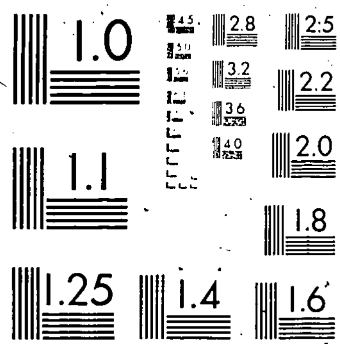

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DOCUMENT RESUME

ED 168 881 SE 027 509

AUTHOR	Blosser, Patricia E., Ed.; Steiner, Robert L., Ed.
TITLE	Investigations in Science Education, Vol. 5, No. 1. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION	Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE	79
NOTE	70p.
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State University, 1200 Chambers Rd., 3rd Floor, Columbus, Ohio 43212 (subscription $6.00, $1.75 single copy)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Discipline; *Educational Research; Inquiry Training; *Instruction; Process Education; *Research; *Research Methodology; *Science Education; Teaching Methods
IDENTIFIERS	*Keller Plan

ABSTRACT

Included in this issue of "Investigations in Science Education" are 10 abstracts all related to instruction. The articles relate to topics such as process skill acquisition, reading, grouping and pacing, inquiry and hunch generation, use of the metric system, and classroom management. Each abstract-analysis includes biographical data, research design and procedure, purpose, research rationale, and the abstractor's analysis of the research. (PEB)

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC® Clearinghouse for Science, Mathematics, and Environmental Education

Published Quarterly by

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

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

Volume 5, Number 1, 1979

As we begin a new volume and publication year, it seems appropriate to attempt another beginning: an issue of I.S.E. devoted to a single theme. Past issues have been either a collection of studies, each on a separate topic, or clusters of studies, with each cluster having a different focus. Volume 5, issue 1, will be focused on studies concentrated on a single theme: instruction.

Given the variety of individual interests, even studies whose authors identify them as related to instruction have different foci or include different variables. These studies are not exceptions. They involve influence of instructional site on acquisition of a process skill (Askham), use of rewritten scientific journal articles with students whose reading skills need improvement (Corey), grouping and pacing (Gabel and Herron), the Keller Plan (Putt), instructional approaches and classroom management orientation (Jones and Harty), inquiry and hunch generation (Wilson and Koran); inquiry vs. lecture (Schmitt and Groves), knowledge of the metric system (Henry and Rowsey), S-APA and Montessori methods (Judge), and active manipulation of materials vs. a more verbal approach related to the development of proportional reasoning (Wollman and Lawson).

Patricia E. Blosser
Editor

Robert L. Steiner
Associate Editor

INSTRUCTION

Askham, Leonard R. "The Effects of Plants on Classification Behavior in an Outdoor Environment." Journal of Research in Science Teaching, 13(1):49-54, 1976.

Descriptors--*Classification; *Educational Research; Outdoor Education; *Plant Identification; Science Education; Secondary Education; *Secondary School Science

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Eugene L. Chiappetta, University of Houston.

Purpose

The purpose of this study was to determine if children, ages 9 to 12 years old, classify plants growing in a semi-natural environment the same way as they classify non-natural objects in a classroom-like setting. In addition, if differences did exist between children's classificatory behavior, was this due to: (a) the type of test employed; (b) the subject's age, sex or race; or (c) the proximity of the objects to the subject during testing?

Rationale

Classificatory ability is a developmental process which becomes more specialized with age and with familiarity of the stimuli under question. Children and adults constantly employ this ability to organize and reduce the complexity of their environment. Researchers have followed two avenues in their search for understanding classificatory behavior. One avenue is descriptive where classifying ability is analyzed only, while the other avenue is experimental where classifying ability is improved through training.

Experimental researchers have been successful in improving children's classificatory skill. Some have reported that training enhances children's ability to group items, to describe their categories, and to employ a variety of criteria in their categorization (Olmstead et al., 1970). Others have reported improvement of children's grouping schemes through training (Irving and Olmstead, 1967). However, Askham believes descriptive studies which analyze the way children naturally classify are the most relevant.

Descriptive researchers identified many useful findings. Inhelder and Piaget (1964) point out that children move from sorting observable attributes to grouping using unseen or inferred characteristics. Peterson and Lowery (1968) formed two different behaviors exhibited by children. Some children continuously and intensely explored new objects if permitted to do so, while other children exhibited a lack of exploratory behavior. It is this type of research and findings that motivated the author of the present research study to investigate children's classificatory behavior under different conditions and to relate these findings to science instruction.

Research Design and Procedure

Sample. A random sample of 95 nine through twelve year old students was selected from classes participating in an education program at the University of California's Botanical Garden in Berkeley, California. One-half of the sample was assigned to classify plants in an outdoor environment, while one-half of the sample was assigned to classify geometric objects in a classroom-type environment.

Procedure. The subjects in each of the two groups were interviewed individually during their classification testing. Group 1 subjects were interviewed outdoors and presented with 20 plants growing in seminatural conditions (Test 1). Group 2 subjects were interviewed in a classroom-like setting and presented with 16 non-natural geometric objects (Test 2). All subjects were repeatedly asked the following questions until all possible responses had been elicited: "Can you think of any ways you could classify these objects?" and "Can you think of any other ways you might classify these objects?"

Design and Analysis. The research design appears to be a split-half factorial procedure where environment, race, sex, age, and subject-to-object proximity were the independent variables and classificatory responses, the dependent variables. Analysis of variance was employed to determine significant ($p \leq .05$) relationships among variables.

Findings

The author reported the following findings:

(1) Significantly more time was spent by subjects in classifying plants growing in a seminatural environment, than was spent in classifying geometrically shaped objects in a classroom-like setting.

(2) Non significant differences were found in subjects' classificatory responses that could be related to race, sex or age.

(3) More classification strategies were used by subjects classifying plants than by subjects classifying geometrically shaped objects.

(4) Mixed categories were used most frequently to classify plants in the seminatural environment.

(5) Shape was used most frequently to classify geometrically shaped objects in the classroom-like environment.

(6) The subjects used more diverse classification strategies than indicated by Inhelder and Piaget studies (1964).

(7) The subjects mixed their categories a significant percentage of the time.

Interpretations

This study showed that the children sampled, ages nine through twelve, used more varied and complex strategies to classify plants growing in a seminatural condition than to classify non-natural geometrically shaped objects in a classroom-like setting.

Indoors, the children restricted their categorizations and descriptions to a limited number of classification schemes. The single most important scheme used was that of mixing categories and the three most prominent categories used were shape, size and color.

The author feels that several implications of the study are noting. Elementary school children appear to employ a greater variety of categories in their classificatory responses than have been described in previous research in this area. This may be a result of most research being conducted in the classroom or laboratory as opposed to the outdoor or natural environment. As a consequence of these findings, more studies should be conducted in relevant settings outside of the classroom to really understand how children view the world of objects and events.

ABSTRACTOR'S ANALYSIS

This article draws attention to a crucial dimension to consider in understanding children's thinking, namely that of the effects of the stimuli on the responses. Children appear to react differently to stimulus situations which are complex than to those that are simple in nature. This word of caution is especially important to a recent trend in science education research which has addressed the development of concepts and mental operations in school-age children. The reason is that the stimuli used to assess student cognition in an experimental setting may not provide generalizations that are directly applicable to classroom settings where the stimulus situation may differ markedly.

One must be careful, however, in interpreting the results of the present study. The reported outcomes may be a little misleading. For example, the author emphasizes the importance of the seminatural or outdoor environment in eliciting numerous classificatory responses in children's behavior. It appears the critical variable was the use of live plants as opposed to man-made geometrically-shaped objects. Obviously the two stimuli were not equivalent in their complexity, yet similar in the concept they were chosen to represent, i.e., plants. If the author was really trying to assess the extent to which stimuli can elicit classificatory behavior, could he not have developed man-made objects which exhibit numerous attributes and complexity, such as a three-dimensional, multi-colored version of Creature Cards from the nationally recognized elementary science program, Elementary Science Study?

Several important explanations were omitted from this research report which make the analysis difficult. There was not an adequate description of the classroom-like setting or the outdoor setting, the plants or the man-made geometrically shaped objects. There was not a description of the sample regarding the socio-economic background or intellectual characteristics. In addition, there was not a results section that presented the statistical procedures and outcomes.

There is still room in the science education research literature for more studies on classificatory behavior. These studies really need to focus on the effects of long-term training with common and contrived stimuli as the training elements. Efforts to alter children's classificatory skill—say over a school year, using common objects and unusual objects—may produce marked improvement in conceptual ability.

REFERENCES

Inhelder, B. and J. Piaget. The Early Growth of Logic in the Child. New York: W. W. Norton and Co., 1964.

Olmstead, P.; C. V. Perks; and A. Rickel. "The Development of Classification Skills in the Preschool Child." International Review of Education, 16(1):67-80, 1970.

Peterson, R. W. and L. F. Lowery. " A Study of Curiosity Factors in First-Grade Children." Science Education, 52(4):347-352, 1968.

Corey, Noble R. "The Use of Rewritten Science Materials in Ninth-Grade Biology." Journal of Research in Science Teaching, 14(2):97-103, 1977.

Descriptors--Academic Achievement; Biology; *Educational Research; *Instruction; *Reading Level; *Scholarly Journals; Science Education; *Science Materials; Secondary Education; *Secondary School Science

Expanded Abstract and Analysis Prepared Especially for I.S.E. by David R. Stronck, University of Victoria.

Purpose

In the secondary schools, scientific journal articles cannot be used for supplementary learning because many students have serious deficiencies in reading skills. One solution to this problem is to rewrite scientific journal articles at an appropriate level of readability. The purpose of this study was to investigate the effect, on students, of reading such rewritten articles. The null hypothesis tested was the following: There are no significant differences in both a measure of comprehension and the reading rate between ninth-grade biology students who read scientific journal articles rewritten for their level of reading and similar students who read the original versions.

Rationale

Most textbooks used in the secondary schools were written one to five years prior to publication. A stated assumption is that up-to-date current material needs to be incorporated into the curriculum as a supplementary learning device. Another stated assumption is that "to a large extent" the secondary schools do not know how to cope with serious deficiencies in reading skills. Another stated assumption is that the rate at which a teacher can proceed depends largely upon the ability of the students to read critically. These three assumptions provide the rationale for rewriting scientific journal articles which were used as a supplementary learning device.

This investigation is related to the research studies in reading which were summarized by Dale: "In the preparation of textbooks and other expository materials much remains to be done to put the materials on the reading level of the learners."

