978). Until these realities are recognized and accommodated studies such as this, which show that instructional materials have not diffused out to a critical mass of educators, will continue to appear.

Caution is needed in accepting reported use as a measure of implementation, as done in this study. The data in this study were suspect due to the number of teachers reporting the use of nonexistent materials. Also, individuals reporting use of an innovation may be communicating an attitude of acceptance though they lack the knowledge and skills needed to implement the innovation (Fullman and Pomfret, 1977). As a result, a reported user often is a nonuser. Furthermore, implementation of the type described in this study is not a bipolar use/nonuse phenomenon as there are various levels of use that can be distinguished through direct observation and assessment (Hall and Loucks, 1977).

In this study, there can be no argument with the conclusion that the population education materials were not being used widely. However, because of the problems with use/nonuse research designs, little can be concluded from knowing that 37.1 percent of the teachers in the 1974 survey reported they used the World Population Data Sheet.

Large numbers of population education teachers in this study were not familiar with the materials listed in this study. Other studies have noted the same situation. A study sponsored by the National Science Foundation reported that 43 percent of the science teachers surveyed indicated they did not receive adequate assistance in obtaining information about instructional materials (Weiss, 1978). The same percentage of teachers reported a need for help in learning new teaching methods. Mechanisms are needed that will assist teachers in learning about new materials and methods. It is obvious the mechanisms developed must involve teachers locally over an extended period of time. There are some questions in regard to where the leadership should emanate. Curriculum coordinators are an apparent source of leadership. However, 63 percent of the school districts in this nation do not have science coordinators (Weiss, 1978). Also, a recent study reported district consultants, along with state and district guides, commercially prepared materials, and even textbooks, had a low level of influence on teachers as they made curriculum decisions (Klein, et al., 1979). Over 75 percent of these teachers indicated their decisions about curriculum were influenced most by their own background, interests, and experiences and the interests and abilities of the students. It is obvious that if a change is to occur, teachers must be involved in all phases of curriculum development and implementation. They must have extended and repeated opportunities to communicate with others who are involved successfully with new programs or materials. Strategies used must recognize that change is a personal process where there must be developmental growth in both feelings and skills if success is to be achieved (Loucks and Pratt, 1979).

This study made no new conceptual or methodological contributions to the study of implementation and/or adoption practices. There was no review of related research or any attempt to relate the findings to other research. Population education teachers were not defined and were selected for the 1974 survey if they had participated in a population education workshop, received a specific population education newsletter, or were named by a principal as a population education teacher. The written report was understandable and the annotated bibliography listing the population education materials was useful.

REFERENCES

Boyd, William. "The Changing Politics of Curriculum Policemaking for American Schools." Review of Educational Research, 48: 577-628, 1978.

Fullman, Michael and Alan Pomfret. "Research in Curriculum and Instruction Implementation." Review of Educational Research, 47: 335-397, 1977.

Hall, Gene and Susan Loucks. "A Developmental Model for Determining Whether the Treatment is Actually Implemented." American Educational Research Journal, 14: 263-276, 1977.

Klein, M. Frances, Kenneth Tye, and Joyce Wright. "A Study of Schooling: Curriculum." Phi Delta Kappan, 61: 244-248, December 1979.

Loucks, Susan and Harold Pratt. "A Concerns-Based Approach to Curriculum Development." Educational Leadership, 37: 212-215, December 1979.

Weiss, Iris. Report of the 1977 National Survey of Science, Mathematics, and Social Studies Education. Washington, D.C.: U.S. Government Printing Office, 1978.

Sunal, Dennis W. "Relationships of Science Skill Performance to Preservice Teaching Behavior." Science Education 62(2): 187-194, 1978.

Descriptors--College Science; *Educational Research; Higher Education; *Preservice Education; *Process Education; *Science Education; Science Teachers; *Teacher Behavior; Teacher Education.

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Charles L. Price, Indiana State University,Evansville.

Purpose

The purpose of this study was to determine if there was a relationship between the ability to perform science process skills and teaching behaviors of preservice elementary teachers. Specifically, the following questions were investigated:

1. What relationship exists between the level of ability of preservice teachers to perform science process skills upon entry into a science methods course and teaching behavior at the end of the methods course and during student teaching?

2. What relationship exists between the level of ability of preservice teachers to perform science process skills measured at the end of a science methods course and teaching behavior at the same time and during student teaching?

Rationale

Previous studies by Jaus (1975) and Kennedy (1973) reported that process skill instruction improved performance in process skills, enhanced attitudes toward science and increased the rate of inclusion of process skill activities in lesson plans. Ashley (1967), Butts and Raun (1969), and Newport and McNeill (1970) found that for teachers who received training in the use of the Science--A Process Approach (SAPA) Program, changes in teaching behaviors and attitude toward the use of process skills in SAPA classroom lessons occurred.

Research Design and Procedure

A one-time series design (Campbell and Stanley, 1963) was utilized:

O X O O

O - SPM Pretest	O - SPM Posttest MSS Checklist	O - SOCAS Questionnaire

Instrumentation consisted of: (1) Prestest and Posttest versions of the Science Process Measure (SPM)--18-item exams which covered all process skill areas experienced by the subjects; (2) Microteaching Skills in

Science (MSS) Checklist--an inference-based observational instrument concerned with lesson planning and evaluation; and (3) Survey of Classroom Activities for Science (SOCAS) Questionnaire which measured subjects' perceptions of teacher behaviors and student process skill behaviors of classroom science lessons. The SPM instrument was developed by the Commission on Science Education, American Association for the Advancement of Science (1969), and the MSD and SOCAS instruments were developed and validated by the investigator.

A stratified random sample of 42 Ss were selected from 169 undergraduate students enrolled in six elementary science methods courses. For final analysis, complete data were available for 36.

The selected sample of students in the six courses was given similar course activities (treatment) in regards to time and orientation of activities toward the philosophy of the new science curricula. One-third of the class time was devoted to activity labs and discussion of science learning theory. Other class activities included instructional methodology, elementary science curricula, and learning theory. As a concluding activity, all students planned and taught a three-lesson science unit for elementary children.

Findings

Subjects' SPM pretest and posttest scores were found to be significantly correlated ($p < 0.001$) and pretest to posttest comparisons showed significant gains ($p < 0.001$). MSS checklist and SOCAS questionnaire results were weakly correlated with SPM pretest and posttest scores; no value was significant at the 90 percent confidence level.

Interperetations

The results of this study indicated that the level of ability to perform science process tasks was not related to those preservice teachers' classroom behaviors consistent with new elementary science curricula. Subjects in the study were found to have increased process skill knowledge by the prescribed treatment; this ability, however, was not significantly related to overt classroom behaviors.

ABSTRACTOR'S ANALYSIS

In the introduction for the study the author notes that "to date there have been no studies of the relationship of process skill instruction on the teaching of preservice teachers in classroom lessons" (p. 188). There certainly is a need for research in this field. Program reviews of science methods courses for preservice elementary teachers seldom reveal a syllabus which does not call for the development of process skills. For what end, save the attainment of process skill abilities, does the training of prospective teachers in process skills serve? This study represents an investigation into the relationship of process skill attainment and classroom behaviors identified as being congruent with teaching philosophies espoused by "new" science curricula.

Design and Methodology For the study, a stratified random sample of subjects from six undergraduate science methods courses was selected. While this random selection procedure is admirable, the validity of the study would have been greatly enhanced by the inclusion of a control group, with subjects randomly assigned to groups. As noted by Campbell and Stanley (1963), a host of sources "offer plausible hypothesis explaining a 0 - 0 difference, rival to the hypothesis that X caused the difference" (p. 7).

Two investigator-developed instruments, the Microteaching Skills in Science (MSS) Checklist and the Survey of Classroom Activities for Science (SOCAS) Questionnaire were utilized to measure "teaching behavior consistent with the new elementary curricula" (p. 190). This reviewer ponders whether there exist behaviors which fit under this rubric; if, instead, the behaviors of a classroom teacher utilizing materials ascribing to the viewpoints of Jerome Bruner may not be sigificantly different from a hierarchical approach a la Robert Gagne?

Descriptions of the educational settings in which subjects performed their student teaching were not provided. A 1974 ERIC study of educational practices revealed that 27 percent of the elementary school sample were using at least one science course improvement project.

Perhaps these conditions can in part be used to explain the low observed correlations in science process test and MSS Checklist and SOCAS Questionnaire scores. It is interesting to note that the largest correlation was -0.26 for SPM posttest and MSS (Teaching Behaviors in the Classroom) checklist. Although short of statistical significance ($N = 36$, P $0.05 = 0.34$), the finding reveals a negative correlation between the variables.

Suggestions. The need to search for observable outcomes of preservice science methods instruction is obvious. It is hoped that instruction in the areas of questioning skills, evaluation, process skills and the host of other topics studied in a methods class leads to more than improved knowledge in these topics. Perhaps the inferential leap that training in any one of these areas leads to changes in overall teacher behavior is too great; instead each of these areas and its relationship to a specific teacher behavior should be considered.

REFERENCES

Ashley, J. P. "A Study of the Impact of an Inservice Education Program on Teacher Behavior." Unpublished dissertation, The University of Texas, Austin, 1967.

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

Campbell, Donald T. and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Company, 1963.

Commission on Science Education, American Association for the Advancement of Science. Science Process Measures for Teachers, Publication 69-6. 1515 Massachusetts Avenue, N.W., Washington, D.C.: U.S. Government Printing Office, 1969.

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

Newport, J. F. and K. A. McNeill. "A Comparison of Teacher-Pupil Verbal Behaviors Evoked by Science--A Process Approach and by Textbooks," Journal of Research in Science Teaching 7(3): 191-195, 1970.

Steiner, Robert L., et al. A Survey of Science Teaching in Public Schools of the United States, Volume 4--Elementary Schools. Columbus, Ohio: ERIC Science, Mathematics and Environmental Education Clearinghouse, The Ohio State University, 1974.

Esquivel, Juan M.; William S. Lashier; and Walter S. Smith. "Effect of Feedback on Questioning of Preservice Teachers in SCIS Microteaching." Science Education 62(2): 209-214, 1978.

Descriptors--College Science; *Educational Research; *Feedback; Higher Education; *Preservice Education; *Science Education; Teacher Attitudes; *Teacher Behavior; Teacher Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Joseph P. Riley, II, University of Georgia.

Purpose

This study investigated the effects of self, peer and supervisory feedback on preservice teachers' proportion of high inquiry questions, probing questions, wait-time, and attitudes toward source of feedback and science teaching.

Rationale

Previous research investigating the effectiveness of various feedback sources has been conflicting. No definitive conclusions about the effect of feedback sources have been made.

Research Design and Procedure

Ninety-two undergraduate preservice elementary teachers were instructed in how to use high order questions, probing questions and wait-time species two. After this instruction, all subjects were randomly assigned to pairs and these pairs were randomly assigned to teach groups of six to ten third, fourth, or fifth graders. Three lessons were taught and audio-recorded to provide a basis for feedback. The lessons were based on Science Curriculum Improvement Study (SCIS) physical science lessons. After each lesson the subjects were given feedback from the source appropriate to their treatment group. Subjects received feedback about their lesson either from themselves, a peer, or a supervisor. A fourth group of subjects served as a control group and observed pupil behaviors which were not related to the criterion outcomes. A form requesting the frequency (often, occasionally or rarely) of low order questions, high order questions, wait-time species one and two, and probing questions was used in each treatment group. The control used a different form focusing on other competencies. Following the treatment feedback lessons, each subject taught a 30-minute non-SCIS lesson involving either air pressure or heat conduction. Audiotapes of these lessons provided the data for the three teacher behavior dependent variables.

