DOCUMENT RESUME

ED 264 136 SE 046 300

AUTHOR Blosser, Patricia E., Ed.; Helgeson, Stanley L., Ed.
TITLE Investigations in Science Education. Volume 11, Number 4.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 85
NOTE 72p.
AVAILABLE FROM Information Reference Center (SMEAC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (U.S. subscription $8.00; $2.75 single copy).
PUB TYPE Collected Works - Serials (022) -- Reports - Research/Technical (143) -- Information Analyses (070)
JOURNAL CIT Investigations in Science Education; v11 n4 1985

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; Chemistry; *Cognitive Development; Elementary Secondary Education; Higher Education; *Problem Solving; *Science Education; *Student Attitudes
IDENTIFIERS *Science Education Research

ABSTRACT

Presented are abstracts and abstractors' analyses of four studies dealing with cogniti-` development, four studies dealing with attitudes, and two studies dealing with problem-solving. The cognitive development studies are: "College Chemistry and Piaget: The Relationship of Aptitude and Achievement Measures" (David Bender and Louis Milakofsky); "Inducing Cognitive Growth in Concrete-Operational Students" (William Thomas and Douglas Grouws); "Effect of Using Analogies on Chemistry Achievement according to Piagetian Level" (Dorothy Gabel and Robert Sherwood); and "The Influence of Content Organization on Students' Cognitive Structure in Thermodynamics" (Marco Moreira and Carlos Santos). The attitude studies are: "Saudi Arabian Students' Chemistry Achievement and Science Attitudes Stemming from Lecture-Demonstration and Small Group Teaching Methods" (H. Harty and N. Al-Faleh); "Relationship between Perceived Levels of Classroom Individualization and Science-Related Attitudes" (Barry Fraser and William Butts); "A Survey of Interest in Science for Participants in a Junior Science and Humanities Symposium" (Jill Wright and Paul Hounshell); and "Do New Science Courses Improve Attitudes toward Science? A Study in Lesotho" (P. Towse), including the author's response to this critique. The problem-solving studies are "Problem-Solving Strategies of Sixth-Grade Students Who are Superior Problem Solvers" (Alan Mandell) and "Proportional Reasoning and Rule-Governed Behavior with the Balance Beam" (D. Maloney). (JN)

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC Clearinghouse for Science, Mathematics, and Environmental Education

INVESTIGATIONS IN
SCIENCE EDUCATION

Volume 11, Number 4, 1985

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: $8.00 per year. Single Copy Price: $2.75.
Add $1.00 for Canadian mailings and $3.00 for foreign mailings.

Volume 11, Number 4, 1985

NOTES FROM THE EDITOR:

The final issue of Volume 11 continues the series of critiques on articles related to cognitive development that was begun in issue three. In this issue four articles relate to cognitive development: an examination of the relationship between a measure of Piagetian development and aptitude and achievement variables (Bender and Milakofsky), the study of an attempt to induce cognitive growth in concrete operational college students (Thomas and Grouws), an investigation of the effect of the use of analogies on chemistry achievement of high school students (Gabel and Sherwood), and the influence of two different content organizations on students' cognitive structure (Moreira and Santos).

The second section of volume four contains critiques of published studies of attitudes: the effect of two teaching methods on students' chemistry achievement and attitudes (Harty and Al-Faleh), the identification of relationships between perceived levels of classroom individualization and science-related attitudes (Fraser and Butts), the development of students' science interests (Wright and Hounshell), and the effect of a new physical science course on attitudes toward science (Towse). A response by Towse to the critique of his article ends this section.

The final section of volume four contains two critiques of articles related to problem solving. One relates to the study of common problem solving behaviors of students classified by their teachers as superior problem solvers (Mandell). The other is an examination of the relationship between college students' proportional reasoning patterns and rule-governed behavior with balance beam problems (Maloney).

Patricia E. Blosser
Editor

Stanley L. Helgeson
Associate Editor

COGNITIVE DEVELOPMENT

Bender, David S. and Louis Milakofsky. "College Chemistry and Piaget: The Relationship of Aptitude and Achievement Measures." Journal of Research in Science Teaching, 19: 205-216, 1982.

Descriptors--*Academic Achievement; *Academic Aptitude; *Chemistry; *Cognitive Development; Cognitive Processes; *College Science; College Students; Developmental Stages; Higher Education; Science Education; Student Characteristics

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

Purpose

The purpose of this work was to investigate the relationship between a measure of Piagetian development (An Inventory of Piaget's Developmental Tasks, the IPDT) and aptitude and achievement variables. It further described the weakest areas of Piagetian development for college students (as measured by the IPDT).

Rationale

This work is located within the Piagetian framework. In this view of cognition, children are assumed to pass through various stages of cognitive functioning. Formal operational devlopment, the highest level, includes the ability to reason abstractly. The authors relate this work to 23 other Piagetian studies, concerning such topics as relationship between Piaget's theories and college science teaching, methods of assessing cognitive maturation, and types of reasoning required in college science courses.

The authors state their belief that previous researchers have demonstrated that many college students are unable to reason abstractly in science courses. Such deficiency is assumed to indicate that those students have not yet passed fully into the formal operational level. It is assumed that constructs such as stages of cognitive maturation exist, are real and measurable, and that their measurement has

pertinence to college science teaching and learning. Citing previous work by one of the authors (Milakofsky), it was accepted that the IPDT is valid, reliable and useful to chemical educators as a measure of cognitive development of students.

Research Design and Procedure

This was a descriptive study. There was no attempt to manipulate instruction or performance. Rather, scores on various performance measures and "aptitude" tests were related and interpreted.

Students. A total of 225 students participated in the study. The students were enrolled in either an introductory/remedial chemistry course (Group 1) or a more advanced "chemical principles" course (Group 2). Group 1 contained 64 students (approximately half male and half female); Group 2 contained 161 students (approximately 10% female). Students were assigned to groups on the basis of "in-house" university-wide chemistry and algebra placement exams. No other information about students was provided.

"Aptitute" and Performance Measures. SAT scores (mathematics and verbal): Scores for all participating students were obtained. Means and standard deviations for Groups 1 and 2 were calculated. SAT's presumably were taken by students in their final year of high school.

Chemistry Placement Exams: No information about these tests was provided, except to claim that these exams tested high school knowledge. In reporting data for Groups 1 and 2, means of standardized T-scores (mean of zero) and standard deviations were reported. These placement tests were presumably taken at some time prior to the beginning of the school term in which data was collected.

IPDT (Inventory of Piaget's Developmental Tasks): This is a 72 item, untimed, multiple-choice, paper-and-pencil inventory. The inventory is divided into 18 subtests (four items in each subtest) grouped into five problem areas: classification, conservation, imagery, proportional reasoning, and relations. Test-retest coefficients and a

split-half reliability coefficient (0.77) were reported. Claims of consistency of the IPDT with other Piagetian research were made. Means and standard deviations were reported for the total score, problem area scores, and individual subtest scores, for both groups. The IPDT was given within the first two weeks of the school term. A sample (64) of the students in the Group 2 retook the IPDT at the end of the school term. No information was given as to how these students were selected for retesting.

Chemistry Grades. End-of-course chemistry grades were reported for Groups 1 and 2 for both laboratory and lecture parts of the courses. Lecture grades were based on "open-ended written quizzes and examinations." Group 1's test questions were 30% quantitative and 70% qualitative; Group 2's exams were 70% quantitative and 30% qualitative. Lab grades were based on lab reports and (for Group 2) some quizzes. The numbers of students in each group who received various grades (A to F) were not reported.

Findings

1) Comparison of Groups 1 and 2 on standardized tests and the IPDT.

   The two groups were compared on these "aptitude" and prior knowledge measures. Significant differences ($p < .0001$) were found on the SAT (mathematics), SAT (verbal) and Chemistry Placement exam. The groups also differed on the IPDT ($p < .01$).

2) Correlation of IPDT and other aptitude or prior knowledge measures.

   Pearson product-moment correlation coefficients were calculated between IPDT and Chemistry Placement, SAT (math) and SAT (verbal). Significant relationships were found ($p < .001$) between IPDT and each of the other measures.

3) Grades in courses vs average IPDT scores.

Course grades for students in Groups 1 and 2 were plotted vs the average IPDT scores for students with those grades. "A" students in both groups had higher average IPDT scores than others in their groups. Total IPDT scores correlated significantly ($p < .05$) with lecture and laboratory grades for Group 2 but not for Group 1. There was only one significant correlation between problem areas and grades for Group 1 (Proportional Reasoning with Lecture Grade); for Group 2, significant correlations were noted for Classification vs Lecture and Lab Grades, Proportional Reasoning vs Lecture and Lab Grades and Relations vs Lab Grades.