Research Design and Procedure

This study used the Posttest-Only Control Group Design described by Campbell and Stanley. After the 152 students of the study read six articles, they completed a multiple-choice test. The test contained 100 items which were composed by the authors of the textbook used by the students in their biology classes. The investigator selected these questions because they appeared relevant to the material included in the six articles. The test has a reliability coefficient of 0.92 by the split-half procedure proposed by Stanley and 0.87 by the general Kuder-Richardson formula. The purpose of the test was "to measure the degree to which pupils were able to recall facts they had read in the science selections, to select the main ideas presented in the selections, and to formulate conclusions from information contained in the selections."

The six scientific articles were selected "for their appropriateness to the existing curriculum." The readability level of these original articles was found to be at grade 12.7. Each of the articles was rewritten according to the procedure suggested by Yoakam. The readability of the rewritten version of the articles was at grade 9.1. The Yoakam Readability Formula was used because "(1) it has been determined a valid technique, (2) it is easy to administer and evaluate, and (3) it provided a standard for determining the relative difficulty upon which materials could be rewritten to a lower level of readability."

The sample for this study consisted of the 152 ninth-grade students' enrolled in five classes of second-semester biology. The students were taught by a single teacher and used the same textbook: <u>Modern</u>

Biology by J. H. Otto and A. Towle, published in 1969. The teacher used a traditional lecture-recitation approach with each of the five classes which were heterogeneously grouped as to ability. These circumstances support the assumption that all students in the sample were receiving very similar instruction in biology.

During the previous year, the 152 pupils of the study completed three tests on reading ability as measured by the Iowa Tests of Basic Skills: (1) Test V: Vocabulary; (2) Test R: Reading; and (3) Total of Test L: Language Skills. The mean of the three scores obtained by each student was calculated and used to rank all of the students. This ranking allowed dividing the students into two groups: (1) those reading at or above grade level and (2) those reading below grade level. At or above grade level was taken as 7.7.

By the application of a table of random numbers, each of these two groups was divided into the following subgroups: (1) two groups from those reading at or above grade level, and (2) two groups from those reading below grade level. One group of those reading at or above grade level and one group of those reading below grade level were assigned to read the original versions of the six articles; they constituted the control groups. The others read the rewritten versions of each article and constituted the experimental groups of this study.

Earlier testing permitted the identification of those students who were reading at a rate of or faster than the mean reading rate, and those who were reading slower than the mean reading rate. The four groups described above were each divided into two groups on the basis of reading rates. This final division generated eight cells of 19 students in each from the total of 152 students.

Analysis of variance was used to test for significant differences among the eight sets. A table of F-distribution demonstrated differences at the 0.01 level of significance. Tukey's W-procedure was used to identify differences between specific pairs of means.

Findings

The investigator provides four tables of data which identify some significant differences among the cells. For the students involved in the study, there were significant differences in comprehension of the scientific articles between those who read the original versions and those who read the rewritten versions. Among students reading at or above grade level, students reading the rewritten versions had higher scores. The same pattern was found among students reading below grade level. The students reading at or above grade level had significantly better comprehension of the scientific articles than those students reading below grade level even when the superior readers read science selections especially prepared for students with poorer reading abilities. There were no significant differences found among the considerations of reading rates.

Interpretations

The investigator made the following generalization based on the findings: "Current scientific journal articles are written at a more difficult reading level than is necessary to adequately present the desired science content for many readers. Those students reading at or above grade-level placement as well as those reading below grade-level placement read with greater speed and better comprehension when materials are rewritten to a lower level of understanding."

The following implications for science teaching were identified by the investigator: "1. Science textbook-selection committees should consider the factor of readability as a criterion in adopting a science textbook or supplementary reading materials for classroom use. 2. Classroom teachers of biology should be encouraged to participate in rewriting difficult science materials to a less difficult level of readability for the poorer readers in their classrooms."

ABSTRACTOR'S ANALYSIS

This study provides excellent statistical support favoring the reduction of the difficulty in reading level in science textbooks and supplementary materials. The general conclusion from this study is that all students will benefit from reading levels which do not exceed grade level. This conclusion is a valuable contribution to a topic of great importance during the recent period of recognizing declining reading abilities. Many studies, including National Assessment, have clearly identified declining reading abilities among secondary school students.

The investigator recognized that some will argue against a simplification of reading levels because the better readers would not be challenged to develop their reading skills. Nevertheless, science teachers are primarily concerned with helping their students to enjoy and to understand scientific concepts. In recent years, science teachers have been selecting textbooks of lower reading levels and encouraging the publishers to provide such materials. The textbook market is now filled with advertisements for materials of relatively low reading levels.

The methodology of the study included an adequate sample of students and an impressive use of statistics to identify significant differences. The written report was inadequate in discussing the data of the study which was simply presented in tables with almost no comments.

The validity of the study may be questioned in terms of the selection of articles which were read by the students. The investigator did not explicitly describe any overlap between the instruction in the biology course and the contents of the six articles. Because all of the students involved in the study were learning biology from a single teacher, the investigator seemed to imply that the six scientific journal articles used in the study considered related biological topics. If there was much overlap, then students with high achievement in the biology course would easily have high scores on the related scientific articles. Because the test questions for the articles were taken from questions prepared for the biology textbook by the authors of the textbook, there seems to be great overlap. Such overlap would obscure the goal of the study, i.e., to identify the impact of materials at different reading levels on the comprehension of the students. Probably the better readers were also the better achievers in biology. Overlap between the content of the course with that of the articles will tend to exaggerate the distinction between the better students and the weaker students. This exaggeration could produce statistically significant differences which do not adequately represent the various reading abilities of the students.

The study used the Posttest-Only Control Group Design. Pretesting would have eliminated the suspicions described in the previous paragraph. Unfortunately the investigator offered no explanation for selecting the Posttest-Only Control Group Design which assumes that all of the students had a similar ignorance of the content of the articles before reading them. Previous comments in this analysis indicate that probably the students had a great variety of understandings about the concepts and vocabulary used in the articles. Pretesting with posttesting could have eliminated this probable variable.

Most science educators agree with the investigator's assumption that up-to-date current material needs to be incorporated into the curriculum as a supplementary learning device. Many secondary school teachers use newspaper articles, or science reports from news magazines, or popular science magazines, e.g., <u>The Scientific American</u>. Such articles are designed for the reading level of the general public, i.e., for people with relatively poor scientific literacy. Unfortunately the investigator ignored such sources and implied that only articles published in scientific research journals will provide up-to-date current material. His advocacy of the scientific journals may be justified if the goal of the science teacher is to consider the methodology of the current scientific researcher. But if the goal is simply up-to-date current information about the latest scientific discoveries, then educators already have an abundance of excellent sources at reasonable reading levels. Some of these sources present the most recent advances of science more quickly than do the typical scientific research journals.

The investigator assumed that "to a large extent" the secondary schools do not know how to cope with serious deficiencies in reading skills. In recent years, curriculum materials for science teachers have offered a wide variety of solutions to serious deficiencies in reading skills. Many schools throughout the nation have adopted programs of individualized instruction which allow for much greater accommodation of reading levels than traditional textbooks. Research studies with the Intermediate Science Curriculum Study have demonstrated that individualized instruction greatly assists the improvement of reading skills. The impressive array of new supplementary science materials for the secondary schools seems to demonstrate that the schools, with little effort, could meet the problem of declining reading levels. The investigator suggests: "Classroom teachers of biology should be encouraged to participate in rewriting difficult science materials to a less difficult level of readability for the poorer readers in their classrooms." A more obvious suggestion seems to be the encouragement to identify appropriate materials which have been prepared by commercial publishers.

Another stated assumption of the investigator is that the rate at which a teacher can proceed depends largely upon the ability of the students to read critically. Certainly this is true in a traditional lecture-recitation approach which was the approach used within this study. On the other hand, the trend toward individualized instruction is reducing this problem. For example, teachers using the Individualized Science Instructional System are able to challenge the bright students while simultaneously caring for the needs of the slow students. The recent recognition of the great variety of abilities in the secondary schools is encouraging the adoption of such flexible new systems.

When this abstractor did the research of his own dissertation, he used the Portland Science Test which attempted to avoid variations in reading skills through the use of diagrams and drawings. Nevertheless this study demonstrated strong correlations between scores on reading tests and scores on the Portland Science Test. Apparently students who do not read well also have difficulties in achieving scientific concepts. The importance of the link between reading and understanding science is now receiving serious attention. Nevertheless, the importance of the topic suggests that many additional studies are needed. For example, highly motivated students will read materials which are far above their reading level. Studies are needed to identify the style of writing which will best motivate students. Additional studies are needed on the long-term use of science materials of relatively low reading levels.

REFERENCES

Dale, E. "The Future of Reading." Journal of Reading, 23:212, 1969-1970.

Stronck, D. R. "Comparisons Between Reading Abilities and Achievements in Two Aspects of Science in Seventh-Grade Students of Various Socio-Economic Levels." California Journal of Educational Research, 23:198-206, 1972.

Yoakam, G. A. Basal Reading Instruction. New York: McGraw-Hill Company, Inc., 1955.

Gabel, D. and J. D. Herron. "The Effects of Grouping and Pacing on Learning Rate, Attitude and Retention in ISCS Classrooms." Journal of Research in Science Teaching, 14(5):385-399, 1977.

Descriptors--Achievement; Attitudes; *Educational Research; Elementary Secondary Education; *Grouping (Instructional Purposes); *Instruction; Junior High School Students; *Pacing; Retention; Science Education; *Secondary School Science

Expanded Abstract and Analysis Prepared Especially for I.S.E. by James A. Shymansky, The University of Iowa.

Purpose

The major purpose of this study was to examine the effect of allowing students to pace themselves to achieve mastery versus imposing a deadline for completion of chapters in the seventh-grade Intermediate Science Curriculum Study (ISCS) materials. Criterion variables in the study were learning rate, retention, and attitude. In addition to examining pacing, the effect of working by oneself or with a partner was also studied.

Rationale

The research reported in this study is tied generally to the mastery learning model of Bloom (1968). This study attempts to extend previous studies of learning rates conducted by Kress and Groper (1964), Merrill, Barton and Wood (1970), and Merrill and Stolurow (1966) which dealt with short time spans in programmed instruction situations and to resolve some of the issues raised by the conflicting results of Wong (1970), Wong and Lindvall (1970), and Yeager and Lindvall (1967) in their learning rate studies of classroom populations.

Research Design and Procedure

The study utilized a post-test only design involving 43 intact seventh-grade ISCS classrooms (1022 students) drawn from four county schools and six city schools. Because of perceived variations in student, parent, and teacher attitudes toward school and differences in teacher/pupil ratios, the data for the two-sub groups were analyzed and reported separately. The 12 teachers involved with the 43 intact classrooms were randomly assigned to either the self-paced strategy or imposed deadline strategy, the main independent variable.