The design can be represented as:

R	X_1	0_1	0_2	0_3	0_4
R	X_2	0_1	0_2	0_3	0_4
R	X_3	0_1	0_2	0_3	0_4
R	X_4	0_1	0_2	0_3	0_4

where;

X_1 = self-feedback group-subjects listened to their own tape and completed a questioning critique form

X_2 = Supervisory feedback group-same procedure as treatment one but done with the subject's supervisor

X_3 = peer feedback group-same procedures as above but a peer took the place of the supervisor

X_4 = control group-followed the same procedures as the self-feedback group but completed a critique form which avoided questioning strategies

and

0_1 = proportion of high order questions

0_2 = proportion of probing questions

0_3 = duration of wait-time species two

0_4 = attitude toward science teaching

0_5 = attitude toward treatment

Seventy-one tapes were analyzed by two raters. Interrater reliability was found to be 0.76 for high order questions, 0.83 for low, 0.83 for probing, and 0.46 for wait-time.

In order to improve the reliability of the wait-time measure the principal investigator rated 12 random tapes and calculated a reliability coefficient with each rater. One was found to be 0.83; the other, 0.29. The rater's data with the highest correlation was used as the basis for analysis. An average of each rater's results was used as the raw data for the other questioning variables.

Analysis

In order to test the significance of the differences among the four groups on the five dependent variables, a one-way multivariate analysis of variance was employed.

Results

The obtained X^2 was not significant. The study failed to reject the null hypothesis of no difference among groups on the five dependent variables.

Conclusions

The authors reported that the results supported previous findings indicating that the source of feedback to teachers makes no difference in the teachers' demonstration of certain behaviors.

ABSTRACTOR'S ANALYSIS

The problem addressed in this study is a practical and important one. Field-based teaching experience is generally thought of as the keystone to an effective teacher education program. While there are many advantages to field-based experiences, one disadvantage appears to be the time demand these programs place on the field supervisor. If research can identify teaching competencies that may be effectively learned through self, peer or pupil feedback, then the supervisor's time might be better allocated in a more cost-efficient manner. The results of this study provide some evidence that certain questioning skills may be as effectively taught by peer and self feedback as by supervisor feedback. The strength of this evidence is influenced by a number of factors.

The study utilizes a posttest-only control group research design. This design eliminates the need for a pretest by equating groups through randomization. However, readers must infer that the subjects were actually randomly assigned to treatments. The authors do report that the subjects were randomly assigned to pairs and to groups of students. Assuming random assignment to treatment, the design is well suited for this particular study.

A problem with instrumentation was described as well as steps taken to correct it. Interrater reliability of two raters on the wait-time measure was found to be unacceptably low (0.46). The principal investigator rated a random sample of 12 tapes and then chose for analysis purposes the results of the rater correlating highest with his own results (0.83 vs 0.29).

The problem with this procedure is that one remains unconvinced about the "true" reliability of the measure. The results of the rater not chosen could plausibly represent the true condition. This problem could be avoided by measuring wait-time with a chart-recorder. There are few opportunities in educational research to measure variables with the precision possible in measuring wait-time using the chart-recorder. The reliability of measurements of a variable as measurable as time should not be as low as 0.29 or 0.46 nor even as low as 0.83. Stopwatch measurements of wait-time are unrealistic, unreliable, and unnecessary.

The authors offer three alternative explanations for their findings. The first two pose threats to the internal validity of the study: 1) the source of feedback may not have varied among groups, the authors could give no assurance that the subjects received feedback from the desired source and only that source; 2) the subjects were in pairs and may have received vicarious, additional instruction; and 3) the concern with the dependent variables may have been overshadowed by survival concerns of these preservice teachers. The first two explanations can be described as multiple treatment interference. They point up the need for monitoring treatments so that procedures are congruent with intents. The third explanation points to a possible area of further research.

The strength of this study's evidence must also be evaluated in light of the fact that the null hypothesis failed to be rejected. How does one interpret failure to reject the null? Most statistics textbooks warn that one never proves H_0, but only disproves it in a probabilistic sense. The finding of no significant difference indicates that the evidence is not sufficiently strong enough to safely exclude chance as a possible explanation. Zero continues to be among the range of plausible parameter values (Hopkins, 1973). Because of this, the authors have been rightfully cautious in stating their conclusions. They state their study supports previous findings and suggest, rather than conclude, that feedback in this study does not seem to differentially affect a teacher's learning a particular behavior.

The contributions of this research is not so much in the strength of its findings as in its procedures. The study is a true experimental investigation with procedures and findings clearly reported. It points out some fruitful, and a few less than fruitful, directions for further research. It clearly builds on previous work and should strengthen future studies in this needed area of research.

REFERENCE

Hopkins, Kenneth D. "Preventing the Number One Misinterpretation of Behavioral Research, or How to Increase Statistical Power." The Journal of Special Education, 7(1): 103-107, 1973.

ACHIEVEMENT

Sabar, Naama and Eugene H. Kaplan. "The Effect of a New Seventh-Grade Biology Curriculum on the Achievements and Attitudes of Intellectually and Culturally Heterogeneous Classes." Journal of Research in Science Teaching, 15(4): 271-276, 1978.

Descriptors--*Achievement; Attitudes; Biology; *Disadvantaged Youth; *Individualized Curriculum; *Instruction; Junior High School Students; Mastery Learning; *Science Education; *Secondary Education; Teaching Models

Expanded abstract and analysis prepared especially for I.S.E. by David H. Ost, California State College, Bakersfield.

Purpose

The authors state the research problem as "Does the new teaching-learning model significantly improve the affective and achievement performance functions in the disadvantaged* segment of the heterogeneous class, while maintaining the performance levels and attitudes of the average and above-average IQ student?"

Rationale

The research problem is closely related to work previously published by the authors (Kaplan and Sabar, 1975a, 1975b). This paper reports on a model biology curriculum which was devised in response to the 1964 efforts of the Israel Ministry of Education to reform programs in an effort to provide quality education to heterogeneous groups of students. Supposedly the model was to maximize the performance of all students; encouraging both culturally deprived and high ability students to perform to the best of their abilities.

It should be noted that one of the major objectives for using the model was to change teacher behavior from instructional strategies characterized as authoritarian to individualized. This change was believed to be an important element for dealing with heterogeneous classes. Although it is reported that teachers were provided an

*The term "disadvantaged" is used synonymously with "culturally deprived."

in-service training course, the content or success of the effort is not reported.

The model proportedly utilized a variety of teaching strategies and related pedagogical techniques in an effort to respond to the needs of students. Reading deficiencies, poor motivation, inadequate self-discipline and a general lack of reasoning ability characterized the specific target group described as "culturally deprived." The authors argue that students with lower I.Q.'s within one standard deviation below 100 were labeled as "gifted culturally deprived." The authors use cultural and I.Q. heterogeneity synonymously.

Research Design and Procedures

From the little information provided in this paper it is concluded that the researchers employed what Campbell and Stanley described as the Nonequivalent Control Group Design. Although pre- and posttests were used, the students were members of established classes in different schools. There is no indication that assignments to the experimental and the control groups were random. Approximately 154 students were in the experimental group and 251 individuals represented the control. For some reason the control was divided as to classes in the same school as that of the experimental group, classes in other schools serving culturally deprived students, and classes in an urban comprehensive school. The control groups (students) studied the original biology program. The earlier papers of Kaplan and Sabar provide considerably more information about the populations and the model. The 1978 paper is merely a massaging of old data.

The authors write, "Because of the relatively small number of classes and teachers, all data were pooled and the variables of schools and classes were statistically controlled." The reader is neither given information regarding those variables nor is there a discussion of the statistical methods employed.

Scores on two tests were used as the dependent variables. Both instruments were administered during the first and last weeks of the semester. The achievement test, a multiple-choice test of 54 items, purportedly evaluated knowledge of concepts common to the traditional program and the experimental curriculum. Since all content was drawn from the same list of concepts published by the Ministry of Education, differences between the programs were primarily in the area of instructional philosophy and strategies. Thus, differences between the control and experimental groups are attributed to teacher philosophy and methods.

The 54 multiple-choice items used in the achievement test were classified by a consensus of five judges. Two sets of criteria were used: (1) level of importance in comprehending principles of ecology; and (2) level of mental functioning according to Bloom's Taxonomy of Educational Objectives. The authors somehow divided the total test into four subtests using the above criteria. No indication was given as to the number of items in each subtest.

The "Attitudes Towards Biology" test was used as a means to measure shifts in attitudes as a result of the new curriculum. This test was/is a modified version of an instrument developed by Wilson, Cohen and Begle (1968) to test attitudes towards mathematics. Students were asked to judge their attitudes towards biology on a 1-6 scale (1 = most positive). The earlier reported work of the authors (1975b) indicated that there were 12 areas of comparison made in the instrument. The 1978 report merely compares (by t-test) the differences in attitudes to biology before and after the program.

Findings

The authors report that the new curriculum, designed around the lower cognitive abilities as defined by Bloom (1966), was generally successful. Despite the provision that it was designed for "culturally deprived" students, the more advanced youngsters were purportedly challenged and interested in the subject matter. A comparison of the experimental and control classes' performances on the achievement test is provided.

The authors summarize the data as follows:

> The experimental group achieved the 80 percent level with respect to "essential objectives," meaning that the mastery level aimed at was attained, and approached this level on the "comprehension" and "basic knowledge" subtests. Only with respect to "application and other higher functions" was a mastery or near-mastery level not attained.

It is of interest to note that neither control group attained 85 percent mastery on any test, and no group attained mastery on the total test.

The authors analyzed the growth scores of the achievement test of those students with I.Q. scores of 90-99 and 110+. They summarized this effort as follows:

> While more than one-half of the experimental population had mastered 78 percent or more of the material, among the Emek Hefer control population, no one of this I.Q. group (90-99) had reached this level of mastery. The achievement of the experimental high-I.Q. group (110+) was also superior to that of the corresponding control group on the mastery sub-test. While about 92 percent of the experimental population demonstrated mastery, only 64 percent of the Emek Hefer control reached this level.

Regarding the Attitudes Towards Biology test, the authors report only that the average rank for biology decreased from 3.60 to 2.96 in the experimental group and from 3.45 to 3.29 for the control group. There was no statistically significant difference in attitude in the two populations on the pre-test. After the program a more favorable attitude $P < 0.05$) was expressed by the experimental population. All experimental I.Q. groups reportedly incurred significant positive changes. No data were shown.

The authors suggest that by dividing the program into basic core material and what was termed "enrichment-optional" activities, the students' time and efforts could be better directed. The 90+ I.Q. group could therefore readily master essential objectives. This approach purportedly allowed children to develop a more positive attitude towards biology and those individuals already interested were not dissuaded. Sabar and Kaplan conclude the report with:

> This study demonstrates that, in classes with heterogeneous populations, methods incorporating individualized instruction have an important influence on achievement when combined with conventional methods.

ABSTRACTOR'S ANALYSIS

It is unclear to this writer that new information or interpretations were developed in this paper that have not been previously reported. The authors' two previous papers were more timely and detailed; the reader had more information with which to work. The data presented in this third paper are in essence the same.

It is ironic that the authors report on a new "teaching-learning model" without any reference to the teachers. There is no indication of how the teachers involved in the study were selected. No data regarding their experience, training or cultural background is provided. There is no evidence that any analysis of teaching strategies or understanding of different methodologies was made. This is particularly troublesome in light of the fact that "The Model" being researched was/is primarily a teaching model and <u>not</u> a learning model.

In an earlier paper (1975b) the authors state:

> It should be emphasized that science teachers in Israel have been trained in the tradition of authoritarian teaching. One of the major objectives of the model was to re-orient the classroom toward the creation of a "child centered" atmosphere.

An inservice training course was given to the participating teachers. No mention is made as to what the course included, although an earlier paper (1975a) does suggest that considerable effort was devoted to bringing about changes in teacher personality so teachers would be more accepting.

The reader is therefore cautioned to regard the conclusions of the authors with healthy skepticism. Not only is the research over ten years old but also potentially important variables are ignored. The literature on the role of teacher attitudes in the instruction of science is extensive; it appears to be summarily dismissed.

As most persons who are familiar with mastery learning will agree, the usual level accepted for "mastery" is 90 percent. However, the authors defined mastery in this study as being attained if the student can answer 80 percent of the test items correctly. No real rationale is provided for establishing a lower level of performance. Of course, if the 90 percent cutoff had been used, the data would have shown the model as unsuccessful. Perhaps there is in this fact some reflection of the researchers' design rationale. Even with the lower "mastery level" the experimental group only attained it on the Mastery of Essential objectives sub-test.