4) Comparison of Groups 1 and 2 on Problem Areas and subtests within areas on the IPDT.

Significant differences in group means were found for the areas of Imagery, Conservation and Proportional Reasoning, but not for Relations and Classification. Significant differences were attributed to differences in specific subtests.

5) Ranking of Problem Areas according to difficulty.

The order of difficulty of the five areas was determined. The six most difficult subtests were identified and described. The percentages of students in each group who missed each of these subtest items were calculated. When some students in Group 2 were retested at the end of the term, they showed significant improvement in only one of these six difficult items.

Interpretations

The authors' claim that their data may be interpreted as follows:

1) The significant relationship between the IPDT and the SATs indicate that the IPDT, a test of cognitive functioning, has "some" validity as a predictor of college grades.
2) Correlations between IPDT and course grades are lower than reported in previous studies.
3) Classification and proportional reasoning appear to make a difference in performance in higher level chemistry courses (lecture and lab). Relations and imagery problem areas appear to be mastered by students and not to influence course grades.
4) Poor performance in specific subtests and in specific problem areas can be related to the inability to achieve in college chemistry. Examples were given of chemistry problems which might use cognitive functions as measured on various subtests.

ABSTRACTOR'S ANALYSIS

One point which should be considered in the analysis of this article concerns the concept of cognitive development and the stages through which Piagetians believe students progress. It is not clear from this article what the authors believe are the characteristics of cognitive functioning at the concrete and the formal operational levels. We are given these labels, but not an explanation of what is meant by them. We are given some examples of test questions from the IPDT which the authors assume can discriminate students who have reached the formal level from those who have not. But other than a glimpse at these tasks and phrases such as "the ability to apply abstract reasoning," the reader must construct his or her own description of what the authors are measuring.

Among the questions which might be asked about the developmental stages of students are the following. Is the cognitive functioning level invariant over all academic tasks? That is, is a person globally concrete or formal operational when dealing with all subject matter? If not, how can one hope to measure it using a non-content specific test? What role does content knowledge and experience play in one's ability to demonstrate "formalness" on a paper-and-pencil, multiple-choice test? Moreover, the "developmental" view of Piaget implies that students progress through stages as they mature cognitively. Why then were so few changes noted with time in this study? Were there no changes because the students didn't mature or because the test couldn't measure the changes which did occur?

A second question relates to the purpose of the study. This was not an experiment in which there was an intent to manipulate some features of the educational process. Rather, it sought to measure in various ways the functioning of students and to seek relationships among those measures. But for what purpose? For example, the IPDT is considered as an "aptitude" measure, along with the SAT math and verbal scores and the in-house departmental chemistry placement exams. There were positive, significant correlations among all these measures (albeit at differing levels). To what use can this information be put? The authors state in their Discussion section that since the IPDT is significantly correlated with the SATs, it does have some validity as a predictor of college performance. The reasoning appears to be: SATs predict college performance for aggregated students, and therefore, tests which correlate with the SATs should also predict college performance. Is this a true statement? Is it only true to the extent to which all correlated measures are measuring the same attributes? Amassing information about correlations of one measure with others allows you to do what?

With respect to the IPDT itself, one can raise a number of questions. For example, the mean difference between Groups 1 and 2 in IPDT scores is 2.5 questions out of 72. While statistically significant, does that have any educational significance? The students

were presumably assigned to groups on the basis of content knowledge as measured by chemistry and math placement exams. Should we expect that the IPDT, which purports to measure cognitive development, tells us anything about the entering content knowledge of students?

Course grades, used as a performance measure, are gross, overall descriptions of student functioning in a course. To do well in a course, one needs a great many intellectual skills, personal characteristics (such as the willingness to do the work and persistence) and some critical prior knowledge. Presumably, one important characteristic for success is the ability to reason abstractly and do whatever "formal operational" students are capable of doing. This article reported few statistically significant correlations between the overall IPDT scores, subtests, and course grades. Should we be surprised by this? Since we know little about the courses themselves, and we don't know exactly what student characteristics the IPDT measures, even if there were a correlation, it would be hard to know on what that correlation rested.

What is the purpose of identifying and analyzing the hardest subtests of the IPDT? The students scored reasonably well on all problem areas, with some specific subtests being relatively more difficult than others. Yet, we do not know the sources of the errors. Are errors the result of bad test questions, incomplete knowledge on the part of students, or deficiencies in mental maturation?

For example, some of the questions which are cited seem rather ambiguous. (Try the question in Figure 3. Unless there are some missing instructions, it isn't clear how that question should be answered.) Moreover, some of the other questions seem to depend on specific pieces of information rather than on some underlying ability to perform abstract or hypothetical thinking. For instance, if a student has a clear idea of conservation of quantity, weight and volume, but misses one or two out of four questions about conservation of distance, does this mean that the student is deficient in "conservation" skills? Which leads us back to an earlier comment in this section: what exactly is the IPDT measuring and what relationship does it have to success in general chemistry courses?

The tasks of the IPDT seem to relate to some of the tasks in general chemistry courses. But the examples given in the discussion section are offered without specifics. For example, the probability subtest in the section in proportional reasoning is linked to the comprehension of statistical chemistry problems. Are statistical questions common in beginning chemistry courses? (Not in my experience.) Moreover, it is not clear how the ability to answer the IPDT questions translates into effective problem solving in chemistry courses. How will poor performance on conservation of volume affect density problems?

The final section of the paper lists some of the critical questions about teaching and learning in school science. The authors correctly point out the need for more information on the "benefits of programs directed toward improving general cognitive skills versus those concentrating on preparation for learning specific subject matter."

Thomas, William E. and Douglas A. Grouws. "Inducing Cognitive Growth in Concrete-Operational Students." School Science and Mathematics, 84 (3): 233-242, 1984.

Descriptors--*Cognitive Development; *College Science; College Students; *Developmental Stages; *Educational Games; Higher Education; *Problem Solving; Science Education; *Science Instruction

Expanded abstract and analysis prepared especially for I.S.E. by Anton E. Lawson, Arizona State University.

Purpose

The purpose of this study was to test the following predictions: Subjects (Ss) who play the commercially available game of "Master Mind" including verbal interaction with an instructor (the Structured-Interaction treatment SI) and Ss who play Master Mind with no instructor (the Neutral-Interaction treatment NI) will show greater cognitive growth than Ss who do not play Master Mind.

Rationale

The authors recognize the problem of many college freshman (20-40%) who do not reason formally yet face advanced courses that may require formal reasoning for success. Thus, they sought to test the utility of the game of Master Mind as a tool to stimulate intellectual development of college Ss in a short period of time. They cite the work of a number of psychologists and educators in support of their belief that experience-based instruction with provisions for reflection upon those experiences will stimulate cognitive growth (e.g., Ausubel, 1968; Bruner, 1963: Inhelder and Piaget, 1985, Kurtz and Karplus, 1979; Lawson and Wollman, 1976).

Master Mind is viewed as a game that may stimulate cognitive growth, particularly if played with an instructor who provides suggestions to stimulate the players' thinking about the game's

strategies, because game success requires use of formal reasoning strategies of hypothesis generation and test via combinational analysis and the control of variables.

Research Design and Procedure

Subjects. A group of 185 college students (freshmen through senior level plus one graduate student) enrolled in general-education science classes were given a pretest of cognitive development. Thirty-nine of the 76 students, who scored in the concrete-operational range, agreed to serve as subjects.

Instrumentation. The test of formal reasoning consisted of four items refined by Renner, Prickett, and Renner (1977). The items were titled the Rock and Scale (combinational analysis), the Frogs (proportional reasoning), the Worms (separation and control of variables), and the Geranium (spearation and control of variables). A composite score of from 4 to 26 was obtained for each S. Ss scoring 13 or less, with no item score in the formal range, were considered concrete-operational. A test-retest reliability coefficient of 0.88 was obtained from the testing and retesting (5 week interval) of a separate group of 70 Ss.

The experiment used the pretest/posttest control group design with the random assignment of Ss into two treatment groups (SI or NI) and the control group. Ss meet with the instructor individually for one hour once a week for four weeks. During the meeting, treatment group Ss played Master Mind with the instructor, while the control Ss played checkers. The major difference between the two treatments was that the SI group was given questions designed to cause Ss to reflect on their game strategies (without imposing the instructor's strategies) while no game related questions or comments were allowed in the NI group.