The dependent variables studied included retention and learning rate as measured by chapter tests developed by the ISCS staff at Florida State University and attitude as measured by A Scale to Measure Attitude Toward Any School Subject.

For the purpose of analysis, the data for the four county schools and six city schools were separated. In the ANOVA procedures applied, data were blocked by ability group (as measured by the Otis-Lennon Test of Mental Ability), student arrangement (individual vs. partner), and ISCS chapter covered. Along with the main effect of pacing, then, a Pacing-Grouping-Ability interaction effect for each criterion variable was also analyzed.

The actual duration of study was not reported except to say that tests for retention and learning rate were administered after each of four ISCS chapters.

ANOVA summary tables are reported for each of the dependent variables (retention, learning rate, and attitude) for both the county and city school samples. In addition, plots of cell means from the ANOVA for learning rate and ability group are presented.

Findings

Learning rate was found to be greater for county school students who worked alone rather than with a partner, but there was no difference in learning rate observed as function of the self-pacing or imposed deadline variable for these same students. In the city school sample, the results were somewhat different. Students who worked in partnerships learned at a faster rate than individuals. It was also found that students working in the self-paced classrooms had higher learning rates. This was particularly true for low ability students and, to a lesser extent, for average ability students. Furthermore, low ability students worked at the lowest rate when working alone with deadlines and best when working alone with self-pacing. For average and high ability students, working with a partner produced the higher rates.

In analyzing retention effects a slightly different blocking arrangement was used. A 2x2x3 design was produced by using the two pacing levels, two ability levels, and three partner groups (individual, heterogeneous, and homogeneous). Results showed significantly greater retention rates for city students in the self-paced classes and a similar trend for the county students, though it was not significant. County children who worked alone had higher retention rates than those in partnerships but no differences were found in city school children.

Low ability students obtained higher retention scores when working in partnerships in both the county and city samples. High ability city children who worked alone scored higher while their county counterparts scored somewhat lower than did those in partnerships on retention.

Analysis of attitude scores produced no discernible trends as a function of pacing or grouping. County school students appeared to like ISCS more than city children, but not significantly so.

Interpretations

The explanation offered for the finding that low ability students seemed to fare batter in self-paced classrooms than in those with imposed deadlines is that the dreaded thought of staying on one topic until mastery level learning was accomplished may have been motivating to these students. In addition, there was greater opportunity for remedial work and extra help in the self-paced classroom.

Though the findings for average and high ability students were not as consistent as for the low ability groups, it appears that the practice of imposing deadlines so that students "move along" probably leads to less learning than insisting that mastery levels of learning be achieved—at least for the low ability students. The authors conclude that apparently low student motivation and lack of interest are at the heart of the problem and that imposed deadlines forcing students to move through the material faster don't solve the problem.

ABSTRACTOR'S ANALYSIS

While it is difficult to comment on the exact niche a study such as this occupies in the matrix of studies of instructional strategies, this study is significant. It attempts to deal with the much highly touted concept called "mastery learning." The study attempts to sort out the myths and the facts related to mastery learning in the context of a junior high school science program. Is there any proof that the Bloom model can be practically applied in a junior high school science classroom? What are the tradeoffs, the limitations, the payoffs? This study addresses these questions and has implications for the broader question relating to the degree of individualization possible in the science classroom.

The authors utilize a unique approach in the research design by including the retention rate as a major dependent variable in addition to the standard variables: achievement and attitude. The question of how well materials and concepts "sink in" is oftentimes overlooked in studies of instructional strategies concerning the effectiveness. The inclusion of retention rate and the overall breadth of the dependent variables studied is the strongest dimension of the investigation.

The article is clearly written. The statistical tables and graphs provided are adequate. The extensive discussion and recommendation sections are especially valuable. The authors include anecdotal data throughout the discussion section which add support to several of the conclusions and interpretations.

The inclusion of several blocking factors and the split of the sample into county and city for the purpose of analysis led to some interaction effects and inconsistent results across the two sub-samples. Thus, generalizations concerning the main effects of pacing and grouping are limited. However, the authors do draw out appropriate conclusions from the analyses.

Perhaps the weakest aspect of the study is in the specification of the treatments. The main effect of self-pacing vs. imposed deadlines seems clear enough on the surface as do the blocking factors, ability, grouping, and chapter. However, the variation implicit in the ongoing instruction of the ISCS classrooms studied is not adequately discussed or recognized in the report. As a result, the reader is left wondering about the level of consistency within treatments. For example, it is erroneous to assume constancy of teacher effect across the sample unless some monitoring is done. It is generally accepted that the teacher plays an equal if not dominant role in overall classroom strategy when compared to that of the curriculum materials being used. In other words, one cannot assume that there is an "ISCS teacher," an "ESS teacher," or any other "curriculum-type" teacher in classroom studies.

The failure to control for teacher effect jeopardizes the external validity of the study. The authors allude to the possible effect of certain "strong teachers" in their discussion of learning rates, but the discussions generally overlook this effect. One cannot be sure that differences occurring between students in the various treatment groups are the result of the independent variables or the unknown teacher effect.

The research questions addressed in this study are clearly important. We need to know what the trade-offs are between self-pacing and group instruction; is mastery learning a viable concept; do student partnerships in the classroom help or hinder learning and attitude? But these factors cannot be studied without reference to the classroom teacher.

There are two basic types of research designs worth considering in dealing with this problem of accounting for teacher effect in the study of instructional variations. Small scale studies involving one teacher represent one alternative. Here teacher effect can be controlled completely, but generalizability is very limited. At best one can explain the control parameters and leave the question of appropriateness of results and interpretations to the reader.

A second alternative is to incorporate the teacher variable into the experimental conditions. This can be accomplished by training teachers to assume specific roles in the classrooms for the purpose of the study. All teachers can be trained to exhibit the same behavioral patterns, thus eliminating the teacher as a design variable, or variations in teaching behavior can be planned to augment the main variables under study or to define a new variable altogether. In either case, an effort must then be made to monitor teacher effect and these data should be incorporated in the report.

Adding the teacher variable to studies of program effectiveness and instructional variations requires a great deal more preparation for the research execution and complicates the research design. Classroom monitoring of experimental conditions either live or via audio/videotapes is costly and time-consuming. But the increase in overall validity and the descriptive information of the critical teacher variable provided far outweight the extra investment. In this otherwise well-designed and executed study, the validity of the findings would have been greatly enhanced by attending more specifically to the teacher variable implicit in the classroom.

REFERENCES

Bloom, B. S. "Learning for Mastery." Educational Comment, 1(2), 1968.

Kress, G. C., Jr. and G. L. Gropper. "Studies in Televised Instruction 4, A Summary Report." Pittsburgh: American Institute for Research in Behavioral Sciences, University of Pittsburgh, 1964. (ERIC Document Reproduction Service No. ED 033 202.)

Merrill, M. D.; K. Barton; and L. E. Wood. "Specific Review in Learning A Hierarchical Imaginary Science." Journal of Educational Psychology, 61(2):102-109, 1970.

Merrill, M. D. and L. M. Stolurow. "Hierarchical Review Versus Problem Oriented Review in Learning an Imaginary Science." American Educational Research Journal, 3(4):251-261, 1966.

Wong, M. C. and C. M. Lindvall. "An Exploratory Investigation of the Carroll Learning Model and the Bloom Strategy for Mastery Learning." Pittsburgh: Research and Development Center, University of Pittsburgh, 1970. (ERIC Document Reproduction Service No. ED 054 983.)

Yeager, J. L. and C. M. Lindvall. "An Exploratory Investigation of Selected Measures of Rate of Learning." Journal of Experimental Education, 36(2):78-81, 1967.

Henry, Loren L. and Robert E. Rowsey. "Comparative Study of the Knowledge of Metric Units of Measure and Their Application." Science Education, 62(3):283-289, 1978.

Descriptors--*Educational Research; *Higher Education; *Instruction; Knowledge Level; *Mathematical Applications; Mathematics Education; *Metric System; Science Education; *Secondary Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Chris Pouler, E. Roosevelt Senior High School, Greenbelt, Maryland.

Purpose

This study was intended to assess the (a) knowledge and (b) application ability of the International System of Units by three groups of students: college preservice education majors, high school seniors, and junior high school eighth graders. Using data from two specific instruments, the answers to the following research questions were determined:

1. Will there be a significant difference between performance on the relationship and the application instruments for college, high school, or junior high school subjects?

2. Will there be a significant difference between (a) college and high school subjects, (b) high school and junior high school subjects, or (c) college and junior high school subjects in their performance on the applications instrument?

3. Will there be a significant difference between (a) college and high school subjects, (b) high school and junior high school subjects, or (c) college and junior high school subjects in their performance on the relationships instrument?

Rationale

Because the metric system is the standard international language of measurement, the United States has, by legislation, adopted its use. Correspondingly, materials and programs for metric education have been developed. This study sought to determine current knowledge and application of the Standard International (SI) units.

Research Design and Procedure

Population. Three groups of students were selected:

1. College seniors enrolled in either methods or curriculum courses designed for prospective mathematics or science teachers. These students attended Auburn University, winter quarter, 1977.
2. High school seniors enrolled in chemistry or physics courses of a southern, city high school which had a total student population of 1433.
3. Junior high eighth graders enrolled in an independent ISCS class. The school was located in an urban setting and served 790 seventh and eighth grade students.

The researchers assumed the high school seniors and the junior high eighth graders represented the upper quartile of students in their respective schools. Therefore, the scores on the two instruments would not be significantly different than any of their classmates not involved in this study.

Instruments. The researchers constructed two instruments to assess (a) metric relationships and (b) metric applications.

The Metric Relationship Instrument (MRI) consisted of 20 items "which were judged most needed by a literate population for everyday, non-scientific use" (p. 284). Subjects were given 20 minutes to complete the test which consisted of metric unit conversion problems. For example, 14,000 meters equals _____ millimeters. Two of the items involved conversions of square and cubic units (e.g. _____ square decimeters equal _____ square meters). The score of each participant was conveniently rated as the number of correct responses. Further, the Kuder-Richardson Formula 20 reliability coefficient was calculated to be 0.94.

The Metric Applications Instrument (MAI) also consisted of 20 items which were rated by the number of correct responses. The items were contained in a slide-tape presentation which first showed the subjects a "commonly known object from the environment" (p. 284) and then, asked them to write the metric units of either length, mass or capacity for a specific characteristic. For example, "about how many kilograms does the chicken weigh?" (p. 288). Since the application questions involved variability within an acceptable range, a panel of qualified experts judged as valid the limits for each item. The Kuder-Richardson Formula 20 reliability coefficient was 0.74.

Procedures. Within each group (college, high school, and junior high), the students were randomly placed into two subgroups. One subgroup was selected to take the MAI prior to the MRI. The test sequence was reversed for the other group. The mean scores of the groups were analyzed with t-tests.