In addition, to this less-than-rigid approach to determining success, the authors' also use phrases such as "approached this (mastery) level" and "near-mastery level" when referring to the experimental group. This would be analogous to stating that experimental results are nearly significant or for a husband to state that his wife was almost pregnant. There is no mid-point possible.

The authors suggest that the model was useful in helping students in the 90-99 I.Q. group attain mastery of the material. Comparison of experimental group individuals with students in the control group support this contention. The implication is that the teaching strategies employed in the model do allow for differential growth of students with differing abilities. Although we know nothing of the teachers assigned to the two groups, it is clear that a distinct performance difference did occur.

Student attitude towards biology in this study is treated in a less than systematic manner. The reader does not know what is occurring. Little additional information is available in the earlier report of the same research (1975b). One learns that the Attitudes Towards Biology test is a modification of an instrument developed to assess student attitudes towards mathematics. Reference is made to twelve areas or subscores within the test. The earlier report indicates that in three of the twelve areas, differences between the experimental and control groups were found. Yet, in the study reported in 1978, no reference is made to the categories, only to "the average rank for biology decreased..." for the experimental group. No data are provided. It is said that a t-test was run and that, "after the program the experimental population showed a significantly more favorable attitude than the control group." It is unclear to this abstractor what is being reported. The statement suggests that the test shows broad changes of attitudes while the earlier report clearly states differences in three of twelve areas. Furthermore, the earlier study reported that these data were subjected to analysis of variance, a powerful tool for reviewing important differences. None of the results of that analysis are made available in any of the papers authored by Sabar and Kaplan. Few of the data are provided for the reader's review.

While it is doubtful that the investigation reported in this paper answered the research problem "Does the new teaching-learning model significantly improve the affective and achievement performance functions in the disadvantaged segment of the heterogeneous class...?" the authors did demonstrate that students respond differentially. The earlier reports of the same investigation related more specifically to the model and provided greater information regarding the design.

As to whether the model maintained "...the performance levels and attitudes of the average and above-average I.Q. students," no conclusion is possible. The authors provided no data or discussion regarding the attitudes towards biology of these groups. The information concerning the achievement of mastery of these two groups indicates some variation but is insufficient to develop conclusions.

REFERENCES

Bloom, B. S. (ed.). Taxonomy of Educational Objectives: Handbook I, Cognitive Domain. New York, N.Y.: David McKay Co., 1966.

Kaplan, E. H. and N. Sabar. "A Model Biology Curriculum for Heterogeneous Seventh Grade Biology Classes Containing Culturally Deprived Students (in Israel). I. The Theoretical Foundations of the Model." Science Education, 59(3): 313-320, 1975a.

Kaplan, E. H. and N. Sabar. "The Construction of the Model." Science Education, 59(3): 321-332, 1975b.

Wilson, J., L. Cohen, and E. Begle (eds.). "Population Test Batteries, Part A." NALSMA Report No. 1. Stanford, Calif.: School Mathematics Study Group, 1965.

Cohen, Elaine L. and Mary Poppino. "Cloze Testing and Study Skills for College Biology." Science Education, 62(4): 443-447, 1978.
Descriptors--*Academic Achievement; Biology; *Cloze Procedure; Courses; Educational Research; Higher Education; *Language Ability; Reading Ability; *Reading Comprehension; Science Education; *Tests

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Joel J. Mintze , University of North Carolina at Wilmington.

Purpose

This study examined the usefulness of the Cloze testing procedure for predicting success in a college-level, introductory biology course, and the effects of a specially designed study skills class on achievement in and attitude toward biology. Specifically, the following questions were addressed:

(1) Is there a relationship between a student's score on the Cloze test and his grade in the biology class?

(2) Can the score on a Cloze test predict the student's success in a biology course?

(3) Do students who participate in both the biology class and the study skills class gain more in expressed confidence in biology skills than those enrolled in the biology class but not in the study skills class?

(4) Are grades earned by students enrolled in the biology class but not enrolled in the study skills class higher than grades of students enrolled in both courses?

Rationale

Although the authors do not expressly indicate it, this effort is part of a rather large and growing body of research on reading and study skills programs at the college level. Closely related studies (Foxe, 1966; Keetz, 1970; Sheldon, 1948; Weinstein and Gipple, 1974) have evaluated the effectiveness of study skills programs on science achievement, and integrated study/reading skills classes on success in college biology (Tomlinson et al., 1974; Tomlinson and Green, 1976). The underlying assumption of the present study is that achievement and confidence in college biology can be boosted through "better coordination of study skills support with content areas."

Research Design and Procedure

One hundred eighty-three (183) students who completed one of two introductory biology courses at Metropolitan State College (Denver, Colorado) served as subjects. Each student was administered a Cloze test and a locally-constructed Confidence Inventory during the first class meeting to ascertain levels of reading comprehension and anxiety in biology. During the second class meeting all students were advised of their Cloze test scores and were invited to enroll in "paired study skills classes." Those who did enroll formed the experimental treatment group, while nonenrollees served as controls. The study skills classes were taught by instructors in the Reading Department and provided students with practice in study-reading techniques using the assigned course textbooks and class lecture notes as primary sources. Students were instructed in the "SQ3R" study method; engaged in skimming and scanning exercises; learned to extract and analyze new vocabulary; practiced listening and notetaking skills using the "Cornell Model," and took biology quizzes similar to those used in the course to develop test-taking skills.

The Confidence Inventory was readiminstered in the biology sections during the last week of classes and data concerning final course grades were obtained from the appropriate biology instructors.

A variety of data analytic procedures was employed, including: Pearson's r (Question 1), Chi-square (Questions 2 and 3) and t-tests (Question 4). Data shrinkage resulted in small Ns for two of the analyses (Nexp. = 10 and Ncont. = 17 for question 3; Nexp. = 12 and Ncont. = 37 for Question 4).

Findings

Findings of the study can be summarized as follows:

(1) A positive relationship was found ($r = .31$, $p < .01$) between students' Cloze test scores and final course grades in biology.

(2) The Cloze test score was found to be a significant predictor ($\chi^2 = 22.31$, $p < .01$) of final course grades.

(3) Students who enrolled in the paired study skills classes demonstrated greater growth in confidence ($\chi^2 = 5.56$, $p < .05$) than control group students.

(4) No differences were found ($t = -.28$, $p < .05$) in final course grades among students enrolled in the experimental and control groups.

Interpretations

The authors conclude that the Cloze testing procedure is a "simple but useful diagnostic instrument which could be employed by biology teachers to identify early those students in their classes who would benefit by content-based study skills support." They go on to say that, "...students in this study who enrolled in paired study skills classes gained increased confidence in biology skills and succeeded in the biology course." They caution, however, that further work is needed to determine whether similar results would be found in other content areas.

ABSTRACTOR'S ANALYSIS

Relationship to Other Studies

Although the authors cite several studies concerning the development of the Cloze testing procedure and interpretation of Cloze scores, the reader is given no background about previous studies which employ this type of diagnostic instrument to predict achievement in the sciences. Has no work been done on the relationship between reading comprehension and achievement in college biology? If previous work has been done, how does this study add to existing knowledge? Why should science educators be concerned about reading comprehension, anyway? If one is concerned about predicting success in college biology, why use reading scores as predictors? (Perhaps other, more powerful, variables exist--SAT scores, IQ scores, scores on tests which assay relevant prior knowledge.) The reader is left wondering.

Similar questions can be raised about the usefulness of study skill classes to enhance achievement. Again the authors do not tell us how this effort fits into the matrix of previous work in the area. As a result, the nonspecialist is left wondering to what extent this study builds upon and adds to that fund of knowledge.

Contributions of the Study

The study does not seem to offer any real conceptual or methodological contributions to science education except perhaps to make us aware of the existence and potential value of the reading-study skills specialist. The fact that reading comprehension is related to achievement and that study skills classes may enhance student confidence should come as no surprise to those familiar with the literature (see Rationale section).

Research Design and Written Report

The authors readily acknowledge several design problems associated with the study, including: nonrandom assignment of subjects to treatment groups and small sample size. These are indeed serious problems which must be reconciled and compensated for. However, the authors failed us in this respect.

(1) We are given virtually no information about the subjects who formed the experimental anc control groups. When assignment to treatment groups is not random, the authors must establish whether the groups vary significantly with respect to important variables that may affect the outcome of the experiment. The present authors give us conflicting information. In stating the problem, they tell us that study-skills classes "attract high achievers striving for perfection." Yet in the discussion of findings we are told that "enrollment included both high and low achievers." Unfortunately, means and standard deviations were omitted for every variable under consideration. Consequently, the reader cannot determine whether

the study skills classes provided a compensatory vehicle (in which case no differences in final course grades is a positive finding) or an enrichment effort (in which case no differences in final course grades is a negative finding). This omission is the most significant shortcoming of the study because it prevents readers from properly interpreting the findings.

(2) We are not told how achievement was measured. Final course grades are a notoriously subjective measure of learning. In this study, the subjectivity was compounded--data from two separate courses (each presumably with idiosyncratic grading policies) were combined.

(3) No information concerning the reliability and validity of the data-gathering instruments was provided.

(4) The authors did not consistently adhere to standard format in the written report, making it difficult at times to read:

(a) no literature review
(b) no tables
(c) sample sizes provided in "Findings" section
(d) references listed but not cited in body of paper
(e) use of specialized jargon without definition (e.g., "Cornell method").

Further Research

If research on the improvement of learning and study skills is to be taken seriously, workers in this field must continue to design and carry out rigorous experimental investigations. The collection of easily available data and the execution of poorly controlled experiments will not provide answers to the important questions.

Further work might profitably focus on the identification of learners who need study skills assistance. However, these effects must be carried out within a rigorous experimental environment, using well-designed techniques and instruments, and preferably, growing out of and adding to a significant theoretical matrix. Interactions between study/reading methods and learner characteristics might also be a fruitful area for further research.

REFERENCES

Foxe, E. K. "An Experimental Investigation of a Brief Study Skills Program for College Chemistry Students." Unpublished Doctoral dissertation, University of Maryland, 1966.

Keetz, M. "An Experimental Investigation of the Effectiveness of a College Reading and Study Skills Course for Freshman Students Enrolled in Scientific Courses of Study." Paper read at International Reading Association Convention, Anaheim, Calif., 1970.

Sheldon, W. "An Evaluation of an Experimental Reading Program for Medical Students." Journal of Educational Psychology, 39: 298-303, 1948.

Tomlinson, B. M. and T. Green. "Integrating Adjunct Reading and Study Skills Classes with the Content Areas." In R. Sugimoto (Ed.), The Spirit of '76: Revolutionizing College Learning Skills. Proceedings of the 9th annual conference, Western College Reading Association, 1976.

Tomlinson, B. M., et al. "A Preliminary Report on a Project Integrating the Teaching of Reading and Studying in Science with Biology and ABC." Unpublished paper, University of California at Riverside, 1974.

COGNITIVE

VARIABLES

Treagust, David F. and Vincent N. Lunetta. "An Analysis of A Four-Category Concept in Science Using the Bower and Trabasso Model of Concept Identification." Journal of Research in Science Teaching, 15(5): 343-354, 1978.

Descriptors--College Students; *Concept Formation; *Educational Research; *Mathematical Models; Models; *Science Education; Scientific Concepts

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

Purpose

The purpose of this investigation was to test extensions of a mathematical model of concept learning proposed by Bower and Trabasso (Bower and Trabasso, 1964; Trabasso and Bower, 1964) to a task in science education. The authors hoped to find an application of that model to instruction in science education.

More specifically, Treagust and Lunetta investigated two hypotheses concerning the responses of subjects prior to the solution of a problem. First, if one were to divide in half all of the responses prior to the last error made in solving the problem, then it may be expected that the first and second halves of the responses would have the same frequency of errors. If this were the case, then stationarity would have occurred. The first hypothesis was stated as follows and tested stationarity:

H_1: The first and second halves of the sequence of correct and incorrect responses prior to the last error will have the same probability distribution outcomes.