Findings

The group means on the posttest were: SI group, 16.1; NI group, 14.6; control group, 13.8. These scores represent respective gains of 4.7, 2.7, and 1.4 points over pretest scores. Using pretest score as a covariate an analysis of covariance was conducted which gave a significant omnibus F value ($p < 0.05$). The Newman-Keuls test was used to test pairwise differences. The only significant pairwise difference was between the SI and Control groups. To test for lasting effects of the treatment a delayed posttest was administered to a small subsample of SI and control Ss (10 Ss in all) one month after the posttest. A significant t value of 6.22 (p 0.001) was obtained in favor of the SI group.

Interpretations

The authors interpreted their results as support for the hypothesis that playing the game of Master Mind under the experimental conditions utilizing problem-related interaction can promote cognitive development in concrete-operational college students.

They believe that playing the game was helpful primarily because it required the player to utilize precisely the same abilities required for doing science and for developing cognitively as assessed by the study's instruments. Both Master Mind and science require the ability to collect and interpret data which in turn requires the capacity to separate and control variables. Similarly, data interpretation requires the capacity to imagine all possible combinations of variables (combinatorial analysis).

The authors acknowledge, however, that the game is limited in its usefulness because some science related abilities such as proportional reasoning which are involved in doing science and in developing intellectually are not required by Master Mind.

ABSTRACTOR'S ANALYSIS

I would certainly agree with the basic hypothesis of this study as it (as quoted by the authors) is, at least in part, mine (Lawson, 1979):

> It would seem that for any approach to be effective it would have to involve activities in which learners were in fact called upon to solve problems which required the use of the reasoning strategies in question. In addition it would seem that the approach would have to have provisions in it which required students to reflect upon their problem-solving attempts and to gradually abstract the correct procedures (p.513).

The study finds support for this hypothesis as have previous studies (e.g., Case and Fry, 1973; Kurtz and Karplus, 1979; Lawson and Wollman, 1976; Wollman and Lawson, 1978; Siegler and Liebert, 1975; Shyers and Cox, 1978; Wollman and Chen, 1982). The primary difference between the present study and these studies, however, rests with the authors' assumption that a training period of the sort utilized here (4 hours of rather indirect and limited instruction in a neutral context) can effect meaningful changes in a concrete-operational students' thinking processes that have become entrenched across the whole of childhood and adolescence. Is it realistic to view the modest gains made on an extremely limited and rather primitive set of four problems after only four hours of playing a game (no matter what the intrinsic merits of the game may be) as meaningful gains in intellectual development from the concrete stage to the formal stage? Indeed the authors reported that three of the SI Ss (21%) changed from concrete to formal reasoning and eight (57%) of the SI Ss changed from concrete to transitional reasoning. I have no objection to the claim that performance on the test has improved, but I object to the implication that Ss have indeed moved from the concrete stage to the formal stage or to even a transitional phase following such a brief and limited intervention.

The literature provides evidence for much more modest interpretations of the authors' results. For example, Danner and Day (1977) administered the bending rods, pendulum, and spinning-wheel tasks to 20 Ss in each of three grades (5, 8, and 12) and followed the testing with brief instruction designed only to prompt use of the control of variables strategy rather than induce its development. These prompts caused significant pre and posttest gains which they claimed implied that many of the Ss were, in fact, not concrete-operational in the first place. Rather the pretests, due essentially to their novelty, underestimated actual subject competence. Perhaps the same could be said of the college Ss in the present study. Perhaps the Ss were formal operational and did poorly on the pretest for similar reasons. Perhaps the game of Master Mind simply prompted use of the strategies which were already developed. The present paper would have been strengthened by acknowledgement of this possibility and by inclusion of a discussion of this issue.

A study by Stone and Day (1978) is also relevant. They administered a modified version of the bending rods task twice in succession with prompts, much like those utilized by Danner and Day, to 28 Ss at each of three ages (9, 11, and 13 years). On the basis of initial use of the control of variables strategy and responsiveness to prompts, Ss were classified as spontaneous (initially controlled variables), latent (controlled following prompts) or nonusers (did not control following prompts). No differences were found between spontaneous and latent users on a backward digit span test (a measure of mental capacity) and on a test of the ability to generate second-order operations. Significant differences were found on these tests between these Ss and the nonusers. In other words, it would seem that spontaneous and latent users have formal competence but the nonusers do not. Thus it would seem that the Piagetian tasks identify important differences in reasoning ability but one must take care not to be fooled by initially poor performance. Successful performance implies competence. Initially unsuccessful performance does not imply the converse (i.e., some formal Ss are latent users).

Further, what of the issue of the interrelatedness of reasoning patterns? The authors do acknowledge that their instruction did not include provisions for aspects of formal reasoning such as proportional reasoning. Yet their introduction and results remain framed in terms of general stages of development and wholesale changes in thinking (i.e., "What is needed is a method for stimulating the intellectual development of college students which takes only a short time to administer and which is largely content free" p. 234). Is this a realistic objective? The literature suggests that it is not. Rather the development of formal reasoning seems to occur on a broad front related to performance in a wide variety of contexts (e.g., Bart, 1971; Capon and Kuhn, 1979; Lawson, Nordland, and DeVito, 1975; Lawson and Wollman, 1977; Hardy-Brown, 1979; Keating and Clark, 1980; Lawson, 1982) and seems hindered by highly intractable factors such as restrictive social environment, field dependence, low mental capacity, and perhaps by an impulsive cognitive style (e.g., Douglas and Wong, 1977; Scardamalia, 1977a, 1977b; Lawson and Wollman, 1977; Linn, Pulos, and Gans, 1981; Neimark, 1975; Case 1974). Further, contrary of the notion that a brief period of training in a context free environment can induce stage change, the extent to which training transfers to novel tasks and other aspects of advanced reasoning seems to depend upon the length and diversity of the training. Brief training may promote limited advance, yet only longer and more diverse training (which indeed more closely approximates stimulating conditions in the general social milieu) seems able to effect the wholesale advances of the sort aimed at by the present authors (e.g., Anderson, 1965; Siegler, Liebert, and Liebert, 1973; Lawson and Wollman, 1976; Howe and Mierzwa, 1977; Case and Fry, 1973; Kuhn and Angelev, 1976).

In summary, although the present study appears to have provided evidence of the utility of Master Mind for increasing performance on tasks which require similar reasoning in novel contexts (specific transfer), it has failed to demonstrate that these advances are stage-like advances in intellectual development as generally conceived of by developmental psychologists or that performance increases are indeed due to anything more than the prompting of the use of already developed cognitive abilities.

REFERENCES

Anderson, R. C. "Can First Graders Learn an Advanced Problem-Solving Skill?" Journal of Educational Psychology, 56: 283-249, 1965.

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

Bart, W. M. "The Factor Structure of Formal Operations." British Journal of Educational Psychology, 41: 70-79, 1971.

Bruner, J. The Process of Education. New York: Vintage, 1963.

Capon, N. and D. Kuhn. "Logical Reasoning in the Supermarket: Adult Females' Use of a Proportional Reasoning Strategy in an Everyday Context." Developmental Psychology, 15: 450-452, 1979.

Case, R. "Structures and Strictures, Some Functional Limitations on the Course of Cognitive Growth." Cognitive Psychology, 6: 554-573, 1974.

Case, R. and C. Fry. "Evaluation of an Attempt to Teach Scientific Inquiry and Criticism in a Working Class High School." Journal of Research in Science Teaching, 10: 135-142, 1973.

Danner, F. W. and M. C. Day. "Eliciting Formal Operations." Child Development, 48: 1600-1606, 1977.

Douglas, J. D. and A. C. Wong. "Formal Operations: Age and Sex Differences in Chinese and American Children." Child Development, 48: 689-692, 1977.

Hardy-Brown, K. "Formal Operations and the Issue of Generalizability: The Analysis of Poetry by College Students." Human Development, 22: 127-136, 1979.

Howe, A. C. and J. Mierzwa. "Promoting the Development of Logical Thinking in the Classroom." Journal of Research in Science Teaching, 14: 467-472, 1977.

Inhelder, B. and J. Piaget. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Keating, D. P. and L. V. Clark. "Development of Physical and Social Reasoning in Adolescence." Developmental Psychology, 16: 23-30, 1980.

Kuhn, D. and J. Angelev. "An Experimental Study of the Development of Formal Operational Thought." Child Development, 47: 697-706, 1976.

Kurtz, B. and R. Karplus. "Intellectual Development Beyond Elementary School VII: Teaching for Proportional Reasoning." School Science and Mathematics, 79: 387-398, 1979.

Lawson, A. E. "The Developmental Learning Paradigm." Journal of Research in Science Teaching, 16: 501-515, 1979.

Lawson, A. E. "Formal Reasoning, Achievement, and Intelligence: An Issue of Importance." Science Education, 66: 77-83, 1982.