Findings

Question 1

> "Will there be a significant difference between performance on MRI and the MAI for college subjects, high school subjects, or junior high school subjects?" (p. 285).

The performance of each of the three groups was significantly different for the MRI when compared to the results of the MAI. In fact, the difference between the relationships and application mean scores were 177 percent higher for college students, 176 percent higher for high school students and 224 percent higher for junior high school students.

Question 2

> "Will there be a significant difference between college subjects and high school subjects, high school subjects and junior high school subjects, or college subjects and junior high school subjects in their performance on the MAI?" (p. 285).

There was no significant difference between mean score of the college students and that of the high school students on this measure.

Significant differences did occur when the college and high school students' scores were compared to those of the junior high students.

Question 3

"Will there be a significant difference between college subjects and high school subjects, high school subjects and junior high school subjects, or college subjects and junior high school subjects in their performance on the MRI?" (p. 286).

No significant difference occurred between the college student group mean score and the high school student group mean score. Similarly, there was no significant difference between the high school group and the junior high group. There was, however, a significant difference between the college student group and the junior high group.

Interpretations

1. In analyzing the test items, the questions dealing with square and cubic units on the MRI were confusing to all participants.

2. All three groups exhibited higher mean scores on the MRI than the MAI. This indicates "inadequate experiences in actually using SI units of measure" (p. 287).

3. The contention that college, senior high school and junior high school students "are able to perform paper-pencil computations with SI units without being able to apply these units to describe the world in which they live" was supported (p. 287).

ABSTRACTOR'S ANALYSIS

This study was intended to assess (a) the knowledge of metric relationships and (b) the application ability of three distinct groups of students. The analysis must, therefore, discuss the results in terms of the sample population and the assessment instruments.

The participants represented secondary and college students who were particularly interested in science and mathematics. Since the researchers did not wish to involve all the junior and senior high students from the two selected schools, they chose only students who (a) were either in eighth grade and enrolled in a self-paced science course or in twelfth grade and taking either chemistry or physics and (b) represented the upper quartile of students in the schools. The assumption that these students would score no lower than students not in the study is reasonable. As for the college group, these students were education majors who wished to teach science or mathematics. Generalizations can, therefore, be made regarding secondary school students interested in science and college science and mathematics education majors. Although hypotheses were missing, it would have been reasonable to assume the college group would have a greater knowledge and application ability of the metric system than either the junior or senior high group. Likewise, the senior high group could have been hypothesized to outperform the junior high group on both measures. Results of no significant difference could be explained by (a) the lack of emphasis on the metric system in coursework; (b) the possibility that the necessary knowledge and metric application are attained by a specific age, thus making academic level irrelevant; or (c) the validity of the instruments. Since the type of coursework and age of acquisition were not specifically determined, an analysis of the instruments would be worthwhile.

The instruments developed for this experiment are interesting. The Kuder-Richardson statistics are impressive as is the care taken to use a qualified panel of experts to judge the items. But are the tests valid indicators?

The relationships test (MRI) appears to measure relationships as determined by metric conversions. Metric conversions are not difficult, providing the student understands units of ten. Thus, it is possible for the bright junior or senior high student to outperform a college student who has never learned the simple unit relationships. The lack of difference between the college and high school groups is not surprising. Further, the lack of significance between high school and junior high groups is also not surprising. The expected difference between college and senior high groups is, perhaps, related to additional coursework. Assuming the test does measure metric relationships, the data indicate coursework as the factor for higher scores and not age. An interesting future study which attempts to correlate the way the metric system is taught, the specific textbook or curriculum utilized, and the MRI score would prove interesting.

The applications test (MAI) appears to measure perception ability, general knowledge of mass, length and volume as well as the ability to apply metric units. When students are asked to express the weight of a chicken in grams, more than application of the metric system is involved. For example, students need to know the weight of a chicken in either pounds or grams before they can express an answer. Since many younger students have had little experience estimating mass, length, and volume, it is not surprising that the junior high group had a significantly lower score than the other groups. A better research hypothesis would have been: that students who can perceive units of measurement within a certain range can apply the metric system better than students who cannot perceive well. Since it appears that perception ability is lacking, the data indicating lower scores on the MAI than the MRI for all groups are not unusual. An interesting follow-up study would be to compare the results of students from a country where the metric system is exclusively used with the results from this study. Further, additional research using items that students can perceive might provide insight.

Specifically, the findings of this study offer some basis for an assertion that the metric system cannot be applied by students. However, sweeping generalizations are not supported! The study was well done. Although the sample sizes were small (24-32), the procedure was sound and has provided data for those interested in metric education.

Jones, Dan R. and Harold Harty. "Instructional and Classroom Management Preferences of Secondary School Science Teachers." Science Education, 62(1):1-9, 1978.

Descriptors--*Educational Research; *Instruction; Instructional Designs; *Perception; Science Education; *Science Teachers; Secondary Education; *Secondary School Science; Teachers

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

Purpose

This investigation was designed to ascertain whether there was a relationship between science teachers' preferences for instructional approaches and their classroom management orientations. The investigators suggested that a teacher expressing a high preference for "inquiry" instruction might also reflect a strong preference for "humanistic" management. Conversely, it was felt that a strong preference for "traditional" instruction might correlate positively with "custodial" classroom management preferences.

Rationale

The investigators cited work of Kerlinger and Walberg as they developed a perspective for this study. There is considerable evidence that attitudes and values held by teachers, along with selected personality variables, influence how they behave in their role as instructional and managerial leaders. The assumption is that teachers who prefer a more "traditional" or "directive" approach to instruction also prefer to manage students in a more "structured" or "restrictive" manner. Likewise, teachers who prefer the more "progressive" or "inquiry" methods of instruction will tend to prefer a more "permissive" or "humanistic" mode of managing students. The researchers in this study were interested in examining these assumptions by correlating teacher responses to their Science Teacher Ideological Preference Scale (STIPS) with responses from the same teachers to an already developed, validated scale developed by Willower et al., called Pupil Control Ideology (PCI).

Research Design and Procedure

The first part of this study was devoted to developing the STIPS instrument. Based on the work of Hurd, Schwab and Joyce and Weil, positive, third-person statements were designed to reflect two generalized instructional strategies; inquiry and traditional. The items were submitted to a group of 18 science education specialists for face validity. From the screening, a list of ten statements for each instructional approach was selected and randomly ordered. The final list of 20 items was then cast in Likert scale format with weights of 5 for "strongly agree," 4 for "agree," 3 for "undecided," 2 for "disagree," and 1 for "strongly disagree" assigned. Respondents to STIPS are scored on each of the two subscales.

The second instrument used in this study was the Pupil Control Ideology (PCI). This 20-item instrument has been shown to provide a measure of the degree of humanistic or custodial pupil control ideology. Reliability and validity estimates have been reported by Willower et al.

A population of 44 inservice science teachers was selected for this study. No sampling plan or randomization techniques were attempted. Participants were selected on their accessibility and their willingness to participate. The teachers in this sample were from inner city and suburban schools. The 44 teachers were from seven high schools, four junior high schools, and one middle school. The sample included 33 males and 11 females with ages ranging from 21 to over 60. Experience ranged from 16 teachers having one to five years to four teachers having 20 years or more. Twelve teachers had a bachelor's degree, 31 had a master's degree, and one had a doctorate.

The subjects were not informed of the purpose of the study nor were they told anything about the underlying nature of the two instruments. They were asked to respond to the items on both instruments in terms of their own classrooms. Both instruments and the demographic checksheet were administered and collected within a one-week period as the schools were visited.

Findings

Data collected using STIPS were scored on the basis of the two subscales: inquiry and traditional. The possible range on both subscales was 10 to 50. Data from the PCI were tabulated so that a single score was produced. The range possible was 20 to 100.

The investigators grouped the frequencies of the STIPS scores into four intervals for "ease of better depicting the frequency distribution of scores." These results were:

Range	Number of Teachers In This Range on Inquiry Subscale	Number of Teachers in This Range on Traditional Subscale
10-20	0	1
21-30	9	18
31-40	31	22
41-50	4	3

The mean inquiry subscale score was 34.16 and the mean traditional subscale score was 31.73. Reliability estimates using coefficient alpha were .73 and .70 respectively for the inquiry and traditional subscales. When teacher scores on the two subscales were correlated, a coefficient of .32 ($p < .05$) was yielded. Subject scores on the PCI ranged from 36 to 72. Looking at four intervals the number of scores within each interval was:

Range	Number of Teachers Within Interval
20-40	1
41-60	26
61-80	17
81-100	0

The mean PCI score was 57.96. A coefficient alpha of .64 was calculated for the 20-item PCI.

Correlation coefficients were calculated to study possible relationships between the inquiry and traditional subscales of STIPS and PCI. The correlation between the inquiry subscale and PCI yielded a −.22 (nonsignificant) whereas the correlation between the traditional subscale of STIPS and PCI was calculated at .32 ($p < .05$).

Correlation coefficients were also computed between various demographic variables and the STIPS two subscales and the PCI. Only two correlations were found to be significant. Sex of the teacher correlated .32 ($p < .05$) with the traditional subscale of STIPS. The number of years of teaching experience correlated .84 ($p < .01$) with scores on the PCI.

Interpretations

The investigators stated that the nonsignificant negative correlation (−.22) between the STIPS inquiry subscale scores and the PCI scores is probably attributed to the predictive inverse directional nature of the scores. Although not significant, they state this trend appears reasonable because inquiry learning environments in science teaching seem to more closely resemble a humanistic pupil control ideology than they resemble a custodial ideology.

The significant positive correlation (.32) between the STIPS traditional subscale scores and the PCI scores was predictable because of the similar directional nature of the two instruments. The authors suggest that this is reasonable since to teach science using a traditional strategy would require more conventional, restrictive, teacher-controlled classroom activities; hence, a more rigid or custodial management orientation.

When the two subscales of the STIPS were correlated, a significant ($p < .05$) positive correlation (.32) was produced. The STIPS, therefore, did not, with these 44 teachers, demonstrate the expected inverse relationship between preference for traditional vs. inquiry teaching. The investigators concluded that the STIPS does not discriminate teachers' preferences toward the two generalized instructional approaches but that it does appear to describe the tendency of a population's orientation toward one of the other of the strategies.

When demographic variables were considered, only one important relationship emerged. The number of years of experience correlated .84 with scores on the PCI. The researchers in this study suggested that the period of time in which the teachers received their college preparation may have influenced their propensity to prefer inquiry or traditional teaching.

ABSTRACTOR'S ANALYSIS

There has recently been a lot of interest in studying belief systems or the "educational philosophies" of teachers and how this relates to their instructional and managerial preferences and behaviors relative to the classroom. The investigators in this study address this general and timely question. They are quick to point out several limitations of the study and to acknowledge that their work is an exploratory endeavor designed to stimulate further research on this topic.