The second presolution hypothesis also dealt with the sequence of responses prior to the last error. Taking two responses at a time, the possible sequences included an error following an error, a correct response following an error, an error following a correct response, and a correct response following a correct response. The second hypothesis was stated as follows:

H_2: The sequence of correct or incorrect responses prior to the last error will form an independent series of observations.

The final hypothesis concerned the solution of a four-category problem as the composite of two two-category subproblems. If the more difficult problem was in fact the whole of two less difficult subproblems, then the probability of the joint occurrence of the two subproblems resulting in the correct completion of the four-category problem would be the product of the probabilities of the correct solution to each of the subproblems. The hypothesis was stated as follows:

H_3: Empirical probabilities of correct response will be predicted for the four-category problem group (SP) by subproblem groups (S and P) using the multiplication rule.

Rationale

Bower and Trabasso proposed a mathematical model for the learning of binary concepts. They believe that an all-or-none learning model is appropriate for the two-category problem in which the subject makes a binary classification of a stimulus using one of two mutually exclusive responses. The subject first identifies a relevant stimulus dimension and then applies this dimension to the classification of some object (in Atkinson, 1964).

A four-category concept identification task involving a combination of two binary classifications may also be attempted by a subject. Trabasso and Bower (1964) proposed that the four-category problem is approached as two subproblems and that the subproblems are solved independently by the subject as all-or-nothing events. Treagust and Lunetta logically pursued the investigation of the Bower and Trabasso model by applying it to a specific content area.

Research Design and Procedure

The conceptual problem of the study was to correctly identify beech, catalpa, maple and sycamore trees when shown close-up, black and white photographic slides of the branch and leaves of the tree. Prior to the actual treatment, the subjects were given directions and were shown slides of the following five binary dimensions relevant to the classification of the leaves: 1) position of the leaf on the stem (opposite or alternate), 2) shape of the leaf (lobe or rounded), 3) venation of the leaf (prominent or not prominent), 4) shape of the leaf tip (pointed or rounded), and 5) base of the leaf (rounded or narrowed). The subjects were told that only two of these dimensions can be used to classify objects and were given with sample illustrations totally unrelated to trees.

The actual instruction consisted of 80 slides or leaves. The first 32 slides contained 8 slides for each of the four types of leaves in random order with the restriction that no two consecutive slides were of the same leaf. A second set of 32 slides arranged in a different sequence was then shown followed by the first 16 slides of the first set of slides. Following the presentation of each slide showing a branch and leaf for 10 seconds, the subject was shown a reply slide for 5 seconds at which time he or she marked a written response on the response sheet. A slide of the correct response was then shown for five seconds.

All of the subjects were randomly assigned to one of three treatment groups. Two of the three treatment groups recieved information regarding one of the dimensions needed to identify the leaves and had to discover the other relevant dimension. Members of group S were told that the position of leaves on the stem could be used in classification and that alternate leaved trees were sycamores and beeches while opposite leaved trees were maples and catalpas. The group S subjects had to discover that the shape of the leaves was important to classification.

Members of group P were told that the shape of the leaves was important in classification and that trees with rounded leaves were beeches and catalpas while maples and sycamores had lobed leaves. The group P subjects had to discover the characteristic of leaf placement on the stem. The third treatment group, SP, learned a four-category problem of which groups S and P were subproblems. The members of group SP were told only that two dimensions were necessary to classify all four trees and had to discover the relevant parameters of leaf shape and placement on the stem.

The subjects were 78 university students majoring in elementary or special education who were enrolled in a science course for nonscience majors. All of the subjects were volunteers with less than eight semester hours of science. The subjects were randomly assigned to one of the three experimental groups (S, P, and SP). All subjects were told that the purpose of the experiment was to determine how students learn to classify trees. Although 26 students were originally in each group some students were eliminated because they made no errors and therefore had no opportunity for learning and others were eliminated because they never met the criterion level of 16 successive correct responses. Therefore, the final sample size was 53 subjects with 16 subjects in groups S and P and 21 subjects in group SP.

The design may be diagrammed as shown below. It is an experimental design involving the random assignment of subjects to groups, no pretest, a different treatment for each group, and a subject response.

	X_S	0
R	X_P	0
	X_{SP}	0

Findings

In testing the first hypothesis, the trials of each subject prior to his or her last error were divided into a first and second half. If the subject was improving, then the number of errors in the first half would be greater than the number of errors in the second half. However, if no improvement was shown, the number of errors in each half would be equal. A paired t-test was used to determine if the error probability of the two halves was equal. No evidence was found to reject hypothesis 1 for groups S and P. However, the stationarity prediction could be rejected for group SP. Thus, stationarity exists for two-category problems but not for four-category problems.

The second hypothesis was tested using a Chi-square test for independence. The probability of success on one trial following success on the preceding trial should be equal to the probability of success on one trial following an error on the preceding trial. If the two probabilities were equal, then the probability of success on any trial was independent of the response on the preceding trial. For all three treatment groups, the observed probability showed that the learning process was not random and that the responses were not independent.

Hypothesis 3 proposed that the probability of correctly solving a four-category problem was the product of the probability of solving each two-category subproblem. The observed probabilities of correct responses were, in every case, greater than the calculated probability based upon the correct responses of each subproblem. Paired t-tests on the differences between the observed and predicted scores of group SP indicated differences beyond the 0.01 level of confidence. Thus, it was implied that the four-category problem was more than simply two independent two-category subproblems.

Interpretations

The strong tendency for stationarity of presolution responses in the two subproblems implied an all-or-none learning process of a binary concept. The more complex, four-category problem did not show stationarity. In addition, the four-category problem did not appear to be learned as two independent subproblems. Although the results were not completely consistent concerning the independence of presolution responses, it appeared that the subjects did not choose their responses randomly. Thus, the Bower and Trabasso model can be applied to the learning of science concepts which are of a binary nature.

ABSTRACTOR'S ANALYSIS

Treagust and Lunetta have applied the research findings of other disciplines to science education. Many prominent researchers have recommended that educational research utilize the methods and conclusions of related areas. The authors have applied the efforts of investigators in psychology and mathematics to science education, spanning the barriers of individual disciplines. Too much energy has been spent duplicating the efforts of others primarily through general ignorance. The model set by Treagust and Lunetta should be followed by more investigators. The review of the literature was well done by the authors and the relationships among the research were clearly shown.

The design, used by Treagust and Lunetta to test the theoretical model of concept identification, served the purpose well. No pretest was given to the subjects to assess entry behavior. However, students who had completed more than eight hours of biology, and who consequently may have known many trees, were eliminated from the study. All students who identified the trees with very little difficulty were also removed, under the assumption that not a great deal of new learning took place. Students who never solved the problem were eliminated since they presumably found the material extremely difficult or were disinterested. The above adjustments to the sample essentially removed the extreme cases and left the central cases. A random assignment to groups helped to insure homogeneity across groups.

The actual treatment was carefully thought out. The students were asked to learn the identification of four different trees given a picture of the leaf and its arrangement on a twig. The leaf arrangements are common characteristics used to identify trees. Most keys present the options of opposite (maple), alternate (beech and sycamore), and whorled (catalpa). However, lobed or rounded are not common, dichotomous characteristics. Lobed or not lobed would be a better description of this characteristic. Describing a beech tree as having rounded and alternately placed leaves is contrary to the general taxonomy of deciduous trees (Preston, 1976). It is always wise to verify all science content with a specialist in that area to avoid confusion. In this particular case, the description of a leaf, which appears toothed, as rounded may delay the learning of the concept.

Analysis of the results was conducted in an orderly and logical manner. Comments on the testing of each hypothesis follows. Hypothesis 1 addressed the probability distribution of the correct and incorrect responses in the first and second halves of the response sequence. Treagust and Lunetta, however, tested the difference between the mean number of errors in the first and second halves of the response sequence, not the error probability. An analysis similar to that used in Atkinson, Bower and Crothers (1965, pp. 40-45) which tested the differences in probabilities might be more appropriate for this hypothesis.

The testing of hypothesis 2, the sequence of correct and incorrect responses prior to the last error will form an independent series of observations, was clearly and correctly presented. Although the results of the analysis of the error probabilities of only treatment group P showed a lack of independence between groups, Treagust and Lunetta were correct in stating that their data did not support the Bower and Trabasso model of independence. The X^2 values would need to be much lower and the conditional probabilities much more equivalent as shown in earlier work by Bower and Trabasso (1964). Treagust and Lunetta offered many explanations for the nonequal presolution frequencies citing relevant research from the literature. One additional possibility may be that the task was not very difficult for the subjects since the success rate was so much higher than the error rate regardless of the previous response.

Hypothesis 3, empirical probabilities of correct responses will be predicted for group SP by subproblem groups S and P using the multiplication rule, was analyzed by comparing the observed mean probability of correct responses by group SP to the calculated probability obtained according to Equations 1 and 2 using the mean probabilities of correct response by group S and those of group P for each five-trial block. The calculated value for group SP, whether by Equation 1 or 2, apparently has no variance since it is a single number based on the product of means of all subjects in groups S and P for any given block of five responses. It was, therefore, inappropriate to use a t-test to determine the significance of differences between the observed and calculated, or expected, mean probabilities as shown in Table IV. Treagust and Lunetta did not give corresponding degrees of freedom to help decipher the origin of the t-value. Further confusion was added when the conclusion was drawn that "Paired t-tests on the differences between observed and predicted scores (using both Equations 1 and 2) of group SP for the 13 five-trial blocks indicated differences beyond the 0.01 level of significance" (Treagust and Lunetta, 1978) when block 13 and equal values (1.00) for the observed probability and the two calculated probabilities. A X^2 test would have been more appropriately applied to this situation (Siegel, 1956).

The conclusions drawn by Treagust and Lunetta were supported by their data and by the literature they cited. This investigation will contribute to concept learning theory and the validity of the mathematical model presented by Trabasso and Bower.

Treagust and Lunetta concluded their paper with a section entitled "Limitations and Implications." It is imperative that more researchers discuss the limitations of their studies since only the researchers themselves know the full limitations of a study. This type of discussion will help others understand the results of the study and will provide insight for those who wish to replicate the work The limitations presented by Treagust and Lunetta add clarity to the presentation of their study and temper their conclusions and implications. Although many potential problems exist with regard to the all-or-none concept identification model, the research conducted by Treagust and Lunetta significantly contributes to the understanding of the Trabasso and Bower model and its application to science education.

REFERENCES

Atkinson, R. C.; G. H. Bower; and E. J. Crothers. An Introduction to Mathematical Learning Theory. New York: Wiley, 1965.

Bower, G. H. and T. R. Trabasso, "Concept Identification." In R. C. Atkinson (Ed.), Studies in Mathematical Psychology. Stanford, CA: Stanford University Press, 1964.

Preston, J. North American Trees. Ames, IA: Iowa State University Press, 1976.

Trabasso, T. R. and G. H. Bower. "Component Learning in the Four-Category Problem." Journal of Mathematical Psychology, 1: 143-169, 1964.

Treagust, D. F. and V. N. Lunetta. "An Analysis of a Four-Category Concept in Science Using the Bower and Trabasso Model of Concept Identification." Journal of Research in Science Teaching, 15(5): 343-345, 1978.

Tamir, Pinchas and Vincent N. Lunetta. "Cognitive Preferences in Biology of a Group of Talented High School Students." Journal of Research in Science Teaching, 15(1): 59-64, 1978.

Descriptors--*Abstraction Levels; *Biology; *Cognitive Processes; Educational Research; Science Course Improvement Project; Science Education; Secondary Education; *Secondary School Science; *Talented Students

Expanded abstract and analysis prepared especially for I.S.E. by Claudia B. Douglass, Central Michigan University.

Purpose

The study conducted by Tamir and Lunetta examined the effects of secondary science curricula and related background variables on the cognitive preferences of talented science students. The curricula included the Biological Sciences Curriculum Study (BSCS) Green, Yellow, and Blue versions and a number of unspecified traditional texts grouped as Modern Biology. The related variables are those listed within the second purpose. The study centered around the following objectives:

1) to compare the results obtained by normative and ipsative procedures;

2) to identify the relationships between cognitive preference patterns and the background variables of gender, year in high school, general achievement, achievement in high school biology, nature of high school biology course, location of residence, hobbies, science reading, and prospective major field of study in college; and

3) to identify the degree of cognitive preference dependence on specific biological topics.