Lawson, A. E., F. H. Nordland and A. De Vito. "Relationships of Formal Reasoning to Achievement, Aptitudes, and Attitudes in Preservice Teachers." Journal of Research in Science Teaching, 12: 423-431, 1975.

Lawson, A. E. and W. Wollman. "Encouraging the Transition from Concrete to Formal Cognitive Functioning-an Experiment." Journal of Research in Science Teaching, 13: 413-430, 1975.

Lawson, A. E. and W. T. Wollman. "Cognitive Level, Cognitive Style, and Value Judgment." Science Education, 61: 397-407, 1977.

Linn, M. C., S. Pulos and A. Gans. "Correlates of Formal Reasoning: Content and Problem Effects." Journal of Research in Science Teaching, 18: 435-447, 1981.

Neimark, E. D. "Longitudinal Development of Formal Operations Thought." Genetic Pyschology Monographs, 91: 171-225, 1975.

Renner, J. W., D. K. Prickett and M. L. Renner. Evaluating Intellectual Development Using Written Responses to Selected Science Problems. A report to the National Science Foundation on grant no. EPP75-19596. Norman, Oklahoma: University of Oklahoma Press, 1977.

Scardamalia, M. "Information Processing Capacity and the Problem of Horizontal Decalage: A Demonstration Using Combinatorial Reasoning Tasks." Child Development, 48: 28-37, 1977a.

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

Siegler, R. S. and R. M. Liebert. "Acquisition of Formal Scientific Reasoning by 10- and 13-Year-Olds: Designing a Factorial Experiment." Developmental Psychology, 11: 401-402, 1975.

Siegler, R. S., D. E. Liebert and R. M. Liebert. "Inhelder and Piaget's Pendulum Problem: Teaching Preadolescents to Act as Scientists." Developmental Psychology, 9: 97-101, 1973.

Stone, C. A. and C. M. Day. "Levels of Availability of a Formal Operational Strategy." Child Development, 49: 1054-1065, 1978.

Wollman, W. T. and B. Chen. "Effects of Structured Social Interaction on Learning to Control Variables: A Classroom Training Study." Science Education, 66: 717-730, 1982.

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

Gabel, Dorothy L. and Robert D. Sherwood. "Effect of Using Analogies on Chemistry Achievement According to Piagetian Level." Science Education, 64 (5): 709-716, 1980.

Descriptors--*Chemistry; *Congitive Ability; *Concept Formation; Educational Research; Science Curriculum; Science Education; *Science Instruction; Secondary Education; *Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by Patricia H. Suter, Del Mar College, Corpus Christi, Texas.

Purpose

A major purpose of this study was to determine whether the use of analogies over a wide variety of topics in high-school chemistry classes for the entire year would improve chemistry achievement. In prior studies the authors felt that the concepts were limited in scope and that no positive results had been achieved for a variety of reasons.

A second purpose was to determine whether there were differential effects on achievement for concrete operational and formal operational students who were exposed to verbal analogies in their study of chemistry.

This study had two other purposes: (1) to determine whether the use of verbal analogies would have an effect on changing students' Piagetian levels during the school year and (2) to determine whether chemistry achievement was related to the students' understanding of the analogies.

Rationale

It would appear that analogies would help make formal concepts concrete by presenting the concepts in terms familiar to the student. Concrete operational students, however, may be unable to relate the analogies to the new concepts because this involves correlational thinking.

Students in the treatment and control groups were provided with study guides on each topic. The treatment groups' guides differed from the controls' in that: (1) they contained pretests to see how familiar students were with the analogies and (2) they had problems utilizing analogies in place of some regular chemistry problems.

Students' levels of logical thinking were assessed using a shortened form of the Longeot test administered during the first and last weeks of the school year. Students' scores from this test were not divided into Piagetian levels, but were kept as raw scores. This permitted analysis of the data by regression analysis.

Chemistry achievement was assessed using two methods. Four to six multiple choice-type items directly related to the content of each topic were embedded in the teachers' regular chapter tests. This resulted in a 68-item composite with a reliability using coefficient α of 0.82. In order to obtain a more general measure of chemistry achievement, the American Chemical Society (ACS) Achievement Test, Form 1975, Part A was administered at the end of the school year. The reliability of this 40 item test was 0.74 using KR-20.

The assessment of the students' familiarity with the verbal analogy was determined by a composite of all analogies used in the units and was administered to the treatment classes and to several control classes at the end of the school year. The composite had a total of 43 items. If students achieved 90 percent or better on the composite, they were judged to have understood the analogies.

Research Design and Procedure

A sample of 277 students enrolled in nine chemistry classes of three teachers in three schools was used. Data were analyzed for students who were present for 7 out of 12 chapter tests, the modified Longeot tests, and who had the same treatment throughout the year. Most students were high-school juniors although there were some sophomores and seniors.

Teacher participants taught at least two chemistry classes. One-half of each teacher's classes were randomly assigned to the experimental treatment, the other half to the control group. The treatment consisted of the teacher presenting 10 chemical concepts in terms of analogies. Detailed instructions with step-by-step procedures were given to each teacher in written teacher guides that varied in length from three to nine pages. Lessons extended over one to four class periods. In order to control the length of the presentations and to give similar treatment in the non-analogy classrooms, teachers were also provided with guides for those classes where additional examples from chemistry were used in place of analogies. Topics taught using analogies were: (1) Laws of Definite and Multiple Proportions, (2) Mole Concept, (3) Stoichiometry (mass-mass, mass-volume, volume-volume relationships), (4) Atomic Theory, (5) Kinetics, (6) Equilibrium, (7) Solutions and Concentrations, (8) Entropy, and (9) Oxidation and Reduction.

Findings

The means and standard deviations of students in the treatment and control groups according to their Longeot test scores on the two dependent measures were determined. The data were analyzed with the student as the unit of analysis.

The control group did significantly better than the treatment group. The additional practice problems seem to have been more beneficial than the use of analogies. The strong Longeot effect was not unexpected by the authors, considering the need for the use of more formal thought on the chemistry topics covered in the study.

The analysis of the data on the more general ACS-NSTA test did indicate an interesting trend. The examination of the data on means indicates that for low scores on the Longeot test, the treatment means are higher than the control group. For higher scores ($\geq$), the control group scores were higher. This would seem to indicate that the analogies

may have been effective for students of lower formal reasoning ability, but not especially useful for the more capable students. Because this interaction was not statistically significant for the other measure of achievement (unit test sum), the result must be interpreted with care.

In order to determine whether the use of analogies by students over an entire school year would have any effect on changing students from the concrete to formal operational stage, pre-post scores on the modified Longeot test were analyzed using t-tests for the experimental and control groups. No significant differences were found between pretest and posttest means for either group indicating no change in cognitive level according to treatment.

Interpretations

One reason why the use of analogies may not enhance chemistry students' achievement is because students are not familiar with the analogies. A surprisingly large number (48%) did not understand 90 percent of the analogies. The t-tests indicated that students who understood the analogies scored significantly higher than those who did not.

The major conclusion of this study is that students who scored higher on the modified Longeot test achieved more chemistry concepts than those whose scores were low.

Results also indicate that using mental analogies does not result in greater achievement for all types of students. Students who scored lower on the modified Longeot test benefitted more from the analogies than did students in the higher stages. Students at the upper end of the scale profited more from extra practice problems than from the analogies. Because these students were already able to think at the formal level, it was unnecessary to make the formal concepts concrete for them.

The authors found that the analogies commonly used in chemistry textbooks are not familiar to students and they fail both to understand them and apply them to the chemical concept. Thus, the use of these analogies needs to be reassessed.

ABSTRACTOR'S ANALYSIS

The authors of this research report studied the effects of using analogies to help students understand the very abstract concepts of chemistry. For those of us who teach beginning chemistry, either at the high school level or college level, the results are discouraging. It would appear that the more formal students do not need analogy and the more concrete operational students do not benefit from the use of analogies. The lack of understanding of the analogy used, or the inability to apply the analogy to the particular concept, is disturbing. The authors' suggestion that the student be encouraged to develop the relationship of the analogy to the particular concept is the latest educational theory. It would seem that the more formal student can do that already.

This study seems to have been well done and valid. The authors suggest that more work be done to determine if concrete operational students are capable of taking a familiar situation and applying it to a new one such as in chemistry. It would also appear that new analogies be sought because a large group (48%) of the students did not understand the analogy.

An excellent reference list is included in the paper which is helpful for those who would like to investigate this problem in their own situation.

Moreira, Marco A. and Carlos A. Santos. "The Influence of Content Organization on Students' Cognitive Structure in Thermodynamics." Journal of Research in Science Teaching, 18 (6): 525-531, 1981.