This paper contains a good review of the literature and develops a solid conceptual framework on which to proceed. It does appear that teachers can be classified into "traditional" versus "progressive" viewpoints and that this represents a legitimate link to how teachers view and behave toward students. The Pupil Control Ideology (PCI) instrument used in this study appears to be a reasonably reliable and valid measure of the degree of permissive vs. custodial pupil control ideology.

One of the most conclusive findings in this study was the high positive correlation between the number of years of teaching experience and scores on the PCI. This suggests, and is in harmony with speculations of the investigators, that teachers with more years of experience tend to possess a more traditional, custodial pupil control ideology. The investigators in this study also reported that sex of teacher correlated .32 ($p < .05$) with the traditional subscale of STIPS, but this was not explained in terms of direction or statistical methodology.

Most of the remaining comments about this study should impinge, I feel, on the validity of the Science Teacher Ideological Preference Scale (STIPS). The scale was developed by the investigators in this study and consists of two 10-item subscales. One subscale was designed to measure preference for the inquiry teaching strategy. The other subscale was designed to measure preference for the traditional teaching strategy. Respondents to the STIPS are scored separately on each subscale; hence, two scores are yielded. Scores range from 10 to 50, with higher scores suggesting a stronger preference for the respective teaching ideology being measured.

The STIPS was developed in a manner somewhat unorthodox to most scales using the Likert format. While it is purported to measure attitude toward science teaching ideology, all of the statements included in the instrument were constructed in a positive stance. Using this method, a total scale score was not possible; two scores were necessary for each subject. When scores for the two subscales were correlated, a coefficient of -.22 was produced. This suggests that less than 5 percent of the variation of one subscale was predicted by the other. By having 10 positively written statements about traditional teaching and 10 positively written statements about inquiry teaching, it was possible to produce scores reflecting agreement (or disagreement) with both ideologies. This was apparently the case here. It does not appear that the two subscales were measuring two different psychological objects ("inquiry" vs. "traditional" teaching), if so, the two subscales were not viewed by these teachers as different constructs.

When instruments like the STIPS are constructed, it is generally considered necessary to factor analyze items in order to develop a case for the validity of claiming two subscales. One would suspect that with this instrument the teachers in this study were reacting to the two sets of items in much the same way, suggesting from at least a statistical standpoint that they were not regarding the two ideologies as contrasting or bipolar entities. Also, had the scale been constructed where a total score reflected a respondent's position along a continuum of "high inquiry orientation" to "low inquiry/orientation," one would be able to develop a sense of distribution for a given population along this "natural" continuum. It is highly possible, and this is a conclusion one could draw from this study, that "inquiry" teaching strategies and "traditional" teaching strategies as represented by the statements in STIPS are not necessarily diametric to each other. While the literature in science education, particularly during the sixties, depicts "inquiry teaching" as an instructional strategy opposite in nature to the traditional directive methods, it is quite possible that this distinction is not concretely held by practitioners.

In looking at items comprising the STIPS it appears that some of Edwards' (1957) guidelines for instrument construction were not followed. Some of the statements appear a bit too lengthy, several going beyond 20 words. In some cases, the sentences are compound, as in the following one: "The primary objective of lab experiences should be the development of manipulative skills and ability to follow directions which lead to planned results." Some of the sentences in the STIPS appeared to me as statements that would likely elicit agreement from most teachers regardless of their philosophical views on teaching methodologies. In studies like this one where new instruments have been developed and are being reported, it is useful for the reader to know something concerning the discriminatory power of the individual items. Since the Likert scale does not presuppose equally appearing intervals, it is also of questionable value to report fresh results in terms of the number of scores within each 10-point (or some other) interval. Normative data such as means, standard deviations and standard scores are usually much more useful.

I think this study represents a fruitful direction in science education research. It was well-written, easy to follow and interesting.

The review of literature was relevant to the study and the use of the PCI appeared well-founded. The investigators carefully documented their sample and presented the results in a thorough fashion. They were quick to acknowledge the fact that the sample was both small and nonrandom and that their results were not generalizable to other populations. For these reasons I view this study as a step in the right direction.

The primary weakness of this investigation is that the validity of the STIPS instrument is questionable. Many steps normally taken to assure high quality and valid instrumentation were missing in this study. While by no means the only guide to instrument development, Edwards' guidelines are generally regarded as good basic steps to follow when constructing attitude scales.

This study elucidates the state of the art in attitude research at this time. Researchers in this field need to be reasonably sure that their instruments for measuring attitudes are reliable and valid. This can be fostered by using techniques recommended by measurement specialists in education and social psychology. New instruments should then be correlated with other valid measures as their validity is sought. Once we have valid measures to use in our research, we can then proceed to look for valid relationships between self reporting, paper and pencil instruments and actual behavior.

Looking at the questions initially raised by investigators in this study, it seems reasonable to expect that preference of teaching methodology correlates positively with preference of pupil control methodology. This may not, however, be the case with practicing science teachers at the secondary school level. Most of the earlier research in this area has been done with elementary school teachers—and they may differ in ideology and preference from their secondary school colleagues. In any event, relationships between these two sets of preferences need to be studied more carefully. Valid instruments need to be developed in order to accomplish this. Once relationships between preferences (ideologies) can be ascertained, then relationships between ideologies and actual teacher behavior can be investigated with more validity. This investigation was a step in this direction.

References

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

Hurd, P. D. New Directions in Teaching Secondary School Science. Chicago: Rand McNally, 1969.

Joyce, B. and M. Weil. Models of Teaching. Englewood Cliffs, NJ: Prentice-Hall, 1972.

Kerlinger, F. N. "Progressivism and Traditionalism: Basic Factors of Educational Attitudes." Journal of Social Psychology, 48:111-135, 1958.

Kerlinger, F. N. "Attitudes Toward Education and Perceptions of Teacher Characteristics." American Educational Research Journal, 3:159-168, 1966.

Schwab, J. J. Biology Teacher's Handbook. New York: John Wiley, 1963.

Walberg, H. J. "Teacher Personality and Classroom Climate." Psychology in the Schools, 5:163-169, 1968.

Willower, D. J.; T. L. Eidell; and W. K. Hoy. The School and Pupil Control Ideology. Pennsylvania State University Studies (No. 24 revised). University Park: Pennsylvania State University, 1973.

Judge, Joan. "Observational Skills of Children in Montessori and 'Science—A Process Approach' Classes." Journal of Research in Science Teaching, 12(4):407-413, 1975.

Descriptors--Early Childhood Education; Educational Research; *Instruction; Learning; *Observational Learning; *Preschool Education; *Science Education

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

Purpose

The author's stated purpose was to compare preschool children from classes using the Montessori method and S-APA in the process skill of observation.

Rationale

The study was initiated when similarities between the Montessori and S-APA programs were noted in the literature. A survey of the two programs identifying common elements was used by the author to justify comparing student competence on the process of observation.

Research Design and Procedure

The 75 subjects included 25 students enrolled in a Montessori school in Dallas, 25 students in a private kindergarten in San Antonio using S-APA, and 25 students from a private kindergarten in San Antonio using neither of these programs. The author states that the subjects were equated on socio-economic level (upper-middle class), number of years in preschool (two or more) and age level of the children (five or six years).

Using Campbell and Stanley notation (1971), the design of this study may be represented as a static group comparison

X_1 — — — — O_1

X_2 — — — — O_2

X_3 O_3

X_1 Montessori Training

X_2 S-APA Training

X_3 Neither Montessori nor S-APA training

and

$O_1 - O_3$ The Science Process Inventory (1970)

The dependent variable was measured using the Science Process Inventory (SPI). The test consisted of 68 tasks assessing the specific behaviors in the process of observation. The test was administered individually to the subjects by the investigator in an effort to control the variable of different test administrators. Prior to testing, the author played a group game with each class involved in the study, using questions similar to the task questions from the SPI. The game was intended to help eliminate test unfamiliarity. The testing procedure allowed the student to continue through the tasks until he/she had three incorrect responses.

No information on test validity or reliability was reported.

Analysis

Multiple t-tests were used to compare the mean scores on the SPI for the three groups. The alpha level was set at .05.

Findings

Significant differences were found between the Montessori and control groups as well as between the S-APA and control groups. No significant differences were found between the Montessori and the S-APA groups. The author reports that the data provide evidence that the children from S-APA classes acquired observation skills in a period of one year whereas the Montessori classes acquired the same skills over a three-year period. It was also reported that the son of a science teacher was the only child in the control group to go beyond Task 27 to Task 62 on the SPI.

Conclusions

The author concludes that acquisition of observational skills may assist learning in other subject areas and implies that the process of observation affects readiness in these subject areas.

Other conclusions concern the possibility of exchanging teaching methods and materials between S-APA and Montessori in implementing each program.

ABSTRACTOR'S ANALYSIS

Design

This study is an example of ex post facto or causal comparative research. This type of research attempts to determine cause and effect relationships by taking advantage of existing contextual differences among groups or individuals.

In this study, existing groups were chosen because they had experienced a S-APA, Montessori or traditional program. The appeal of this design is in its unobtrusiveness. It makes little or no demand on the subject because the independent variable is not actually manipulated. A caveat is attached to the use of this design. That is: the validity, both internal and external, rests heavily on the efforts of the investigator to equate comparison groups and control extraneous variables. The author provides descriptive information about the sample population and reports efforts to control such important variables as socioeconomic level, number of years in preschool and age level of the children. Confidence in the equivalence of the three groups would have been enhanced if this information had been expanded and broken down by groups. Without randomization of subjects to treatments, the investigator is forced to case-build and convince readers that these groups would have been equivalent had it not been for their enrollment in these programs. The case could have been more strongly stated.

Analysis

Inappropriate procedures were used in analyzing the data. The three null hypotheses ($\mu_1 = \mu_2$, $\mu_1 = \mu_3$, and $\mu_2 = \mu_3$) were tested separately using the t-tests on each pair of means. This procedure takes advantage of chance differences resulting in an inaccurate estimation of the probability of a Type I error.

When more than one t-test is computed, the probability of one or more Type I errors is greater than .05. The actual probability is $p = 1 - (1 - \alpha)^c$ (Winer, 1962, p. 69), as the number of comparisons (c) increases the probabilities based on t-test tabled values become progressively more inaccurate.

In this study, with three groups and three pair-wise comparisons, the probability that at least one of the differences found between the three groups was actually a chance difference is .14 rather than the reported .05.

Analysis of variance should have been employed. ANOVA yields an accurate and known Type I error probability and is more powerful than multiple t-tests when α is held constant (Hopkins, 1978).