Rationale

The recent emphasis in the development of science curricula has been inquiry-oriented. Importance has been placed on concept formation, scientific principles, and the processes of scientific thinking. Implicit in these developmental activities is the assumption that the curriculum and teacher can influence the student's thought processes and cognitive style. Tamir and Lunetta investigated this assumption using an instrument which Tamir developed.

Research Design and Procedure

The subjects were 177 high school students who attended a secondary science training program (SSTP) at the University of Iowa during the summer of 1976.

The 40-item biology cognitive preference test (BCPT) was administered to the students one time. Each of the 40 items had four responses or "extension statements" of an original statement. Each extension represented one of the four modes of cognitive preference: recall (R), principles (P), questioning (Q), and application (A). The students were instructed to rank each of the four extension statements according to their own personal preference. They were told that there was no correct answer. This is a typical item (Tamir, 1975);

Penicillin was discovered by Alexander Fleming in 1928.

(A) a. Penicillin is extensively used by doctors in the treatment of bacterial infections.

(R) b. Penicillin is an antibiotic drug.

(P) c. Penicillin, like other antibiotics, is produced by a living microorganism and is effective in the control of other microorganisms.

(Q) d. Penicillin is not an effective treatment of all bacterial diseases and a whole range of other compounds has been developed to treat bacterial diseases.

The student's overall cognitive preference pattern was represented by his total score in each of the four cognitive preference areas. Two forms of the test were administered. One form was arranged so that the first 20 items required a ranking of the responses on a four-point scale (normative). The second form had 20 normative items first followed by 20 ipsative items.

Findings

The findings for the study were the following:

1) Mean scores for each of six subtests (R, P, Q, A, Q-R, and P-A) were determined for the normative and ipsative items. Statistically significant differences were found between the normative and ipsative items in the areas of recall, application, and principle-application. The intercorrelations between the ipsative and normative items were all positive and, with the exception of the application items, statistically significant. No statistically significant differences were found between form A and form B.

2) Students following the BSCS Blue and Yellow versions were found to have higher questioning and lower recall cognitive preference patterns than students following more traditional curricula.

3) Students who received a grade of "A" in high school biology showed a higher preference for principles than students who received a grade of "B."

4) Non-science majors had the lowest preference for recall while premedical students showed a generally higher preference for principles and questioning.

5) Students in the higher grades showed a greater preference for principles and a lower preference for recall.

6) When the students were grouped according to the frequency with which they read the scientific literature, those who were non-readers and those who were frequent readers formed one homogeneous

group while all other students formed a middle group with different cognitive preference characteristics. Those in the middle group had a lower preference for applications and principles than did the other two groups.

7) A few isolated topics of interest were identified and related to cognitive preference.

8) Students showed higher questioning and application scores for items relating to form and structure while for items relating to process and function, their scores were higher in the areas of recall and principles.

9) This sample of students, when compared to Israeli students, was exceptionally high in their preference for critical questioning and exceptionally low in their preference for recall.

Implications

The finding that high ability science students have a high cognitive preference for recall indicated that these students exhibited a high level of curiosity and a desire to learn more. These data have implications for biology teaching when questioning or recall preferences are valued.

ABSTRACTOR'S ANALYSIS

The rationale for the study conducted by Tamir and Lunetta was based upon the assumption that a goal of modern science education is to encourage and develop critical thinking in students. The development of recent science curricula would certainly support this assumption. Tamir and Lunetta then proceeded after the model of Heath (1964) and turned their attention to the cognitive preference characteristics of the students rather than their subject matter achievement. An evaluation based upon the thought processes of the students indeed seems most appropriate for the inquiry-oriented curricula. As Tamir and Lunetta presented the rationale for their study, they used interchangeably the terms "cognitive preference" and "cognitive style." Usually, the cognitive style of the student is considered to be a rather stable personality characteristic such as reflexivity-impulsivity or field-dependence-independence (Kagan 1966; Witkin et al., 1967). Cognitive preferences on the other hand are more flexible and refer to a preferential mode of learning (Nunny, 1975) or of extending an idea or concept (Heath, 1964). This was a minor problem and the differentiation between the concepts of cognitive style and cognitive preference may not yet be clear. This research is a natural extension of Tamir's other work.

The presentation of the research would have been improved if the independent variables had been operationally defined. Specifically, a definition of general achievement was lacking. Also, the definition of an inquiry-oriented curriculum or one emphasizing "concept formation, scientific principles, and the processes of scientific thinking" was lacking. BSCS texts are assumed to meet these criteria, but according to what criteria were curricula placed into the category of "Modern Biology" and how can the differences between the Green and Blue versions be explained? In the introduction to their paper and in other papers (Tamir, 1975), Tamir and Lunetta acknowledge the effect of teacher attitude. However, it is omitted from the current research design. BSCS texts were written with the intention of meeting the above criteria emphasizing questioning skills; however, without endorsement from the classroom teacher these goals may never be met and a BSCS text may be taught in the most traditional, dictatorial manner. The remaining variables were clearly presented and well defined.

The biological cognitive preference test (BCPT) was developed and validated in a previous study (Tamir, 1975). The four response areas of recall, principle, questioning, and application were well presented. However, the combined areas of questioning-recall and principle-application were presented with very little explanation. In the discussion of the results, they were referred to as an indication of curiosity (questioning-recall) and general interest in applications (principle-application). These generalizations arose as a synthesis of the work of others (Kempa and Dube, 1973) and earlier work by Tamir

(1975). However, they were not adequately justified in this paper and conclusions were not drawn regarding their discrete or unique identity from the other response areas. A more complete explanation of the questioning-recall and principle-application areas is required to differentiate the curiosity aspect from questioning alone and the interest in applications from the area of application alone.

Table I presented the data regarding the ipsative and normative portions of the BCPT. Although the last column was not labeled, one gathered from the discussion of the paper that these were the t-values comparing the ipsative and normative aspects of the test. Although it was never overtly stated, a general case for the equivalence of the normative and ipsative response methods was made. Tamir and Lunetta minimized the significant differences between the two procedures. However, later in their paper, differences of this size and smaller were regarded with importance. The inconsistent treatment of the statistical analyses was somewhat apparent and should have been avoided. The intercorrelations showed a range of overlap between the ipsative and normative procedures up to 26 percent. In all cases except one, Form A, presenting the ipsative items first, was more reliable. It appeared that Table I could have been interpreted more completely and that an expanded discussion of these data was warranted since the remainder of the paper depended upon this instrument and the treatment of the resultant scores.

As one studied Table II, it was interesting to note that the cognitive preference characteristics of the BSCS Green Version students resembled more those of the Modern Biology students in all areas than the Yellow and Blue Version BSCS students. It is unfortunate that the sample size could not have been larger, permitting analysis of this observation. Looking at the results of the biology course grade analysis, one wonders why all students receiving a grade less than B (N = 20) could not have been analyzed as a group. Perhaps that would be too heterogeneous a grouping to be of any value. The analysis of the prospective field of study should have had a multiple range test, such as a Scheffé or Newman-Kuels following the analysis of variance rather than a series of t-tests (Scheffé, 1959). In spite of these problems with the data contained in Table II, the results were extremely well presented.

The discussion of the relationship between the biological topics of the BCPT and the student's cognitive preference was sketchy. A complete listing of the topics was not included nor was a rationale for the analysis given. The results of the relationship between the emphasis on structure vs. function and cognitive preference were somewhat confusing. It was unclear how a low score in the cognitive preference area of questioning and a high score in the area of recall could result in the conclusion that students were more curious about the processes and functions of biology resulting in a high combined questioning-recall score. The reader would be aided by a more complete discussion of the analyses of, and conclusions from, these data.

The sample for this study was a select and fairly homogeneous group of students. The generalizability of the study was therefore restricted. The authors acknowledged this drawback and tried to compare this group of SSTP students to Israeli students. There was no rationale for the cross-cultural comparison and, although it may have been of value, it should have been done with a more complete description of the Israeli students.

In general, the study was well planned and conducted. The results were carefully analyzed and clearly presented. A strength of the study was the attention given to sample size and the authors' concern for the power of the statistical tests. In this regard, this study could be a model for others.

Future research in this area might include an investigation of the effects of various curricula on student cognitive preferences. This research could continue the cross-cultural comparisons of Tamir and Lunetta with broad samples of students. Similarities and differences in cognitive preferences may also be explored across disciplines and longitudinally throughout a student's maturity. The specific function of cognitive preference assessment needs to be evaluated in a practical way.

REFERENCES

Heath, R. W. "Curriculum, Cognition, and Educational Achievement." Educational and Psychological Measurement, 24(2): 239-253, 1964.

Kagan, J. "Reflection-Impulsivity: The Generality and Dynamics of Conceptual Tempo." Journal of Abnormal Psychology, 71: 17-24, 1966.

Kempa, R. F. and G. E. Dube. "Cognitive Preference Orientation in Students of Chemistry." British Journal of Educational Psychology, 43: 279-288, 1973.

Nunney, D. N. Educational Cognitive Style: A Basis for Personalizing Instruction. Unpublished manuscript. Oakland Community College, 1975.

Scheffé, H. The Analysis of Variance. New York: Wiley, 1959.

Tamir, P. "The Relationship Among Cognitive Preference, School Environment, Teachers' Curricular Bias, Curriculum, and Subject Matter." American Educational Research Journal, 12(3): 235-264, 1975.

Witkin, H. A.; D. R. Goodenough; and S. A. Karp. "Stability of Cognitive Style from Childhood to Young Adulthood." Journal of Personality and Social Psychology, 7: 291-300, 1967.

54

Marjoribanks, Kevin. "The Relation Between Students' Convergent and Divergent Abilities, Their Academic Performance, and School-Related Affective Characteristics." Journal of Research in Science Teaching, 15(3): 197-207, 1978.

Descriptors--*Achievement; *Attitudes; *Convergent Thinking; *Divergent Thinking; Educational Research; Elementary Education; *Elementary School Science; *Junior High School Students; Predictor Variables; Science Education

Expanded abstract and analysis prepared especially for I.S.E. by Rosalie Grant and John W. Renner, University of Oklahoma.

Purpose

The primary purpose of this research was to examine the academic performance and school-related affective characteristics of 12-year-old English students who had been classified as convergers, divergers or all-rounders. Convergers are defined as those who excel on conventional intelligence tests but do not do well on open-ended tests (here identified as tests of creativity). Divergers are the converse of convergers, and all-rounders are those children who do equally well or poorly on both kinds of tests. The research tested these specific hypotheses:

1. the mathematics and physical science scores of convergers are higher than those of divergers while divergers perform better than convergers in English and French;
2. all-rounders, who have both high convergent and divergent abilities, perform as well as convergers in mathematics and physical science, and perform as well as divergers on measures of English and French;
3. there are no differences between convergers, divergers, and high-scoring all-rounders in performance in biology; and
4. convergers and high-scoring all-rounders have more positive school-related attitudes and higher self-adjustment in school scores, than do divergers and low-scoring all-rounders.

Rationale

The research addresses the concept that children do not approach school problems at the absolute limit of their cognitive abilities but are greatly influenced by being a converger, a diverger or an all-rounder. The researcher proposes that convergers tend to perform well in the physical sciences and mathematics. Divergers, according to the researcher's assumption, perform well in English and modern languages and are attracted to the arts. Biology and general arts courses are assumed to attract both convergers and divergers in about equal proportions. The researcher also suggests that convergers are conscientious and have conforming and conservative attitudes while divergers are likely to be more rebellious and intellectually independent and have more liberal and nonauthoritarian attitudes. Because prior research has produced inconclusive evidence regarding gender differences in cognitive and affective development of children, the hypotheses were examined separately for boys and girls. The research was suggested by research published in 1966 and 1968 by L. Hudson which was done with English schoolchildren. That work was published in London under the titles Contrary Imagination and Frames of Mind by Methuen. The research is also related to many studies which have investigated the relationship between convergent and divergent abilities, and the academic achievement and affective characteristics of children. Those studies, according to Marjoribanks, relied on the use of restricted statistical techniques. Therefore Marjoribanks analyzed his data using regression-surfaces, generated from a multiple-regression model.