Descriptors--*College Science; Conventional Instruction; Engineering Education; Higher Education; *Learning Theories; Organization; Physics; Science Curriculum; Science Education; *Science Instruction; *Sequential Approach; Teaching Methods; *Thermodynamics; Vertical Organization

Expanded abstract and analysis prepared especially for I.S.E. by Ubiratan D'Ambrosio, Universidade Estadual de Campinas, Brazil.

Purpose

The authors propose to study the influcence of two different content organizations of thermodynamics, one traditional and the other based on Ausubel's learning theory, on students's cognitive structure. The research uses a word association test analyzed through hierarchial clustering analysis.

Rationale

Most of current approaches to achievement in physics tend to focus on cognitive abilities related to problem solving. Instead, the authors' focus is on the relationship between the organization of knowledge being transmitted and the student's cognitive structure. It is assumed that a correlation is present. This leads to looking for a model of cognitive structure, instruments to get information about it, and ways of representing it. Previous research in this direction from the first author, as well as by S. C. Johnson, R. J. Shavelson, J. D. Novak (among others) is acknowledged and referred to.

Research Design and Procedure

The basic framework in which the research is designed is D. Ausubel's cognitive structure as synthetizing a hierarchial organization of concepts. In this framework of cognitive structure there is a hierarchial configuration in which concepts of a given body of knowledge are represented by points and the degree of relationship among concepts is measured by distances. With respect to content organization two approaches were utilized: one based on "Fundamentals of Physics" by Halliday and Resnick (1970) and other based on the learning theory of David Ausubel (1978). Basically, in this last one it is said that the most inclusive/the most general idea, phenomena and concepts should be presented as conceptual anchorage for subsequent learning. In the specific case of thermodynamics, which is the subject under study, the first approach introduces thermodynamics evolving linearly from the concepts of temperature and the zeroth law and finishing with the second law and entrophy. In the Ausubelian approach all the concepts and basic laws of thermodynamics are introduced at the beginning of the instruction and then differentiated, and instruction progresses.

The research was carried on in the state of Rio Grande do Sul, during the first semester of 1978. Subjects were 58 engineering students randomly selected and belonging to two different sections of the same course. One of the groups followed the traditional approach and the other studied the same content in the Ausubelian approach. The content organization was the independent variable and the degree of conceptual relationship, as measured by the word association test, was the dependent variable. The tests were organized in such way that a stimulus word of the interest of the researcher would elicit a series of responses. In this specific research, each stimulus word was a physical concept appearing at the top of the booklet test. Subjects were asked to write down on the corresponding page as many words in the field of physics as they could think of in association with the stimulus word. Stimulus words were: heat, energy, work, temperature,

entropy, state variable, thermal equilibrium, and specific heat. Then the authors computed the relatedness coefficients among pairs of concepts which depend on the number of responses to a given stimulus, the order and the overlap of response lists. This gives rise to a similarity matrix, which was investigated by the method of hierarchial clustering.

Findings

Conceptual hierarchial clustering of the group following the traditional approach before and after instruction showed practically no difference in the behavior of subjects. In both cases the clustering is almost linear, in the sense that such concept was associated to the already existing cluster.

Now, looking into the conceptual hierarchial clustering for those following the Ausubelian approach before and after instruction, there is a clear difference. Even if the clustering is also almost linear, there are evident changes in the cluster "energy-work" which changed into "energy-work-heat", which are the concepts involved in the first law of thermodynamics. Also, there are differences in the clustering temperature and heat to temperature and thermal equilibrium after instruction, and of entropy and state variable, which got closer after instruction.

Interpretations

The results show that the hierarchial clustering order changed more in the group following the Ausubelian approach. This suggests that the Ausubelian approach to the organization of content of thermodynamics influences the student's cognitive structure in such a way that the conceptual hierarchies are more coherent with the basic laws and the conceptual structure of the subject. Similar results on electricity and magnetism were found by the first author and other researchers in previous works.

ABSTRACTOR'S ANALYSIS

This line of research provides a sound basis for curriculum design. Although the authors have touched only content organization, similar research might be designed to obtain information on the effects of adopting different methods and even on broader objectives of the pedagogical practice. As a result, mutual implications of curriculum design and cognitive structure might be better understood.

Although the authors tend to minimize the implications of the research in effective teaching, it is our belief that this kind of research can have a direct effect on planning strategies for teaching aimed at long term results. Surely, subsequent testing aimed at analyzing the degree of retention of concepts would be an important follow-up of this kind of research.

Arrangements of content in defining curriculum is probably one of the most crucial issues facing science educators. There is a practice, with deep roots, to organize content in a linear way, building-up concepts upon previously acquired former concepts, supposedly less complex. This is based on nothing more than tradition. Each redefinition of the syllabus is based on previous syllabi and it becomes hard to substantially change the order of presentation. Ausubel's approach is refreshing, in the sense it proposes to present more general and most inclusive concepts before going into details. This research, together with other related studies shows the advantages of this approach to the traditional one.

ATTITUDES

Harty, H. and N. Al-Faleh. "Saudi Arabian Students' Chemistry Achievement and Science Attitudes Stemming from Lecture - Demonstration and Small Group Teaching Methods." Journal of Research in Science Teaching, 20 (9): 861-866.

Descriptors--*Academic Achievement; *Chemistry; Conventional Instruction; Grade 11; High Schools; Lecture Method; Science Education; *Science Instruction; *Secondary School Science; *Small Group Instruction; *Student Attitudes; Teaching Methods

Expanded abstract and analysis prepared especially for I.S.E. by Willis J. Horak, The University of Arizona.

Purpose

This study analyzed the effects of two distinct teaching methods upon chemistry students' achievement and attitudes. The two methods utilized were the large group lecture - demonstration and the small group laboratory approach. The study was carried out in Saudi Arabia where the predominant teaching method at the secondary level continues to be lecture based.

Rationale

The research cited points out that science teaching strategies have varied effects upon students' achievement and students' attitudes. Many of the studies have shown that small group and experimentation methods offer numerous advantages over large group lecture methods. However, most of the studies have been carried out in the United States. This study was undertaken to see if the previous results could be extrapolated to other cultures. If possible, then it may be conjectured that other instructional methods also offer advantages to Saudi Arabian students.

Research Design and Procedures

The sample consisted of 84 male chemistry students in a high school in Saudi Arabia. The students were randomly assigned to the two treatment groups. The treatments were basically described as a lecture - demonstration instructional method and a small group (4 students per group) laboratory instructional method. The content of instruction included six chemistry concepts which were a regular part of the national chemistry syllabus.

The criterion instruments used in the study were a self developed chemistry achievement test and the previously developed instrument entitled, "An Attitude Survey for Junior High Science" (Fisher, 1973). The instruments were administered immediately after the six week instructional period and also one week later.

After testing the data for normality of scores, a t-test was employed to test for significant differences. The 0.05 level was chosen. Four analyses were conducted. These tested for differences in attitude and achievement for both the immediate and the delayed test administration.

Findings

The statistical analysis revealed significant differences in achievement at the 0.001 level for both the immediate and the delayed post-tests. It also revealed no significant differences in attitudes for both the immediate and delayed post-tests. There were also no significant correlations revealed between achievement and attitude for either the lecture - demonstration group or the laboratory group. The correlations were computed for both the immediate and the delayed post-test.

Interpretations

The analyses show that chemistry achievement may be enhanced by employing laboratory - centered, small - group activities. Much of the variance between the two scores, however, may be attributed to individuals feeling challenged and performing beyond expectation and also to the halo and Hawthorne effects. The fact that attitude scores dropped for both groups on the delayed post-test measure may also be due to the weakening of these three effects.

ABSTRACTOR'S ANALYSIS

This study probably raises as many questions as it answers. First of all, it appears as if neither of the methods was highly successful. As mentioned in the article, the pre-test to immediate post-test scores only increased by 3.9 for the lecture/demonstration method and by 6.8 for the laboratory method. This does not appear to be a large increase for six full weeks of instruction. The authors do state that the individuals felt challenged and were performing beyond expectations. This appears questionable for a test where the high possible score was 40, and the overall achievement increase was so minimal. In fact, the students showed an increase between 20 to 40 percent during the week between the immediate and the delayed post-tests when no related instruction on the topics was being conducted. I feel the authors needed to address this in their discussion.

The reported correlation between achievement and attitude also raises other questions that could be addressed. Why did the correlation change direction for both the lecture/demonstration and the laboratory group? They actually went in opposite directions. The lecture/demonstration group went from a positive correlation to a negative correlation. The laboratory group went from a negative correlation to a positive correlation. This apparent anomaly needs to be considered.