Findings

A number of the reported findings and conclusions go beyond the scope of the study. The reported findings on the amount of time required for the acquisition of observation skills by the Montessori group is not substantiated. The study was not designed to answer how long it took this group to acquire observation skills. Had they been tested after the first or second year of Montessori, the results might well have shown an earlier acquisition of the skill.

Inclusion of the occupation of one of the subject's fathers under the heading "findings" was unfortunate. What was meant to be a human interest note becomes confused with the reporting of a cause and effect relationship between the father's position as a science teacher and the student's achievement.

Conclusions

The design of the study and the reported results provide no logical basis for concluding that observational skills may assist learning in other subject areas nor for implying the process of observation affects the readiness in content areas.

The lack of strong conclusions and implications may result in part from the unfocused rationale for the study. No previous research was cited nor was any theoretical framework provided which would support or extend the findings.

A second factor limiting the strength of the findings and conclusions of this study can be found within th- research design itself. Due to inability to control for selection bias, cause and effect interpretations of relationships established using ex post facto research must be considered extremely tenuous.

REFERENCES

Campbell, Donald T. and Julian C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally and Company, 1971.

Hopkins, Kenneth D. and Gene Glass. Basic Statistics for the Behavioral Sciences. New Jersey: Prentice-Hall, Inc., 1978.

"Science Process Inventory." American Association for the Advancement of Science, 1970.

Winer, B. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Putt, Graeme D. "Testing the Mastery Concept of Self-Paced Learning in Physics." American Journal of Physics, 45(5):472-475, 1977.
Descriptors--Autoinstructional Programs; College Science; Educational Research; *Higher Education; *Instruction; *Mastery Learning; *Pacing; *Physics; Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Gerald G. Neufeld, Brandon University.

Purpose

The study was designed to investigate one of the basic assumptions underlying personalized systems of instruction (PSI) such as the Keller Plan—that the amount of instructional time rather than student ability is the important variable in mastery learning.

Rationale

The study grew out of the author's concern that previous studies of personalized systems of instruction focused on evaluations of the approach using measures of student participation, opinion and performance, and neglected to examine the basic assumptions of a mastery learning approach. This study examined the assumption that instructional time, rather than student ability, was the important variable in mastery learning. The author felt that this was best examined part way through a course because, at this stage, the students would be relatively free of time pressures.

Research Design and Procedures

The research design was a modified static-group comparison design. As indicated in Figure 1, the mastery test (M-test) was administered at different points in the course for the experimental and control groups.

Weeks of Instruction		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Experimental	(140-K)	K	K	K	K	K	K	K	O	K	K	K	K	K	K
Control	(140)	L	L	L	L	L	L	L	L	L	L	L	O	L	L

K = 1 week Keller-plan
L = 1 week lectures
O = M-test administration

Figure 1: Experimental Design.

The 619 subjects were taking a 14-week physics course in Newtonian mechanics and introductory relativity at the University of Michigan. Initially there were 518 students in the control group (Physics 140) and 101 in the experimental, Keller-plan group (Physics 140-K). Assignment of students to treatments was non-random. Most of the Keller-plan students had volunteered for the treatment. However, about 20 students were late registrants and had been forced to enroll for Physics 140-K because the lecture sections were filled. Many of these were nominal registrants and dropped the course after one or two sessions. Since assignment to treatment was non-random, the two groups were compared using SAT scores and high school averages to determine whether the groups were comparable. The author stated that the differences of one-third and one-seventh of a standard deviation, respectively, both in favor of the Keller-plan class were not considered significant for the purposes of the study.

The mastery test instrument (M-test) used in the study was an author-prepared, 16-item, multiple-choice test on Newtonian mechanics. There was one item for each of the 16 units of the Physics 140-K course that the fastest student had completed when the test was given. One item was deleted from the test for the Physics 140 students because that topic had been deleted from the lectures. No information is provided regarding the validity or reliability of the test.

The Physics 140 students were taught by one instructor in two lecture sections of about equal size. These students attended two one-hour lectures and two one-hour recitations per week. Their final grades were based on three one-hour progress examinations and one two-hour final examination.

The Physics 140-K course consisted of 15 compulsory units and 6 optional units. The students attended two two-hour sessions per week. Twenty-four Physics 140/140-K students served as tutors. After a student completed all 17 units on Newtonian mechanics, he/she was given a review test. Final grades were based on accumulated points: 1 point per unit passed, 10 points for the review test, and 20 points for the final two-hour multiple choice test. To minimize the need for memorizing, these students were given a crib sheet listing all the relevant formulas.

Physics 140-K students took the M-test during the eighth week of the semester. They were told to attempt only those questions that related to the units that they had already completed. There was no time limit on the test but they indicated on their test papers about how much time they spent completing it.

Physics 140 students took the M-test during the twelfth week of the semester. In one section students were given 40 minutes to complete the test and asked to attempt all the questions. In the other section students were given 20 minutes to complete it and asked to answer only those questions they felt competent to attempt. Despite these instructions most students attempted all 15 questions. Since the expected spread did not appear for the 20-minute time limit section, only the results for the 40-minute time limit section are reported.

Since the M-test was not part of the formal testing program for either Physics 140 or 140-K, it was administered only to those students who volunteered to take it. The number of test takers and the total number of students in each group were: 47 out of 88 Keller-plan students; 85 out of 257 in the 40-minute time limit lecture section; and 94 out of 261 in the 20-minute time limit lecture section (results not reported for this section). To determine whether the volunteer test takers in the 40-minute time limit lecture section were typical of all control group students, a comparison was made of the final course grades of the 72 volunteers who signed their test papers and received a final grade and the 469 (out of 518) Physics-140 course registrants who received a final grade. The average GPA's were very similar (2.90 and 2.89, respectively) so the author felt the volunteer test takers were typical.

Data Analysis. The M-tests were scored by the author. Mean scores and standard deviations were calculated for the Physics 140 and 140-K sections. The mean scores were then divided by the number of weeks of instruction prior to the test administration (7.5 for 140-K and 11.0 for 140) to determine an average score per week of instruction. No statistical comparisons of mean scores or average score per week are presented.

Student mastery levels were also determined. These were found by comparing the number of test items attempted (the number of units passed for Physics 140-K students) and the student's M-test score. A student who attempted four items and obtained a total score of four was rated as having perfect mastery. A student who attempted 6 items and obtained a score of 5 was rated as having 1 incorrect or obtaining 83 percent correct.

Findings

The mean M-test score for the Physics 140 students was considerably higher than that for the 140-K students (P140, $\overline{X} = 10.18$, $\sigma = 2.33$; P140-K, $\overline{X} = 7.79$, $\sigma = 2.54$). However, when these mean scores were divided by the number of weeks of instruction prior to testing (P140, $t = 11.0$; P140-K, $t = 7.5$), the average score per week was higher for the Physics 140-K students (P140, AS/W = 0.93; P140-K, AS/W = 1.04).

A comparison of mastery levels on the M-test indicated very large differences in favor of the Physics 140-K students. The correlation between mastery levels on the M-test and the final course grades was relatively high for the Physics 140 students (0.71) and quite low for the Physics 140-K students (0.19).

Interpretations

The author feels that the results of this study indicate that time is indeed the relevant variable determining mastery in an introductory physics course. This would imply that Keller-plan courses should be truly self-paced and not have a final completion date so that all students will accrue maximum benefit by progressing as fast or as slow as they can master the material.

ABSTRACTOR'S ANALYSIS

The use of personalized systems of instruction such as the Keller Plan is not widespread but they have become increasingly popular due to the greater emphasis on competency-based education and accountability during the last decade. The ideas and techniques that form the basis for these instructional systems are not recent. Self-paced, individualized instruction formed the basis of the Pueblo Plan (Search, 1894), the Dalton Laboratory Plan (Parkhurst, 1922), and the Winnetka Plan (Washburne and Marland, 1963).

The development of a conceptual model of school learning by John B. Carroll (1963) tied together many ideas about individualized instruction and provided a theoretical framework from which to work. This model predicts that the degree of student learning is a function of: the instructional time allowed, the quality of instruction, and an individual student's perserverance, aptitude for the subject, and ability to understand instruction. Bloom (1968) transformed Carroll's conceptual model into a working model for mastery learning.

Bloom (1971) indicates that if students' aptitudes for a subject are normally distributed and all students receive the same quality and quantity of instruction, their achievement will be normally distributed and there will be a high correlation (0.70 or higher) between aptitude and achievement. However, if these same students were provided with the kind, quality, and quantity of instruction suited to each learner's needs, the majority of students would achieve subject mastery and the correlation between aptitude and achievement should approach zero.

Both Carroll's conceptual model and Bloom's working model of mastery learning have generated a great deal of related research. This research has been the subject of several reviews including Block (1971), Bloom (1976), and Dolan (1977/78). The premise that instructional time, rather than student ability, is the important variable in mastery learning has been investigated by numerous researchers including: Airasion (1967), Atkinson (1968), Behr (1967), Block (1970), Cronbach and Snow (1969), Kim (1968), Sjogren (1967), and Yaeger and Kissel (1969). In general, the research has tended to support this premise. Unfortunately, this study does not appear to relate to, let alone build on, any of this previous work. In fact, the author fails to even mention Carroll's or Bloom's models that serve as the conceptual framework for the study.

Although the research design chosen for the study (a static-group comparison design) is a relatively weak design, it is often the only possible design that can be used when investigating educational questions. What the design lacks in rigor is often more than compensated for by the fact that the research is conducted in a "real-world" classroom setting.

The validity of the study is seriously weakened by a number of factors including: the failure to control or measure important variables, the use of volunteer subjects, the choice of testing instrument, the administration of the test, and the inappropriateness of some of the data analysis procedures.

Carroll's and Bloom's models of school learning indicate that the degree of student learning is a function of several variables including: the quality of instruction, the student's perserverance, his aptitude for the subject, and his ability to understand instruction. It appears that these variables were not controlled, or even measured, in this study. In fact, it appears that even the content of the two courses was not closely controlled because one test item had to be deleted because that topic had not been covered during the lectures.

Many educational studies have to be conducted in situations where random assignment to treatments is not possible and, as a result, have to make use of volunteer subjects. In this study, however, there are three levels of volunteering. Students initially volunteered for the treatment, they then volunteered to write the mastery test, and then volunteered to write their names on the test. Despite the author's assurances that the volunteers were typical of the entire groups, this reviewer still feels uneasy about the validity of the findings. This is reinforced by the author's statement that a difference of one-third of a standard deviation in the mean SAT scores of the two groups in favor of the Keller class was not considered significant for the purposes of this study.

The use of a non-standard testing instrument of unknown reliability and validity further weakens the study. When researchers do have to invent a new test, they should provide some validity and reliability data so that a reader has some means of judging the suitability of the testing instrument.