Research Design and Procedures

Four variables were measured; those variables are convergent ability, divergent ability, academic achievement, and school-related affective characteristics. The verbal reasoning test (Verbal Test EF) and a nonverbal-reasoning test (Spatial Test 3) from the National Foundation for Educational Research in England (N.F.E.R.) were used to measure *convergent thinking*. *Divergent ability* was measured using material developed by Torrance and previously used with English schoolchildren. The adopted Torrance material was scored for fluency, flexibility and originality just as is done in Torrance's Minnesota Tests of Creative Thinking. Convergent and divergent abilities were measured at the beginning of the year. *School-related affective characteristics* were measured during the year using a questionnaire developed by the N.F.E.R. personnel who provided the reliability and validity data. The factor analysis done by the N.F.E.R. generated two factors which were labeled school-related attitudes and self-adjustment in school; the theta reliabilities of those factors are 0.85 and 0.82, respectively. *Academic achievement* was measured at the end of the school year using mathematics and English tests designed and standardized by the N.F.E.R. Teachers from all the schools involved designed tests to measure achievement in French, physical science and biological science. Reliability estimates for those tests were 0.94, 0.94, and 0.93, respectively.

The research was performed in 1972 in four junior high schools of an English provincial town. A total of 450 12-year-old children (210 males and 219 females), who were the new students being taken into the schools, comprised the sample. The children had been randomly assigned to the schools and studied a common curriculum designed by teams of teachers from the four schools.

In summary, the data for the research were gathered throughout the course of one year. Convergent and divergent abilities were measured at the beginning of the year, school-related affective characteristics were measured during the year, and academic achievement was measured at the end of the year.

Regression surface analysis which examined possible linear, curvilinear and interaction relations among the variables was used to test the hypotheses. In addition, the analysis-of-zero-order-correlations technique was employed. In employing regression surface analysis a simple random sample was not used. The Jackknife technique was used to estimate the design effect for each raw regression weight. The Jackknife estimates were based upon a subsample taken from each participating school. A second use of the regression analysis and application of the Jackknife technique resulted in the deletion of those variables that no longer had significant relations with the variables of academic achievement and school-related affective characteristics.

Findings

The zero-order correlations procedure allowed the researcher to assess the current validity of the measures used. The results demonstrated that convergent and divergent abilities measures have stronger relation with academic achievement measures than with affective characteristics. Generally, the measures of convergent and divergent abilities have moderate to high concurrent validities in relation to the achievement and affective scores.

The regression surfaces generated indicate that achievement and affective scores generally have only linear relations with the convergent and divergent abilities. There were some situations, however, in which the associations are more complex.

The regression surfaces also demonstrated some findings with respect to gender. Girls who are convergers (high convergent ability-low divergent ability) scored higher on the physical science test than did the divergers. The highest physical science scores were obtained by girls who were high all-rounders. For boys, the analysis reflected a significant interaction between convergent and divergent abilities. The highest fitted scores in physical science are associated with boys who are convergers; the all-rounders and divergers have similar scores. The academic scores in English demonstrate that for both genders, the performance of convergers is higher than that of the divergers. Girls who are convergers have more positive school-related attitudes than do divergers. Girls who are high-scoring all-rounders have very positive school-related attitudes. Similar results were found for the relations between girls' abilities and self-adjustment-in-school scores. Boys present a quite different profile. At each level convergent ability, increments in divergent ability are associated with increases in school attitude scores until a threshold of divergent ability is reached. From that point, further increments in divergent ability are associated with decreases in the attitude scores.

Interpretations

In general, the hypotheses relating bias in abilities to academic achievement and affective characteristics are only partially supported. While it was found that convergers performed better than divergers on the mathematics and physical science tests, which provides support for the acceptance of the first hypothesis, it was found also that convergers performed better than divergers in English and French. For most subjects, high all-rounders performed as well as or better than convergers and generally much better than divergers, which only provides partial support for the acceptance of the second hypothesis. The third hypothesis, which suggests that there are no differences between convergers, divergers, and high-scoring all-rounders in biological science, was not supported. Instead, high all-rounders performed better than convergers, who in turn had higher biology scores than divergers. For girls, high-scoring all-rounders and convergers tended to have more positive school-related attitudes and higher self-adjustment in school scores than did divergers and low-scoring all-rounders, which provided support for the acceptance of the final hypothesis. But for boys the relations with affective characteristics were more complex. For example, boys who were convergers (assessed by nonverbal ability) and who were academically successful exhibited negative school-related attitudes and a low level of self-adjustment in schools. Thus, the final hypothesis was only partially supported.

Therefore, while convergers appear to be extremely successful in physical science and mathematics, their pattern of abilities does not prevent them from succeeding in arts subjects. Alternatively, divergers tended not to succeed in arts subjects nor in the physical science area.

Thus, the general theoretical position that convergers and divergers display differential academic success was not supported in the present examination of 12-year-old children.

ABSTRACTOR'S ANALYSIS

The value of this research lies, we believe, in the ability of the instruments to measure what they purport to measure. The investigator reports moderate to high validity for the instruments used and provides quantitative data to support those claims. The results of the study perhaps reflect the "moderate to high" validities of the instruments: the results seem very mixed. Perhaps instruments other than teacher-made tests and conventional intelligence tests could help separate the results from such a study into cleaner categories.

The statistical techniques used in the research are complex and difficult to understand. If the researcher had more fully explained and justified regression surface analysis, the reader who is unfamiliar with that statistical methodology would now fully appreciate the study. Certainly the author stated that regression surfaces based upon a particular multiple-regression model would enable a more complete analysis of the relationships among the variables. But, the author doesn't explain why regression surface analysis was better than other analyses based on multiple-regression models. Perhaps the explanation is simple if you understand the statistical methodology used in the study, but an explanation would have been helpful to the reader who is unfamiliar with the statistics. Because regression surface analysis is a statistical procedure which is not commonly available, the author should have cited some commonly accepted references.

Furthermore, because the author doesn't justify the particular statistical procedures used in the research, the reader is left wondering whether there are any limitations on the procedures, and consequently on the interpretations of the results. For example, could a cubic, rather than a quadratic, equation have been used for the multiple-regression model? Would a cubic equation be expected to reveal more relationships among the variables than a quadratic equation? If so, why was the model based on a quadratic equation? Perhaps the author tried to fit the data with polynomial expressions up to, say, the fourth power but found that the quadratic was the most parsimonious, according to some criteria. If so, numbering such information would be relevant and useful to the reader.

The statistical procedures appear to have been used more for descriptive purposes than for making inferences. Although hypotheses were stated, no statistical tests of significance were reported. The regression weights, regression-fitted scores and means appear to have been used as descriptive, not inferential, statistics. Perhaps regression surface analysis is a technique which is not readily used for hypothesis testing. The technique may, possibly, be more appropriate for describing the best-fitting surface for a set of data. If the author had either cited a well-known reference for the statistical procedures or more clearly described the analysis, the reader who is unfamiliar with the statistical analysis could clarify the aforementioned ambiguity. Without an appreciation of the statistics used in the research, the reader could wonder why the design of the study demanded such complex statistical techniques be employed.

The principal value of the research, in our opinion, is the recognition of the importance of convergent and divergent thinking to success in school. Today's schools do demand mostly conformity and convergent thinking. This research demonstrates that those traits can be validly measured and that they have an impact upon student achievement, attitude, and affective characteristics. There is need, we believe, for further study of the three classes of thinkers identified here to ascertain how such thought types influence the learning that occurs in science. Consider that convergers are most successful in physical science and mathematics and also successful in the arts. Is that unexpected finding due to the disciplines themselves, the manner in which they were taught, or the nature of convergent thinking? Perhaps even the most divergent thinking--such as is believed to be useful in art--has an element of convergence in it and the instruments used in the research detected it. In our opinion, this research has isolated an important area and demonstrates that further study is needed.

Pringle, Rhodell G. and Ashley G. Morgan. "The Effects of Laboratory-Oriented Experience in SCIS on Stability of Cognitive Style of Teachers." Journal of Research in Science Teaching 15(1): 47-51, 1978.

Descriptors--*Cognitive Processes; *Cognitive Tests; Educational Research; Laboratory Experiments; *Science Course Improvement Project; Science Education; *Teacher Education; Thought Processes.

Expanded abstract and analysis prepared especially for I.S.E. by Avi Hofstein, The Weizmann Institute of Science.

Purpose

This paper reports a study that examined the question whether, and by which method, the cognitive style of a group of graduate students could be modified.

In particular, the study attempted to examine the influence of laboratory-oriented experiences in the Science Curriculum Improvement Study (SCIS) program on the stability of cognitive style (field dependence versus field independence) of teachers.

Rationale

The researchers claim that research studies concerning the cognitive styles of teachers are scarce and that very little is known about the impact of curricular experiences on cognitive styles. Cognitive style is defined by the researchers as the "consequence of the interaction between the learner's constitutional characteristic (inheritance) and environmental experiences."

The authors cited researchers who have categorized cognitive style along a biopolar continuum with poles labeled field dependent--field independent, analytic--non analytic, conceptual--perceptual motor, inferential--categorical, and strong--weak automatization (Broverman 1960; Gardner, 1959; Kagan, et al., 1963; Witkin, et al., 1962).

They suggested that the Science Curriculum Improvement Study (SCIS) program seeks to enhance the cognitive style continuum: field dependent versus field independent. The authors also suggest that getting insight into the possible influence of SCIS on cognitive style could provide information about modifying cognitive (perceptual and intellectual) functioning.

Research Design and Procedures

The study involved four groups ($N \sim 20$) of graduate students enrolled in the School of Education of Georgia State University during the summer quarter of 1975.

Two groups taught by the same instructor were enrolled in four to five-week laboratory-oriented activities in SCIS courses. The two control groups enrolled in a non-SCIS course in which they were engaged in group activities (problem solving) involving curricular models.

The Group Embedded Figures Test (G.E.F.T.) was administered to students in the following treatment groups:

SCIS Groups (Treatment B)	No SCIS Groups (Treatment A)
Group 1B: pre/post testing	Group 1A: pre/post testing
Group 2B: post-testing	Group 2A: post-testing

The G.E.F.T. is a test designed to assess students' competency to locate a single form when it is hidden within a complex pattern.

Findings

The results obtained from one-way analysis of variance conducted on the mean gain pre/post score on the G.E.F.T. test showed a significant ($P < 0.05$) gain in cognitive styles for both the SCIS and the non-SCIS groups.

No significant differences were obtained on comparison of the two instructional treatments (two-way analysis of variance).

Interpretations

On the bases of the research findings the authors suggested that stability of cognitive style could be modified. However, they could not imply that the SCIS treatment is a better or worse method to modify this stability.

They concluded that, via the two courses taught, students "experienced something, probably the problem-solving tasks which appeared to modify cognitive style."

ABSTRACTOR'S ANALYSIS

This research study appears to have been well-designed in terms of treatment and control. However, it is very difficult to arrive at generalizable conclusions since the results were obtained from a small ($N < 20$) groups of subjects. By using such small treatment and control groups the researchers introduced vagueness to their research design and hence, to their educational interpretation.

The research report is also vague in the review of the literature. New approaches and techniques to measure cognitive style like "cognitive preference" (Heath, 1964; Tamir, 1975; Ben-Zvi, et al., 1979) were not included or mentioned in the review of the literature. Although the authors claimed that, "While there exist extensive literature regarding cognitive styles. . .", no reference or examples were given.