One other point that might have been more fully explained was the use of an attitude scale that was designed for junior high students. Again the authors state that potential unforeseen cultural biases were a consideration. However, they do not state why they felt an instrument designed for a younger age American population would ameliorate this situation. Additionally, I would have liked to see the reliability and validity checked on the version that had been translated into the Arabic language.

Overall, the study is one of the type needed if we are going to generalize findings to diverse cultural groups. It was done thoroughly and the analyses of the data was appropriate. I would have liked to see a table of means and standard deviations in the paper. They are in the text, but it is quite tedious to sort them all out. From such a table, one can much easier see the overall trends and magnitudes of differences. I agree with the authors that this initial study could serve as a first step for further research into science teaching methods in Saudi Arabia.

REFERENCE

Fisher, T. H. "The Development of an Attitude Survey for Junior High Science." School Science and Mathematics, 73: 647-652, 1973.

Fraser, Barry J. and William L. Butts. "Relationship Between Perceived Levels of Classroom Individualization and Science-Related Attitudes." Journal of Research in Science Teaching, 19 (2): 143-154, 1982.

Descriptors--*Attitude Measures; *Classroom Environment; *Individualized Instruction; Junior High School Students; Measures Individuals; Science Education; Science Instruction; Secondary Education; *Secondary School Science; *Student Attitudes

Expanded abstract and analysis prepared especially for I.S.E. by Marvin Bratt, The Ohio State University.

Purpose

The purpose of this study was to identify relationships between perceived levels of classroom individualization and science-related attitudes as measured by the Individualized Classroom Environment Questionnaire (ICEQ) and the Test of Science-Related Attitudes (TOSRA).

There were five dimensions measured by the ICEQ and seven distinct attitudes measured by the TOSRA.

Eight significant relationships were found.

In addition, further support for the validity and reliability of the ICEQ and TOSRA instruments was provided.

An hypothesis examining classroom perceptions was tested to demonstrate the sensitivity of the ICEQ scales between classrooms.

Rationale

An extensive historical review of prior research concerning individualization and open education methodologies is provided as is a description of the measurement of classroom individualization (developmental description of the ICEQ). Similiar descriptions of previous and ongoing research on the measurement of science-related attitudes are presented. The present study follows a similiar pattern of data collection and analysis for attitude measurement.

There is very little contextual framework for the assessment of individualized instruction in science as measured by the ICEQ because it is new. In particular, the article is directed to a concern identified by NARST and NSTA in relationship to research on the effects of individualized instruction in science on science-related attitudes. These literature reviews were provided in support of the two instruments for the evaluation of the effect of individualized instruction and the assessment of attitudes.

Assumptions regarding individualized instruction differ. This paper defines indvidualization with the five criteria used in the ICEQ.

This study is related to studies that look at reasons for poor enrollment in science education courses in high school and college. It is related to studies of attitudes. It is also related to research on learning environments and classroom environments.

Research Design

The sample was composed of 712 students enrolled in grades 7-9 in 30 classes in a suburban area of Sidney, Australia. While not randomly selected, the sample was representative of the Sidney suburban schools. According to the authors, the sample is representative of a reasonable range of the degree of individualization of instruction in science in Australia.

Design of the ICEQ is described in previous studies and was guided by three criteria which make it relevant to this study. The test contains five dimensions of ten items each. A good discussion of the validity and reliability of the instrument is provided. In addition, a statistical discussion of the desirability of the instrument is provided.

The TOSRA is a seven scale test containing ten items per scale. Good information is given relative to the validity and reliability of the scales for this and previous studies. The TOSRA was given as a pre and posttest, while the ICEQ was given midway in the one-year study period.

Since the purpose of the study was to identify relationships between individualized instruction and attitude, a multiple regression analysis technique was selected for the following reasons:

- the predictors were known to be intercorrelated, statistical power could be maximized by maintaining continuous variables as such instead of reducing them,
- an estimate of the amount of variance accounted for in each criterio.. by each predictor variable is provided automatically (R-squared),
- and finally, multiple regression could be employed to test the combined effects of sets of variables.

ICEQ scales included Personalization, Participation, Independence, Investigation and Differentiation.

TOSRA scales included Social implications of Science, Normality of Scientists, Attitude toward Inquiry, Adoption of Scientific Attitudes, Enjoyment of Science Lessons, Leisure Interest in Science and Career Interest in Science.

Intact classes were used as the primary sampling units and the means for classes were used for statistical analyses.

Findings

R-squared values were reported for each of the seven attitude scales for pretest, the five ICEQ variables and the total. R-squared values ranged from 30.4 to 38.5 for pretest, 7.4 to 29.1 for ICEQ and 42.3 to 61.7 for the total. The total values were all significant and, as one would expect, the pretest values were all significant.

Four of the ICEQ values were significant, including Social Implications of Science, Enjoyment of Science Lessons, Leisure Interest in Science and Career Interest in Science. An examination of the b-weights for each ICEQ scale revealed eight significant relationships between an individual posttest attitude outcome and an individual classroom dimension.

When the corresponding pretest was controlled:

- Social Implications of Science scores were greater in classrooms perceived as having greater Personalization, Independence and Investigation.

- Greater Enjoyment of Science Lessons was expressed in classrooms perceived as having greater Personalization, Participation and Independence.

- Higher Leisure Interest in Science scores and Career Interest in Science scores were obtained in classes perceived as having greater Participation.

In each of the eight cases, perceptions of greater individualization were associated with more positive attitudes.

Interpretations

Student perceptions on a set of five dimensions of individualization accounted for a significant increment in the variance in end-of-year attitude scores beyond that attributable to pretest scores for four of seven scales.

All significant associations between an individualized dimension and an attitudinal dimension were positive.

Further research in this area was strongly encouraged especially with other samples.

ABSTRACTOR'S ANALYSIS

The relationship of this study to other studies in attitude measurement is particularly important because it adds to our understanding of the effects of individualized instruction. It is also important because the authors were quite careful in defining both criteria for evaluating individualized instruction and in stating the attitudes they were measuring. The more researchers probe the different types of

instruction and relate it to other variables, the more this increases our understanding of the processes of learning related to the types of instruction used. This can lay the foundation for experimental studies of specific learning styles or teaching strategies. We know that attitudes among the general public and in children, in particular, are not positive regarding science. This trend must be reversed and one of the only places to do this is in the schools.

The study is important because it addresses a specific need identified by international learned societies in science education. There is a distinct lack of research regarding individualization of instruction in science education. Researchers interested in developing studies in the area can and should follow mandates for research established by national and international learned societies. Far too many criticisms have been leveled at science education research for unnecessarily duplicating attitude scale construction and for poor instrument documentation. These can only serve to erode the credibility of science education as a legitimate discipline.

While not introducing new methods to the research community, it should be noted that few studies are found that are as complete and carefully done. This is particularly important because it helps new researchers (and more established ones) see the need for and importance of doing careful work.

The report also provides a method for doing such a study. This can be of vital importance to new researchers attempting to build skill in measurement and design techniques in science education.

Measurement of attitudes has been continually criticized because of the problem of self reporting instrumentation, validity and reliability. This study illustrates up-to-date methodlogy that should be used in studies of attitude. It is the best method so far developed.

There are always questions that may be raised while reviewing any study. Some questions that arise out of this study are...

- Were there differences between boys and girls on the variables related in this study?

- The junior high school student is in a very peculiar stage in development of attitude. How would the data for high school students be different, or, for that matter, for young adults, college students, preservice teachers? What recommendations can be drawn for those school systems that have experienced drops in enrollment of high school science courses?

- Will the students who have positive attitudes or who are educated in classes utilizing individualized instruction take more coursework in science?

- What should be done to assist schools in making decisions concerning the assignment of science classes to teachers whose primary motivation in education is not science (coaching, for instance?) These teachers cannot devote complete energy to either educational area, let alone do justice to either.

Wright, Jill D. and Paul B. Hounshell. "A Survey of Interest in Science for Participants in a Junior Science and Humanities Symposium." School Science and Mathematics, 81 (5): 378-382, 1981.

Descriptors--Interests; Science Education; *Science Interests; Secondary Education; *Secondary School Science; Secondary School Students; *Student Interests; Surveys

Expanded abstract and analysis prepared especially for I.S.E. by Joan Schumm.

Purpose

This study proposed to investigate the development of science interests in a select group of high school students. The group chosen to participate in this study was composed of those high school students attending the 1978 North Carolina Junior Science and Humanities Symposium (NCJSHS). The students in this group had been chosen to attend the NCJSHS on the basis of their interest in science as well as on their past achievement in science courses. These students completed a questionnaire to assess their science interests and the development of those interests. The items of the questionnaire probed the initiation and maintenance of science interest, the intensity of that interest, and the scientific area of most interest to each student. Gender differences in the science interest patterns were then examined.