The method of administering the test raises a number of questions. Why were the tests given at different times during the semester? Why were students in the control group given a 20-minute time limit while the experimental subjects were given no time limit? Why were the control subjects told to attempt all the questions while the experimental subjects were told only to attempt those questions that related to units they had already passed? These differences in test administration must have had significant effects on the findings.

The author's definition and use of "average score per week" and "mastery level" as indicators of student performance seem inappropriate. The use of the average score per week indicator assumes that the test scores form a ratio scale and that learning occurs as a linear function of time. In addition, this indicator is very sensitive to such extraneous factors as how rapidly a lecturer "covers" the course content. The way in which "mastery level" was defined (items attempted-number correct) and the way the test was administered (control-"attempt all questions"; experimental-"attempt only those questions relating to the units you have passed") appear to have seriously biased the results in favor of the Keller plan students and make any comparisons totally meaningless.

Despite the weaknesses of the study, the report was well written. The procedures were clearly described and the results were concisely summarized on a series of graphs.

It is encouraging to see research that attempts to test a theoretical model of learning in a "real-world" science education context. All too often science education research has focused on practical problems and had little or no relation to learning theories or models. We must never forget the old maxim—"There is nothing so practical as a good theory." A careful study of Bloom's (1976) recent extension and elaboration of the mastery learning model should provide science education researchers with many "practical" research ideas.

REFERENCES

Airasion, Peter W. "An Application of a Modified Version of John Carroll's Model of School Learning." Unpublished master's thesis, University of Chicago, 1967.

Atkinson, Richard C. "Computer-based Instruction in Initial Reading." Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton: Educational Testing Service, 55-67, 1968.

Behr, Merlyn J. "A Study of Interactions Between Structure of Intellect Factors and Two Methods of Presenting Concepts of Modulus Seven Arithmetic." Unpublished doctoral dissertation, Florida State University, 1967.

Block, James H. "The Effects of Various Levels of Performance on Selected Cognitive, Affective, and Time Variables." Unpublished doctoral dissertation, University of Chicago, 1970.

Block, James E. (ed.). Mastery Learning: Theory and Practice. New York: Holt, Rinehart and Winston, 1971.

Bloom, Benjamin S. "Learning for Mastery," UCLA-CSEIP Evaluation Comment, 1(2), 1968.

Bloom, Benjamin S. "Mastery Learning," Mastery Learning: Theory and Practice, James H. Block (ed.). New York: Holt, Rinehart and Winston, 47-63, 1971.

Bloom, Benjamin S. Human Characteristics and School Learning. New York: McGraw Hill, 1976.

Carroll, John B. "A Model of School Learning." Teachers College Record, 64:723-733, 1963.

Dolan, Lawrence. "The Status of Mastery Learning Research and Practice." Administrator's Notebook, 26(3), 1977/78.

Kim, Hogwon. "Learning Rates, Aptitudes, and Achievements." Unpublished doctoral dissertation, University of Chicago, 1968.

Parkhurst, Helen. Education on the Dalton Plan. New York: Dutton, 1922.

Search, Preston. "The Pueblo Plan of Individual Teaching." Educational Review, 8:84-85, June, 1894.

Sjogren, Douglas D. "Achievement as a Function of Study Time." American Educational Research Journal, 4:337-344, 1967.

Washburne, Carleton and Sidney P. Marland, Jr. Winnetka: The History and Significance of an Educational Experiment. Englewood Cliffs, N.J.: Prentice-Hall, 1963.

Yeager, John L. and Mary Ann Kissel. An Investigation of the Relationship Between Selected Student Characteristics and Time Required to Achieve Unit Mastery, Working Paper No. 46. Pittsburgh: Learning Research and Development Center, University of Pittsburgh.

Schmitt, Robert M. and David L. Groves. "A Comparison Between Educational Approaches to Teaching Forestry and Tree Identification in a Resident Camp Setting." Science Education, 60(4):485-491, 1976.

Descriptors--*Educational Research; Forestry; *Inquiry Training; Instruction; Science Education; Secondary Education; *Secondary School Science; *Teaching Methods; *Trees

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Lynn W. Glass, Iowa State University.

Purpose

This research was designed to investigate the effectiveness of two methods of teaching forestry concepts in a resident 4-H camp setting. The authors tested the null hypothesis that there were no differences in knowledge gained about forestry by adolescents using the inquiry-process approach versus the lecture-demonstration approach when subjects were stratified by sex and age.

Rationale

It is recognized that adolescent learning takes place in many arenas; these usually can be placed into the dichotomy of formal educational programs and informal educational programs. The reason for success in any educational program is the degree of coordination existing between the given program and prior relevant learning experiences. Stated another way, programs that build upon prior knowledge and skills possessed by the learner will have greater success than programs that do not. Nature study, especially the study of trees and use of keys, is one such area that is included in both formal and informal educational programs and can benefit if the two settings can be coordinated. Traditionally, nature study in a resident camp setting has been taught as a walking-lecture through a natural area. A major difficulty encountered with such a program has been the diversity of backgrounds found within a class.

This research contributes to a large body of existing knowledge about lecture-demonstration and inquiry-process approaches to teaching. The innovative aspect of this research is that it investigates student growth on a "school-type" topic in a resident camp setting.

Research Design and Procedure

Seventy-four boys and girls (ages 9 through 14) who chose to participate in the camp nature study class were involved in this study. An additional 32 boys and girls were selected to serve as control subjects. The two treatments used for teaching forestry and tree identification were the inquiry-process approach and the lecture-demonstration approach. Each instructional approach was used on a separate week and was selected randomly for use on a given week.

With both instructional groups, day one consisted of a pretest followed by instruction. Days two and three were devoted to the instructional treatment, and day four was used for the posttest. All instruction was conducted by the same teacher. A lesson plan to insure that both treatment groups covered the same concepts was used. The pretesting and posttesting were accomplished with a 23-item matching and true-false test in the areas of forestry, tree characteristics, and sight identification of trees. The test was constructed by the researcher and was reported to have a KR-20 index of reliability of 0.73.

The sample was stratified by sex and by age, with 9 to 11 year olds forming one group and 12 to 14 year olds forming the second group. Data are reported as mean group scores. A pretest-posttest control group design was used and can be diagrammatically represented as thus:

$$\begin{array}{cccc} R & O & X_1 & O \\ R & O & X_1 & O \\ R & O & & O \end{array}$$

where i represents an inquiry-process approach and 1 represents a lecture-demonstration approach. A multivariate analysis of variance with a Duncan Multiple Range test to check for differences among all means was used to analyze the data.

Findings

Analysis of the data suggested that adolescents in the age group 9 to 11 responded better to a lecture-demonstration approach, while adolescents in the age group 12 to 14 responded better to an inquiry process approach. These data can be summarized as thus:

	AGE			
	9-11		12-14	
Treatment	Pre	Post	Pre	Post
Lecture-demonstration				
mean	13.2	16.4	14.1	13.6
n	22	22	19	19
sd	4.0	4.1	2.8	3.9
Inquiry-process				
mean	12.5	12.7	13.4	16.1
n	16	16	17	17
sd	3.2	3.9	3.9	3.9

No statistically significant sex differences were reported.

Interpretations

The authors interpret their results to suggest that different instructional methods should be used with different age groups to obtain greater gains in knowledge about forestry and tree identification in a resident camp setting. The traditional approach of lecture-demonstration used in camping to teach nature study works best with the younger age group while the inquiry-process approach works best with the older age group.

ABSTRACTOR'S ANALYSIS

This reviewer believes that there is a need for more research in all areas of informal education. The study reported herein is one such example. This study, undoubtedly, will be considered much more valuable by those persons charged with the responsibility for delivering nature-study programs in a resident camp setting than it will be by those who are attempting to build a model on human learning. Practitioners should find that the article can be read and interpreted easily.

It is difficult to ascertain how the two experimental treatments differed. Statements in the inquiry-process section such as: "...the instructor structured the discussion so as to stimulate additional questions to provide a dialogue" make it difficult to see how this approach differed from the lecture-demonstration approach. Further complicating the matter is the fact that the treatment period was less than three days long. It is hard to be assured that differences in the outcome measure are attributable to the different instructional procedures when students have been exposed to such a small amount of instruction.

When such a short experimental time is to be used, researchers must guard against main effects of pretesting and the interaction of pretesting and the experimental treatment. A Solomon Four-Group Design would be a better research design for this study. Diagrammatically, the design could be represented for this study as thus:

R	O	X_i	O
R	O	X_1	O
R	O		O
R		X_i	O
R		X_1	O
R			O

where i represents an inquiry-process treatment and 1 represents a lecture-demonstration treatment. This design could be utilized with only a moderate increase in the size of the study population.

The authors, in the first paragraph of their article, indicate that a key to studying interrelationships between formal and informal educational programs is identifying instructional elements that provide continuity between the two programs. It would have improved their study greatly if they would have attempted to determine the nature of the previous science program experienced by each subject. Since the age categories selected for this study broke between elementary school and junior high school, knowing the nature of previous science instruction may have provided valuable insight in interpreting the findings of the study.

Wilson, John T. and John J. Koran, Jr. "Effects of Generating Hunches on Subsequent Search Activity When Learning by Inquiry." Journal of Research in Science Teaching, 13(6):479-488, 1976.
Descriptors--*Educational Research; *Elementary Education; Elementary School Science; *Inquiry Training; *Instruction; Learning; *Problem Solving; Science Education

Expanded Abstract and Analysis Prepared Especially for I.S.E. by Jerry G. Horn, Kansas State University.

Purpose

The object of this study was to investigate the effects that generating hunches in symbolic form may have on subsequent search behavior occurring during an inquiry activity. The following questions were of primary concern in this study:

(1) How does hunch generation affect the learner's discrimination and selection of procedures when searching for plausible solutions?

(2) What kind of relationship exists between the quality of the solutions produced?

Rationale

Inquiry, as a popular teaching style, has been justified by a number of researchers, including Suchman and Bruner. Among the claims for inquiry are increased applications of scientific techniques and findings, enlightened attitude toward science, increase of intellectual potency, shift from extrinsic to intrinsic motivation and a means to individualize instruction. Yet there remains the lack of a common description which indicates what constitutes inquiry and what does not.

Whether or not inquiry is an effective model for classroom instruction has been an issue in many research efforts. However, too many of these efforts intended to demonstrate only that inquiry *per se* is better than others. Little effort has been made to identify specific instructional elements or to describe the nature of the learner's responses, either of which could be considered to be unique to inquiry. Wilson's Process Model of Inquiry, as reported by John T. Wilson in 1974, is one paradigm which has identified several sets of processes as key elements of inquiry. As part of the search behavior by learners, they perform some process which permits them to gain information, but they may also tentatively identify possible causes and explanations for the observed situations. Wilson identified this latter activity as "hunch generation." It has been hypothesized in reports by Hadamand, Shockley, Long and McDonald that generating hunches has a facilitative effect on both the search activity and the construction of plausible solutions. Conditions, such as generating hunches in symbolic form, may encourage the performance of relevant processes and should also positively influence the search activity and the quality of solutions posed.