It would have been preferable to include in the review of the literature other methods used to measure cognitive style that would help the reader to understand why the investigators decided to choose a certain measuring instrument. Another point worthwhile mentioning is the need to give the reader sample items from the measuring instrument so that it will be clearer why and how the instrument is measuring cognitive style. It is not clear why the G.E.F.T. was chosen and why the G.E.F.T. is a valid measure to answer the research questions the investigators raised. The article is also vague in the way the authors presented both the methodology and the research results. The following is a quote from the methodology:

> The ECI 740-1 and ECI 848-1 groups were administered the Group Embedded Figures Test (G.E.F.T.) during the first week of classes. The ECI 740-1 and ECI 740-2 groups experienced laboratory-oriented activities in SCIS according to the following schedule:
>
> ECI740-2 4.5 weeks, 2 days/week (M&Th), 2 hours/day (2:00-4:00 p.m.) = 18 hours
>
> ECI740-2 4.5 weeks, 2 days/week (M&Th), 2 hours/day (9:00-11:00 a.m.) = 18 hours
>
> The ECI 740 groups were taught by the same professor. The ECI 848 groups were taught by different professors. All of the intact groups were posttested after 4.5 weeks for measures of cognitive style using the G.E.F.T.
>
> An equal number (N = 20) of subjects were randomly selected from the groups with N's greater than 20. The following represents the initial distribution: ECI 740-1 = 20 Ss, ECI 740-2 = 23 Ss, ECI 848-1 = 20 Ss, and ECI 848-2 = 26 Ss.

It is not clear why the authors have to torture the readers with all the details of course numbers in their university. It is suggested that in the future, researchers should try to simplify the description of methodology in order to help the reader to understand both the methodology and the research results.

In the conclusion to their report the investigators suggested that through enrollment both in SCIS and the non-SCIS programs the students experienced "something, probably the problem-solving task which appeared to modify cognitive style."

Problem solving in the context of science learning appears to be a complicated, and thus needed, research area (Getzels, 1964; Larkin and Reif, 1979). Therefore, it seems that in their conclusion the authors should elaborate and expand why they think, and what evidence they have got, to suggest that students have developed such skills.

More content analysis and examples concerning the two instructional methods could help in understanding how and why problem-solving abilities could have been developed.

Since this research study (according to the authors) is only exploratory in nature and in which the results were obtained on the basis of <u>one</u> research instrument (measure), it is very difficult to use the results and to employ them either to classroom instruction or to teachers training.

Again, it seems that in their concluding remarks the authors should have suggested how to elaborate the research in order to obtain generalizable findings. It is very difficult to convince the reader that this research study provided more insight to the area of cognitive style in general and to the understanding of the question of how students solve problems in science in particular.

In conclusion, it seems that, in the future, research reports should include a more organized and clear description of the methodology and results, a more clear analysis of the instructional method, and detailed description of research instruments. And, conclusions should be stated in a clear and scholarly manner.

REFERENCES

Ben-Zvi, R., A. Hofstein, Y. Solomon, and D. Samuel. "Cognitive Preferences and Modes of Instruction in High School Chemistry." Journal of Research in Science Teaching, 16: 569-574, 1979.

Broverman, D. M. "Cognitive Styles and Intra-Individual Variation in Abilities." Journal of Personality, 28: 240-256, 1960.

Gardner, R. W., P. S. Holzman, G. S. Klein, H. B. Linton, and D. P. Spence. "Cognitive Control: A Study of Individual Consistencies in Cognitive Behavior." Psychological Issues, 1(4), 1959.

Getzels, J. W. "Creative Thinking, Problem Solving and Instruction." In E. R. Hilgard (Ed.), Theories of Learning and Instruction. The 63rd Yearbook, National Society for the Study of Education. Chicago: The University of Chicago Press, 1964.

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

Kagan, J., B. Rosman, D. Day, J. Albert, and W. Philips. "Information Processing in the Child: Significance of Analytic and Reflective Attitudes." Psychological Monographs, General and Applied, 78(1): (Whole No. 578); 1-37, 1964.

Larkin, J. H. and F. Reif. "Understanding and Teaching Problem Solving in Physics." European Journal of Science Education, 1: 191-203, 1979.

Tamir, P. "The Relationship Among Cognitive Preference, School Environment, Teachers' Curriculum Bias, Curriculum and Subject Matter." American Educational Research Journal, 12: 235-264, 1975.

Witkin, H. A., D. R. Goodenough, and S. A. Karp. "Stability of Cognitive Style from Childhood to Young Adulthood." Journal of Personality and Social Psych 'ogy, 7: 291-300, 1967.

MISCELLANEOUS

Gardner, Robert E. "Women in Engineering: The Impact of Attitudinal Differences on Educational Institutions." Engineering Education, 67(3): 233-240, 1976.

Descriptors--Attitudes; *Educational Research; Engineering; *Engineering Education; *Females; *Higher Education; *Student Attitudes; *Sex Differences

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Walter S. Smith, University of Kansas.

Purpose

The main purpose of this study was to ascertain whether women students, who had been entering engineering in increasing numbers, might tax the institution's resources differently than a population of nearly all men would. A second purpose, logically prior to the first purpose, was to find out whether women engineering students differ from men in their choice of majors within engineering and academic performance, and whether the institution affected the women's attitudes toward their education and job selection.

Rationale

Based on previous studies which had indicated that the attitudes of women engineering students toward their education was different than men's attitudes, this study grew out of a concern that increasing numbers of women students might affect institutions' use of resources. Specifically, if women continued to constitute an increasingly larger proportion of engineering students and if their pattern of specialization differed significantly from that of men, then the overall distribution of students to specialties would change. In turn, institutions would need to commit more faculty to some areas and less faculty to others. However, institutions may be unable to readjust their faculty distribution for a variety of reasons (e.g., tight budgets preclude new hiring; tenure rules preclude cutting staff).

Research Design and Procedure

Part I on the College Student Questionnaire (CSQ) and a questionnaire developed locally were administered to women and men in the study at the start of the freshman year; and the CSQ Part II was given at the end of that year to women only. Items were selected from the CSQ to measure attitude toward education and job selection. Grade point average and attrition information were collected for all subjects at the end of each of the first four college semesters to measure academic performance. Major selection was recorded for all subjects at the start of college and at the end of the fourth college semester.

Grade point average, attrition, and major selection of women and men were compared for the four times where data were collected. CSQ data for women from the start and end of the freshman year were compared to ascertain the impact of the institution on selected attitudes. (CSQ data for men was not reported.)

The population was composed of women freshman engineering enrollees (N = 73) at Cornell University in 1974-75 and a random sample of men freshman enrollees (total N = 575; number selected for the study was unspecified). Of the 68 women who responded to CSQ Part I, 9 were from ethnic minorities and were dropped from the study as being atypical; 36 of the 59 nonminority women responded to CSQ Part II. Eighty-one responded to CSQ Part I. Grade point average, major selection, and attrition data were collected for the 81 men and 59 women who responded to the initial survey.

Findings

(a) Academic performance. Mean grade point averages for the women and men were reported for each of four semesters and for the entire two years. No statistics beyond mean g.p.a. were reported; but a difference of 0.14 in the two-year overall grade point averages on a 4.30 grading scale (men's g.p.a. = 2.75 and women's g.p.a. = 2.61) led the author to conclude that these two groups performed academically at approximately the same level.

Attrition over the two years was nearly identical for women (26.5 percent) and men (25 percent). However, attrition due to academic failure did differ for the two groups (six men and one woman were dropped for academic reasons over the two years of the study).

(b) Choice of major. These data were reported as the percentage of men and women who chose each of nine engineering majors at the end of the second year.

Greatest differences between women and men occurred in civil and environmental engineering (chosen by 26 percent of women and 14 percent of men) and in electrical engineering (chosen by 14 percent of women and 28 percent of men). Although no statistical tests were reported, the author concluded that the pattern of choices of major was different for women and men.

Changes in major between the start of college and the end of the second year were also analyzed. Based on data indicating 48.8 percent of men and 45.3 percent of women changed majors, the author concluded that women and men were not different in their astuteness in preparation for choice of major prior to the freshman year.

Data were included to indicate the percentage of women and men changing to each of the nine major areas. For example, 23 percent of the men who changed fields switched to electrical engineering whereas only 10 percent of the women who changed fields selected electrical engineering as their new major. Although no statistics beyond percentage were reported, the author concluded that the pattern of changes of majors differed for women and men.

(c) Attitude. Twelve CSQ items were included on both Parts I and II; and of thse 12 items, 7 items changed significantly between the two test administrations. Although no statistical analysis was described, the author concluded that the attitudes of women engineering students changed during their freshman year and became more like the attitudes of men engineering students and people practicing the profession.

The author concluded that over the course of the freshman year women students became far less interested in being helpful to others and society as an anticipated source of job satisfaction and more interested in above average income, being creative and original, and working with people. Also, these women students differed between what they had hoped for and what they actually found to be their greatest personal satisfaction in college. A majority (61 percent of those responding to the CSQ Part II) indicated they had hoped for academics as their greatest satisfaction, but only 19 percent had found their greatest satisfaction from coursework.

Interpretation

Despite differences in attitudes (found in previous studies), women engineering students differ little from men in behaviors (i.e., academic performance and choice of major) that might have adverse effects on institutional resources. At least, any differences were small enough to produce negligible effects. Moreover, some important attitudes of these women students were changed in college to become more like attitudes of men. The author concluded that even if women made up as much as 50 percent of the engineering students, their presence would have little impact in the distribution of faculty resources except possibly in chemical engineering and mechanical and aerospace engineering.

In interpreting the results of this study, the author frequently employed a cause-effect model which assumed that attitudinal differences ought to cause behavioral differences. For instance, when discussing the apparent lack of difference in academic performance, the author stated that the two groups performed at the same level "despite differences in attitudes toward academic achievement" (p. 234).

ABSTRACTOR'S ANALYSIS

This study of women engineering students is timely and of potential benefit both to institutions and their students. However, the time lapse between the study (1974-75) and the present requires a note of caution in applying the study's findings. The finding that academic performance of these women engineering students was similar to that of men contrasts with other studies cited by the author and therefore provides important evidence for those evaluating women's increasing numbers in engineering.

The analysis and interpretation of the data could have been strengthened in several areas; and because of these weaknesses, either in the analysis or in the reporting of the analysis, the study's conclusions must be regarded cautiously. As was pointed out in the preceding section on findings, almost no analysis other than "eyeballing" the data was done. For instance, grade point averages were shown, but no comparative analysis (e.g., t-test) was reported; and the percentage of women and men in each engineering major was shown, but no comparative analysis (e.g., chi square) was reported. In one instance (comparison of attitude items from CSQ Parts I and II) probabilities of difference were reported, but no data were provided for the reader to verify the results and there was no description of how the probabilities were computed.

The author exhibited a disconcerting tendency to change hypotheses when discussing various points in the article. For example, when describing the purpose of the study, a null hypothesis was stated, "The attrition rate, as a summation of all sources of attrition for the two-year study period, will not be significantly different for the two groups" (p. 233). However, when presenting the data relative to this point, the hypothesis was reworded; "Observed attitudinal differences will produce no significant differences in attrition rates over the two-year period" (p. 234). The data which were presented were appropriate as a basis for testing the first hypothesis but inadequate for testing the reworded cause-effect hypothesis.

Additionally, the author's conclusions were not aways supported by the data that were presented. For example, in discussing the women's anticipated source of job satisfaction, the author concluded that "women became <u>far less interested</u> in being helpful to others and society" (emphasis added, p. 235). However, the data presented show only how many women in the two CSQ surveys indicated each of nine motivations as their "most <u>important</u> anticipated source of job satisfaction" (emphasis added, p. 235). The fact that the number who chose "be helpful to others" fell from 15 to 6, in the context of all the responses, indicated only that fewer women chose this source as their most important anticipated source of job satisfaction: In point of fact, their interest in being helpful to others may have increased over the year, but not as much as their interest in other job satisfaction motivations; or a number of other scenarios are possible. The point is that the reported data do not indicate degree of interest, and a conclusion relative to degree of interest is not warranted.