Rationale

No rationale was given for this paper. Previous studies describing the science interests of high school students were not cited, nor were studies investigating gender differences in science interests mentioned.

Research Design and Procedure

A survey design (ex post facto) was used in this study. Student interest in science was measured with a questionnaire given to 147 students (79 males and 86 females) attending the 1978 NCJSHS. Interest in science was measured as "high," "medium," or "low." Other items on the questionnaire asked students their age at first interest in science, whether or not interest developed through school activities, factors most influential in developing science interest, areas of interest in science, and proposed college major. Each of these categories (the dependent variables) was then analyzed for gender differences (gender--the independent variable) using the chi square statistic and alpha levels of 0.05, 0.01, and 0.001.

Although the item "interest in science" was described as having alternatives of "high," "medium," and "low," the alternatives for other items on the instrument were not given. Instrument reliability and validity were not reported, nor were the procedu. s utilized by the investigators given.

Findings

Seventy-three percent of the students reported "high" interest while 27% reported "medium" and 0% reported "low" interest in science. Science interest among males was higher than among females. Males became interested in science at an earlier age than did females. Science interest was developed through school for 74 percent of the students. School was more influential in developing the interest of females than of males. Students reported the following school factors as important in developing their science interest: high school teachers (71% of the students), junior high school teachers (40%), school textbooks (41%), field trips (40%), student science projects (39%), and books and magazines in the shcool library (39%). Outside factors reported by these students as being important in developing their science interests include: books and magazines read outside of school (65% of the students), parents (40%), and television (35%).

Science interest among males was developed more often by books and magazines read inside of and outside of school than among females. Science interest among females was more influenced by friends than among males. All students planned to enter college or technical school. Males were more interested in an engineering major than were females while females were more interested in a medicine/health major than were males. Finally, males were more interested in physics, chemistry, and astronomy than were females.

Interpretations

From these results, the investigators concluded that the teacher's role was important in stimulating student interest in science. Of these high school students, 74 percent claimed that their interest in science was developed through school, where both junior and senior high school teachers were singled out as the most important influence in developing that interest. From the results of this study, the investigators inferred the following: (1) Science teaching in grades K-6 should be improved; (2) Female students should be exposed to more science-related books and magazines; (3) Science projects and field trips should be included in the science curriculum; and, (4) More time should be devoted to informing students about careers in science.

ABSTRACTOR'S ANALYSIS

Theoretical Context. This study investigated the science interests of a special group of high school students; namely, those students in North Carolina chosen on the basis of their interest and achievement in science to attend the 1978 NCJSHS. The investigators attempted to define that science interest in terms of its initiation, growth, and development.

Hypotheses regarding the science interests of these students were not stated. Earlier research was not used in developing a theoretical basis for this study. The study would have been much stronger had these items been included.

Research Design and Validity. The students responded to an instrument assessing their interest in science at the present time as well as asking them to remember when they first became interested in science. They also reported which person(s) and material(s) were most influential in developing that science interest. The questionnaire, developed by the investigators for this study, could have been described more fully. For instance, were the students to respond to questions by choosing appropriate alternatives, or were the questions open-ended? Were the students allowed to choose all of the areas of science that interested them, or only the area of most interest? Were the students given a choice of college majors and asked to choose which one they would be most interested in pursuing, or were they allowed to state their proposed major? What validity did the instrument have in terms of assessing science interest? What was the reliability of the instrument? Perhaps some of the items in the instrument could have been presented in a Likert-type format. For instance, when responding to the item concerning which area of science students were most interested in, students could have rated their interest in one area of science independently of other areas.

The high school students chosen to participate in this study were a select group of students who were already interested in science. While the investigators did mention that these students were a special group, they failed to note the possibility of a second bias. It is possible that the students' answers may have been biased because the science inventory was administered during a science symposium. These design problems limit the generalizability of this study because the results were based on data obtained from a select group of students answering a specific inventory under conditions that could bias their responses.

Results and Interpretations. The investigators report that school is, perhaps, a very influential factor in developing students' science interests--especially for young women. They also report that males become interested in science at an earlier age than do females. Females do not tend to develop as much interest as males do in majors or careers traditionally chosen by males. From these results, the investigators have proposed that science teaching in elementary school should be improved and science projects and field trips should be included in the teaching process. They also have proposed exposing female students to more science-related books and magazines. Finally, they have suggested that more time be devoted to informing students about science career opportunities. These proposals would have been more strongly supported by also reporting results from previous research.

The proposal that science teaching can be improved in elementary school may be supported in part by the following study: Brush (1979) found that students' stereotypes of scientists are largely based on the information found in their school textbooks. These stereotypes may influence students to feel that they are not like scientists and so do not take interest in a scientific career. In addition to the general conceptions young students have about scientists, O'Brien and Corder-Boltz (1978) have found that by age seven children already have established sex-role stereotypes which may influence the development of their interests in science.

Hardin and Dede (1973) found that teachers play significant roles in developing student interests. They suggest that teachers, by their attitudes and actions, may lessen some of the effects of sex-role stereotyping on the development of science interests among female students. Teachers, then, may be important role models for their students.

Other role models also may encourage females to pursue non-traditional careers. For example, Stake and Granger (1973) suggest that female students, introduced to female role models in non-traditional careers, may develop more positive interests in pursuing non-traditional careers. McLure and Piel (1978) agree that female role models have a positive influence on young women's choices of majors and careers.

Farmer's (1976) review of recent research may help explain why female students are not as interested in science as are males. According to Farmer, females have lower levels of academic self-confidence and competitiveness than do males. Females perform as well as males in math and science until ten years of age. When they become older than ten, they begin to perform at a lower level than do males.

In summary, the lower interest in science courses and careers among females is a very complex problem. It involves perceptions of one's ability, the way one interacts with and is accepted and encouraged by others, the role models with whom one interacts, and the courses one chooses. Simple answers to this complex problem are unlikely. More research, including a multivariate approach, is necessary.

Suggestions for Future Research. While the chi square statistic was appropriate as used in this study, further analysis may have proven fruitful. For instance, hypothesized models tested in path analysis would allow the investigators to tie together important factors in developing student science interest, to discern the relationships between these factors, and to relate the amount each factor adds to the proposed model. Path analysis may have resulted in a more definitive model of factors influencing the development of science interest in these high school students. It may have been easier to draw implications from the model, as the importance of each factor would have been determined.

The research reported here adds to the body of research on the development of high school students' interests in science. It supports results obtained in earlier studies. In addition, the study focuses on high school students who already are interested in science. The implication of this study is two-fold: (1) We should determine which factors play significant roles in the development of science interests; and, (2) We should implement those factors to encourage high school students (especially young women) to develop interests in science.

REFERENCES

Brush, L. R. "Avoidance of Science and Stereotypes of Scientists." Journal of Research in Science Teaching, 16 (3): 237-241, 1979.

Farmer, H. S. "What Inhibits Achievement and Career Motivation in Women?" The Counseling Psychologist, 6 (2): 12-15, 1976.

Hardin, J. and C. J. Dede. "Discrimination Against Women in Science Education." The Science Teacher, 18-21, December, 1973.

McLure, G. T. and E. Piel. "College-bound Girls and Science Careers: Perceptions of Barriers and Facilitating Factors." Journal of Vocational Behavior, 12, 172-183, 1978.

O'Brien, S. L. and C. R. Corder-Boltz. "The Effect of Television on Children's Stereotyping of Women's Work Roles." Journal of Vocational Behavior, 12, 233-244, 1978.

Stake, J. E. and C. R. Granger. "Same Sex and Opposite Sex Teacher Model Influence on Science Career Commitment Among High School Students." Journal of Educational Psychology, 70 (2): 180-186, 1978.

Towse, P. J. "Do New Science Courses Improve Attitudes Toward Science? A Study in Lesotho." Science Education, 67 (2): 159-169, 1983.
Descriptors--*Attitude Change; Foreign Countries; *Program Effectiveness; Questionnaires; Science Curriculum; Science Education; *Science Programs; Secondary Education; *Secondary School Science; Sex Differences; *Student Attitudes

Expanded abstract and analysis prepared especially for I.S.E. by Thomas R. Koballa, Jr., The University of Texas at Austin.

Purpose

The investigation had a two-fold purpose. The primary purpose of the investigation was to test whether or not a new physical science course, the Lesotho Introductory Science Improvement Programme (LISIP), promoted more positive attitudes toward science among junior secondary students in Lesotho than did the "traditional" physical science course already in use. The development of a Likert-type scale to measure students' attitudes toward science was the secondary purpose.