Research Design and Procedure

The sample consisted of 45 children, ages nine through eleven, from middle-income families, attending an elementary school in Austin, Texas, where science was a regular activity in the program. The ratio of boys to girls in the sample was about 1:1.

The experimental materials consisted of a discrepant event, a set of 15 investigations and corresponding equipment, and a criterion measure. The discrepant event consisted of four blocks of wood, each painted a different color with an equal length of string and weight attached. The four blocks were lined up at a "starting line" and allowed to slide along a table as if in a race. One block always won and another always lost. The key variable was the sliding surfaces on the blocks, which the subjects were never allowed to see.

There were 15 investigations constructed so that each procedure investigated one discrete variable. Eight of the investigations were "relevant" as they dealt with a property or variable that was dealt with in the original discrepant event. An example of these investigations in question form is found below.

"Does the weight of the block help it win the race?"

The criterion measure presented a drawing of a block of wood with a string and weight attached. Instructions directed the subjects to make the block the winner. Each modification which would make the block slide faster was awarded +1 point, changes which would slow the block received -1 point, and changes which would produce no effect were given 0 points.

Additionally, three versions of printed materials were used. One page common to all versions presented brief statements about the 15 investigations, and subjects were asked to identify procedures as "useful" or "not useful" for finding a solution to the situation; "why the winner won and the loser lost." Three forms of a second response included one asking subjects to write hunches about "why the winner won and the loser lost," one that directed subjects to read a set of hunches, and another that directed subjects to continue to the next page. Booklets containing written materials were assembled to coincide with the three different treatment conditions (wrote hunches, read hunches, no hunches).

The subjects were randomly assigned to one of three treatment groups. The subjects completed the introductory materials (personal information page and an introductory explanation about hunches) and then viewed the discrepant event. The experimental subjects either wrote or read hunches and the control group performed no hunch generation. All subjects then identified which investigation procedures were "useful" or "not useful." Each subject was then given additional investigation material and equipment for performing those procedures he or she identified as useful.

The experimental design followed the structure of a posttest-only control-group design, with the first dependent measure requiring subjects to identify procedures as "useful" or "not useful." The second dependent measure was the criterion test in which subjects modified a drawing of a block of wood which in turn was rated according to a prescribed protocol. The scoring reliability was .95 using an analysis of variance technique and an estimate of content validity, using Gullickson's procedure, was .97.

Findings

Mean scores and standard deviations for dependent variables were reported. An inspection of the means of the three treatment conditions identified potential differences between group differences. Using analysis of variance and Tukey HSD techniques, it was found that subjects who

(1) wrote hunches selected significantly ($p < .05$) fewer procedures as useful, spent significantly ($p < .05$) more time performing the selected procedures, and exhibited a significantly ($p < .01$) higher quality of solution on the criterion test.

(2) wrote hunches selected significantly ($p < .05$) fewer relevant procedures. No significance between group differences was found for the percent of selected procedures that were relevant. No significant differences were found to occur between subjects who read hunches and subjects in the control group who neither read or wrote hunches.

Interpretations

Generally, the findings supported the expectations that generating hunches would positively influence the quality of the solution formulated; learners who generated hunches generally seemed to benefit in terms of the amount learned during the search behavior. The findings of this study suggest that differences probably exist in the mental processing activity associated with the generation of hunches in symbolic form, in this case a verbal-written form.

Inquiry incorporates a number of instructional elements, of which only a few have been shown to be effective means to facilitate learning. The mental processing activity performed by the learner during inquiry, as in other types of instruction, is sensitive to external elements within the instructional situation. These elements can be manipulated in a way that promotes appropriate mental processing, relative to defined instructional outcomes. Relative to hunch generation, teachers may elicit various observations, inferences, and predictions to facilitate hunch generation.

The results disclosed here generally support the hypothesis proposed by Wilson in his Process Model of Inquiry in that generating hunches facilitates the processing activity required in the performance of scientific inquiry.

ABSTRACTOR'S ANALYSIS

This study by Wilson and Koran is obviously one part of a rather extensive effort to develop a theoretical base for both inquiry teaching and an explanation of the mental processing by learners as they encounter new opportunities for learning. Evidence of this conclusion is clearly found in the text of the research report and the bibliographical listing. The impetus for the study and the consequential report of the results reflect serious consideration for previous work by the authors as well as other researchers.

In terms of the research design, there are some weaknesses that must be considered. The subjects for the study appear to be from an intact group found in one elementary school in Austin, Texas. While it was reported that "personal information" was obtained from the subjects, there was never any mention of these data at any other point in the report. For purposes of generalizing the results of the study and giving greater credibility for the developing theory, a more fully defined description of the sample would be very beneficial. While the point was made that the children participated in the experiments as if it were one of their regular science activities, one may question the accuracy of this statement since there were booklets and other activities related to the experiment that were probably very different from a usual class activity. It is also not clear whether there was discussion among the teacher, researcher and students during the course of the experiment. One presumes that the activities of the experiments were presented to a group rather than in individual settings.

As to the report itself, apparently the authors did not intend it for teachers or even program specialists at the local level due to the complexity of the report. As mentioned earlier, the report builds on a developing theory, but it lends little support for the practitioner as he/she deals with elementary school age children on a daily basis. The previous statement is not intended as a criticism but is merely a notation of the rather narrow audience that could make application of the findings.

While the authors suggest that immediate applications of the findings should be made with caution, I fail to see the risk in such an effort, since they clearly point out in an earlier section that "inquiry is one of the most popular, widely known teaching styles in education," and that "hunches are tentative ideas that serve to direct the activity of both empirical and conceptual inquiry."

Certainly, this study is one of the stronger efforts in science education to synthesize theoretical models and expand our knowledge about instructional techniques. Generating hunches, as an instructional technique, may be novel to the elementary age student, and it could serve as a positive motivating force for individual exploration and inquiry. This writer perceives lack of motivation by students as one of the most serious obstacles to learning. Research that addresses motivation and its relationship to novel approaches, such as hunch generation, would be beneficial to teachers, curriculum planners/developers and teacher educators.

Wollman, W. and A. Lawson. "The Influence of Instruction on Proportional Reasoning in Seventh Graders." Journal of Research in Science Teaching, 15(3):227-232, 1978.

Descriptors--*Abstract Reasoning; *Cognitive Processes; Educational Research; Elementary Education; *Instruction; *Junior High Schools; Mathematics Education; Science Education; *Secondary School Science; *Teaching Methods

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

Purpose

The purpose of this research was to compare two methods of teaching proportional reasoning to seventh grade students.

Rationale

This is one in a series of papers which report the application of Piagetian theory to problems that bear directly or indirectly on science education. In this case the problem attacked was the teaching and learning of a mental operation that is necessary for the understanding of many science concepts.

Research Design and Procedure

Twenty-eight seventh grade students, from two mathematics classes, were pretested on three conservation tasks and assigned to three levels on the basis of test results. They were then randomly assigned within levels to two training groups, forming two comparable groups of 14 students each.

In the "active" group each subject met with an experimenter for four 30-40 minute sessions over a period of approximately two weeks. Manipulable materials were used at each session, starting with arrangements of rods in ratios of 1:2, 2:3, etc. At each session symbolic notion was introduced after problems had been solved by use of the manipulable materials.

In the "verbal" group, instruction (or training) was also carried out individually in four sessions. The work was based on a standard textbook and included reading, discussion with the instructor, and completion of the homework exercises in the book. Topics covered were (a) comparing sets, (b) comparing ratios, (c) computing with ratios, (d) ratio, proportion, and scale drawings. Algorithms were presented and followed by applications.

A posttest followed immediately after completion of training. The posttest consisted of problems as follows: (1) an individually administered task, based on apparatus (Disks) not previously used, (2) a multiple choice ratio problem, (3) the well-known Mr. Tall-Mr. Short problem, and (4) four written problems about the work done by a machine.

A delayed posttest was administered one month later. This consisted of two tasks: (1) Mr. Tall-Mr. Short and (2) two written problems similar to the written problems of the immediate posttest.

On the immediate posttest the scores of the two groups were approximately the same on the Disks problem and the multiple choice ratio problem. On the Mr. Tall-Mr. Short problem and on the written problems about the machine, the active group scored significantly higher. On the delayed posttest the active group again scored significantly higher ($p < .05$) than the verbal group on both tasks.

Inspection of test results shows that the verbal group scored very low (17 percent) on the first encounter with the Mr. Tall-Mr. Short problem while the active group scored a much higher 75 percent. On second encounter with this problem, one month later, the active group scored even higher (89 percent) but the verbal group mean score jumped up to 56 percent without intervening instruction. On the written problems, in contrast, the scores of both groups dropped. It is interesting that the Mr. Tall-Mr. Short problem apparently engaged the minds of the students and stimulated thinking in a way that the written problems did not.

Interpretations

The authors believe that this study supports the view that instruction that is intended to improve reasoning should parallel the process of internalization of actions by having students work first with materials that model or illustrate the principle to be learned in "concrete, flexible, action-oriented" contexts. They would have students work with symbolic representations only after they have had the opportunity to manipulate materials, to use their own words to describe their actions, and to bring their own mental resources to bear on problem solving.

ABSTRACTOR'S ANALYSIS

This rather modest study was carried out in a straightforward, deceptively simple way. The sample was not large but sufficient to make the point. It is the kind of research that "anyone" can do but few people actually do. Many problems were solved by conducting the training and, presumably most of the testing, on an individual basis, but the use of that method required a large investment of time.

This study does not break new ground but confirms what many educators have been saying for a very long time: that children learn more when they are active participants in the process of abstracting ideas from concrete exemplars. The manipulation of objects should precede the manipulation of symbols.

We can thank the authors for doing their bit to demystify Piaget. Not too long ago there were those who thought that the operations defined and described by Piaget appeared as if by magic and that nothing at all could be done to change a predetermined course of events. While it is true that attempts to teach the concept of conservation have been notably unsuccessful, it is now clear that both concrete and formal operations are susceptible to training under certain conditions. This is important for us to know. We should spread the word at every opportunity that children can be taught in ways that will stimulate the growth of logical thinking. It is not true that schools make no difference or that teachers make no difference. Some teaching methods are better than others. The results presented here support the use of inductive instructional methods which lead students from the specific and concrete to the abstract and general.

What we need now are more studies of this nature to form a foundation for instructional theory and practice. Beyond that we need to translate these ideas into methods that will work in classrooms and to train teachers in their use.