Despite these problems, the study does suggest several interesting avenues of investigation. A study which tests the author's underlying concern with cause-effect relationships between attitude and academic performance and choice of major could be fruitful. The author's speculation about what kind of special program is necessary in order to counteract attrition of women engineering students would be very useful. Of particular interest is the notion that a "critical mass" of women students can supply a mutually supportive environment that would combat attrition problems and that anything short of that "critical mass" will result in high attrition rates. Finally, the author suggested that women engineering students had broader interests than those of men and that those broader interests led the women to leave engineering more often than men, who left engineering more often for academic reasons. This statement needs verification. However, if it is correct, then rather than denigrating women for leaving engineering due to lack of specificity in career interests, consideration ought to be given to broadening the career interests of men who select engineering, just as attention has been given (quite correctly) to broadening the career interests of women.

73

Palmer, Glenn A. "Students' Relative Interest in Selected Science Categories." School Science and Mathematics 77(6): 477-483, October, 1977.

Descriptors--Biological Sciences; Earth Science; *Educational Research; *Grade 8; Junior High Schools; Physical Sciences; *Science Education; Secondary School Science; Statistical Analysis; *Student Science Interests

Expanded Abstract and Analysis Prepared Especially for I.S.E. by John P. Smith, University of Washington.

Purpose

The purpose of the investigation was to identify within specified science categories junior high school students' interests using four science interest inventories developed by the author.

Rationale

As the author indicates, studies investigating students' science interests have taken many forms. Previous studies have for the most part focused on soliciting student-initiated responses rather than selecting from an investigator-generated list of science categories.

Science topics used in the author's inventories were selected from several of the sciences. The reduction of ambiguity among the categories enabled the investigator to measure internal consistency of response sets for each student and to measure agreement across response sets for each student. Previously reported results suggest a student's science interests are not based on random selection--that some internal consistency is apparent. These results also indicate "statistically significant reliability across form and format" for the investigator developed inventories.

Research Design and Procedure

Two inventories, Forms I and II, were developed giving approximately equal weight to biological, physical, and earth-space sciences. Each form included statements from the following 15 categories: A mechanics and kinematics, B ecology, C cytology and embryology, D physical geology, E atomic Theory, F radiant energy, G historical geology, H astronomy, I chemical change, J heredity and evolution, K physiology and morphology, L electricity, magnetism, and electronics, M biological taxonomy, N kinetic theory, and O meteorology.

Form I emphasized the manipulative aspects of science in the sense of how does one gain knowledge through direct experience. Items in the inventory are grouped in triads as in the following investigator-provided example:

A. Perform tests to determine the effects of wind blowing on different kinds of soil.

B. Perform tests to determine how much larger a lens makes objects appear.

C. Perform tests to determine how air movement influences the rate of evaporation.

Form II inventory was designed to assess student interests in nonmanipulative or passive aspects of gaining scientific information, i.e., obtaining information from reports of popularized articles. Form II items were also grouped in triads as shown in the following investigator-provided example:

A. Read a book which has the title, Light and Color.

B. Read a book which has the title, Natural Forces Changing the Earth.

C. Read a book which has the title, Plants, Animals, and Man: Their Effect on Each Other.

A second set of inventories using a paired comparison format was developed from Forms I and II. Paired items were developed simply by combining items from each triad of Forms I and II in the following manner: Items A and B represented the first pair, A and C the second pair, and B and C the third.

The four inventories were administered to nine eighth-grade science classes selected from Wisconsin junior high and middle schools. (Details of the class selection process were not reported.) The total number of students involved in the study was 218.

Each student was asked to complete Forms I and II, the triad comparison inventories, and only one of the paired comparison inventories. In responding to each triad, students were asked to identify both the activity they believed to be most interesting to do and the activity they felt would be least interesting to do. Students were asked to respond to only one of the two paired comparison inventories. For the paired comparison inventory, students only marked the activity of most interest to them for each pair.

Findings

A rank ordering of categories based on individual student choices was made for each inventory. These individual rankings were not discussed in the article.

Class rankings of categories were reported and were calculated by combining individual rankings within each class. Class ranks were ordered on the basis of frequency of choice with the category chosen the most interesting with the highest frequency relative to other categories ranked first, and so on. The investigator does not mention how the markings of least interesting on the triad inventories was used, if at all, in determining the ranking for each category

A coefficient of agreement (K) was calculated for each class based on individual choices from the paired comparisons inventories. While the coefficients are admittedly small, the investigator stated, "It is unlikely that the agreement found within any of the classes could have come about through random allocation of preferences."

The similarity among rankings was calculated using Kendall's coefficient of concordance, W. The coefficient $W = 0.833$ for class rankings has a probability $p < 0.001$ of occurring by a chance ordering of ranks.

There was considerable variability in composite category rank orders among classes, but several categories emerged as clearly showing student preferences. Categories G historical geology, J heredity and evolution, and K physiology and morphology ranked sixth or higher in student interest among all classes. Categories E atomic Theory, I chemical change, L electricity, magnetism, and electronics, and N kinetic theory ranked tenth or lower among all classes.

Further comparison of rank orders by row totals across classes for all four inventory forms resulted in rankings very similar to those cited above, i.e., earth-space and biology categories tended to rank highest and physical science categories lowest.

Combining the rank orders of Forms I and II of the paired comparison inventories resulted in the following order for the 15 categories:

Rank	Designation	Category
1	K	physiology and morphology
2	J	heredity and evolution
3	G	historical geology
4	H	astronomy
5	B	ecology
6	M	biological taxonomy
7	C	cytology and embryology
8	D	physical geology
9	F	radiant energy
10	O	meteorology
11	A	mechanics and kinematics
12	I	chemical change
13	N	kinetic theory
14	E	atomic theory
15	L	electricity, magnetism, and electronics

Again, the rankings reflect the tendency of students to rate biological and earth-space sciences higher in interest than physical sciences.

Interpretations

The results of the four interest inventories led the investigator to make inferences about why biological and earth-space sciences were rated higher in interest among Wisonsin eighth-grade science students than were physical sciences. First, the investigator proposes and seemingly rejects the notion that the four inventories were written in such a way as to be biased against the physical sciences. The similarity of results between this study and other investigations would tend to support the investigator's position. Second, the inference that teacher background may have influenced student interest is given some support by the investigator. A final inference connecting student maturation with science interests is made but not discussed by the investigator.

The investigator concludes eighth-grade students may be expected to show less interest in physical sciences than in other areas of science.

ABSTRACTOR'S ANALYSIS

This study is similar to many other studies designed to identify areas of science most interesting to students. Investigations of this type have surveyed students from a wide range of social and economic backgrounds, different school sizes, rural and urban communities, etc., with generally the same results, i.e., students as a group generally demonstrate a greater interest in biological sciences than in physical sciences.

In my opinion, the greatest shortcoming of "interest" research is its lack of a theoretical base. The formulation of the problem in this study is in no way unique and, therefore, contributes nothing to the resolution of the problem. In fact, in this particular study, there is no presentation of a problem or statement made that in any way hypothesizes a relationship between student interest and some other aspect of education. The author fails to establish any link between the theme of his study and education theory. It is simply a report without a conceptual framework.

The author does make a contribution to the field through the procedures of using multiple interest inventories. The use of several inventories allows an investigator to calculate the internal consistency of choices for each individual and to measure the degree of concordance across inventories for each individual. This procedure is not new but is apparently seldom used in interest investigations.

The validity or generalizability of the study is difficult to judge due to lack of information in the report. The investigator says nothing about the student/class selection procedure beyond reporting they are Wisconsin eighth graders. One even has to look in the tables to find the number of science classes and number of students surveyed! It seems fair to ask how the science classes were selected. Was it a random sample from some specifiable population? In essence, to what extent are the students in this study representative of Wisconsin eighth graders? Of all U.S. eighth graders? There was an extensive analysis of the inventory results but the reader really knows nothing about the students surveyed.

One would similarly like to know more about the procedure for selecting the 15 science categories and the development of associated items used in the triad comparisons inventories The paired comparisons forms are derived from the triad comparisons forms. Was there a check for item bias? If so, in what way?

In the report section entitled Implications, the author limits his discussion to his inferences about why students rate certain fields of science as more interesting than others but fails to make any connection between these inferences and the data from the study.

Future studies focusing on the roots of student science interests would seem to hold greater promise for contributing to knowledge than would additional studies of student science interests per se. We would all admit to, I believe, likes and dislikes; to strong interest in some field of science and little or no interest in others. But we are not beyond rising through the threshold of excitement when exposed to an idea not previously considered, when hearing of a new discovery in a field remote from our own, or when topics previously thought dull are related with enthusiasm and sensitivity to the listener. Students are not beyond becoming excited about topics previously thought dull and uninteresting, either. This factor, however, seems to be overloooked in interest research. The interest investigator seems only interested in what is rather than why or what might be.

Moore, Kenneth D. "A Science Needs Assessment of Elementary Teachers," in Piper, M. and K. Moore, Eds. Attitudes Toward Science: Investigations. Columbus, OH: SMEAC Information Reference Center, The Ohio State University, 1977.

Descriptors--Educational Research; *Elementary School Teachers; *Inservice Teaching; Inservice Teacher Education; *Needs Assessment; Research Methodology; Science Education

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

Purpose

The purposes of the study were to: 1) assess the utlity of factor analysis techniques to identify science needs of elementary teachers; 2) determine the degree of intensity of these needs; and, 3) to assess the relationship between needs and the grade level and experience of the teachers.

Rationale

Changes in the elementary science curricula have changed the role of the elementary science teacher. The changes are reflected in the teachers' needs. Procedures for identifying and categorizing these needs should be a major concern. Also, inservice education should be based on these identified needs and should reflect their intensity.

Research Design and Procedures

Needs assessment data were collected from 107 elementary teachers in 21 school districts in Texas. This represented 53.5 percent of the 200 teachers randomly selected to respond. Each subject responded to 117 needs statements on a Likert scale from 1 (no need) to 4 (much need). Data were also collected on the grade level and experience of each teacher.

Teacher responses were submitted to principal components factor analysis. The mean scores on each of the stable factors were then calculated and one-way analyses of variance procedures were carried out to determine differences related to grade level or experience.

Findings

The factor analysis resulted in the identification of 13 factors with three or more items that loaded over 0.5. The need to provide realistic science experiences ranked first as a need; while additional study of the history and philosophy ranked last. The author also reported that primary teachers had the greatest level of need across the 13 areas and that teachers with less than four years experience had more than a moderate need for help on nearly one-half of the factors.

Interpretations

The author concluded that the generic science needs of elementary teachers could be categorized and identified through factor analysis, and inservice programs should be designed to meet these identified needs. In addition, primary level teachers and teachers with less than four years experience have unique needs that demand special attention.

ABSTRACTOR'S ANALYSIS

The identification of teacher needs to guide inservice programs is an important activity. In this study the effort to empirically categorize, through factor analysis, the needs into a manageable interpretable number is particularly laudable. The fact that the 13 categories are logical and interpretable should make organization of inservice programs a bit more systematic.

Although the list of needs factors appears to be comprehensive, there is no assurance from the procedures reported by the author that this is true. An inspection of the instrument used is needed to logically examine the comprehensiveness of the original needs assessment statements. The criterion for retaining a factor was for it to have three or more items loading over 0.50. It is possible that some generic need had only two (or one) items representing it on the instrument.

The author reported that there is indeed a difference in the degree of intensity of need of the identified factor needs. There is little question that the intensity of the needs are different at the extrenes, but one has to question whether there is a nonchance difference between any near adjacent or some not so near adjacent factor means. No statistical tests or dispersion data were provided; but there is little question that the confidence intervals around many of the means would overlap other inteivals within a very liberal alpha level.

In addressing the questions of grade level and experience differences in needs, the researcher appears to have been a bit over-zealous to support a "need" to find differences. Although no nonchance diffeiences existed in the data (i.e., differences in any set of means could be attributed to nothing more than sampling error), detailed discussion and interpretations of the differences were carried out. The desire to find significant differences in some sets of data is understandable; but, to explicitly interpret chance results is to assume the existence of a Type II error. If this is the case, the energies should be spent at the front end of the study to beef-up the power of the design and analysis.

In general, the study is a good one and provides some interesting and useful results on the generic needs of elementary teachers. It is, in fact, better if there are no differences between grade level and experience groups; the task of inservice will be easier.