Rationale

The investigation was predicated on the assumption that highly student-centered, investigative science courses promote more positive attitudes toward science than courses that do not possess such attributes.

Research Design and Procedure

The Lesotho Attitude to Science Test (LAST) was developed and analyzed in a manner similar to the Science Attitude Questionnaire (Skurnik & Jeffs, 1970) and the Science Opinion Poll (Laughton & Wilkinson, 1970). First, a pool of 80 items was created that focused on two affective objectives defined in the introduction of the LISIP

course: (1) students' interest and enjoyment in science and (2) their awareness of the contribution of science to the social and economic life of the community. The items were then reviewed by a panel of 20 science educators, language specialists, and others and subjected to two factor analyses. The final version of the scale contained 33 items grouped into four subscales: (1) interest, (2) teacher, (3) social implications, and (4) difficulty. Students responded to each item using one of five responses (strongly agree, agree, not sure, disagree, strongly disagree).

A posttest only control group design was used in the investigation. The attitude scale was administered to 647 students (172 boys and 167 girls in the LISIP group, and 138 boys and 170 girls in the control group) early during the 1976 school year.

Findings

No significant differences were found between the groups when the scores from the total scale and its subscales were compared. When the groups were segregated by gender, no significant differences were found between boys in either group. Girls, however, were found to differ significantly on the difficulty subscale; science proved to be more difficult for girls studying LISIP. When subscale items were examined individually, one significant difference was identified. Girls in either group were more subject to the influence of their science teachers than were boys.

All subscales (interest, teacher, social implications, and difficulty) were found to be highly correlated. Of the four subscales, the highest correlation with science performance as measured by the Junior Certificate Examination was with interest in science.

Interpretations

Both boys and girls were very interested in science. Girls were more influenced by their science teacher.

The correlations between subscales were similar to those reported for several studies carried out in Britain by Nuttall (1970), Meyer (1970), and others.

Interest in science appears to be the strongest predictor of students' science performance on the Junior Certificate Examination. Although interest correlated the most significantly with science performance, it is impossible to suggest if interest is caused by good performance or good performance causes interest.

The investigation suggests that the LISIP course did not promote more positive attitudes toward science among Lesotho junior secondary students than did the "traditional" physical science course used as the control. Three explanations were offered to account for the lack of difference in attitudes toward science between the two groups:

1) Teachers assigned to the control group, being just as enthusiastic as their LISIP counterparts, were successful in promoting equally positive attitudes.
2) The LISIP course imposed added pressures on the teachers and students, which prevented the development of more positive attitudes.
3) Teachers in neither group considered affective objectives to be as important as cognitive ones.

ABSTRACTOR'S ANALYSIS

The popularity of comparing courses (student-centered, investigative vs. traditional) in an effort to determine those that are most likely to promote more positive attitudes toward science is borne out by the reviews of Peterson and Carlson (1979) and Schibeci (1984). It may be suggested, however, that the use of courses as treatment, more often

alike than they are different, has a major liability: their multi-dimensional nature. A shotgun-type variable lacking a theoretical elegence, the variables are typically an aggregate of interacting stimuli--all the many components that science teachers include in a syllabus. Even when a significant difference between mean scores is the outcome in an investigation of this type, there is no way of knowing which of the many dimensions (and their interactions) is responsible for the results. A whole course may not be needed to bring about attitude change.

Also an inherent problem in using a course as an independent variable is the length of treatment. Treatment periods of weeks or months, long considered an asset among many educational researchers, may well be a liability to attitude research; time can exacerbate the problem of undefined dimensions within the treatment. Long treatment periods in attitude research is an artifact of educational research--not the practice of attitude theorists.

Not being able to administer the attitude scale at the end of the treatment period (the 1975 school year) proved to be a major problem encountered in the investigation. Students who did not pass the Junior Certificate Examination, used to select students for the final two years of secondary education, were not among the respondents to the attitude scale. With a sizable portion of the subjects who composed the two treatment groups not responding, the findings of the investigation must be interpreted with caution.

Science educators should base attitude investigations on theoretical models in social psychology. A 40-year legacy of theoretical development by many attitude theorists (e.g., Carl Hovland, Gordon Allport, and Phillip Zimbardo) is the major advantage of this option over testing courses or instructional modes as independent variables. A survey of the literature (Shrigley, 1983) and several basic investigations (Shrigley, 1976; Koballa, 1984) suggest that Hovland's learning theory approach appears compatible with the goals of science education.

Also there is a precedent for science educators borrowing from social psychology. Witness our use of their measuring techniques: Likert's summated rating and Osgood's semantic differential. It could be argued that the use of these measuring techniques is valid when attitude investigations in science education are based upon the theoretical underpinnings of social psychology.

Our purpose for studying attitude is its relationship to science related behaviors. Attitude serves as a mediating variable, an indirect but convenient means for measuring behavior. Using the attitude scale developed as part of this investigation and other scales to predict students' scores on the Junior Certificate Examination is in line with this purpose and warrants further investigation.

REFERENCES

Koballa, T. R., Jr. "Changing Attitudes Toward Energy Conservation: The Effect of Developmental Advancement on the Salience of One-sided and Two-sided Persuasive Communications." Journal of Research in Science Teaching, 21: 659-668, 1984.

Laughton, W. H., and W. J. Wilkinson. "The Science Opinion Poll. A Study of Attitudes for Use in G.C.E. Science Courses." Educational Science, 37: 25-27, 1970.

Meyer, G. R. "Reactions of Pupils to Nuffield Science Teaching Project Trial Materials in England at the Ordinary Level of the General Certificate of Education." Journal of Research in Science Teaching, 7: 283-302, 1970.

Nuttall, D. L. Administrator's Manual for Science Attitude Questionnaire. Slough, Great Britain: National Foundation for Educational Research in England and Wales, 1970.

Peterson, R. W. and G. R. Carlson. "A Summary of Research in Science Education--1977." Science Education, 63: 429-550, 1979.

Schibeci, R. A. "Attitude to Science: An Update." Studies in Science Education, 11: 26-59, 1984.

Shrigley, R. L. "Credibility of the Elementary Science Methods Instructor as Perceived by Students: A Model for Attitude Modification." Journal of Research in Science Teaching, 13: 449-453, 1976.

Shrigley, R. L. "The Attitude Concept and Science Teaching." Science Education, 67: 425-442, 1984.

Skurnik, L. S. and P. M. Jeffs. Science Attitude Questionnaire. Slough, Great Britain: National Foundation for Education in England and Wales, 1970.

IN RESPONSE TO THE ANALYSIS OF

Towse, P. J. "Do New Science Courses Improve Attitudes Toward Science? A Study in Lesotho." by Thomas R. Koballa, Jr. Investigations in Science Education 11 (4): 48-53, 1985.

Peter J. Towse
University of Zimbabwe

Koballa is right to point out the dangers of using a whole course as a uni-dimensional variable to determine attitude change. However, the course was not introduced as a single variable to see if it would improve attitudes toward science. Rather, the attitudes were tested on completion of the course to see if they had been improved by it, whatever the causative factor or factors in the course responsible for such a change might be. This is an important distinction, for the decision to introduce the course throughout Lesotho had been taken long before the investigation was carried out. As indicated in the introduction to the paper, one of the purposes of the investigation was to see if some of the rationale for the course was justified.

As the attitudes were assessed at the end of the course, Koballa's argument that "long treatment periods in attitude research is (sic) an artifact of educational research--not the practice of attitude theorists" seems somewhat confusing and irrelevant.

Dramatic changes within the Ministry of Education in Lesotho made it impossible to administer the evaluation as planned for December 1975, when all those who had completed the new course were still in school. For this, and for a whole variety of other reasons all too familiar to those working the the Third World, the evaluation had to be delayed for about three months, by which time only those selected for the final two years of secondary education were still in school. The implications of this and the resultant need to interpret the results of the study with caution were clearly stated in the discussion of the results. Koballa does not acknowledge this, but merely makes the same point in his analysis in a way which suggests that the investigator was unaware of

of this need for caution and that conclusions have been drawn too hastily. The investigator was all too well aware of the limited interpretation this forced on the results of the study and wishes that it could have been otherwise.

The list of references at the end of the analysis is Koballa's, not mine.

PROBLEM SOLVING

Mandell, Alan. "Problem-Solving Strategies of Sixth-Grade Students Who are Superior Problem Solvers." Science Education, 64 (2): 203-211, 1980.

Descriptors--Cognitive Processes; *Educational Research; Elementary Education; *Elementary School Science; *Grade 6; *High Achievement; *Problem Solving; Science Education

Expanded abstract and analysis prepared especially for I.S.E. by William C. Robertson, University of Colorado.

Purpose

The main purpose of the study was to find common problem-solving behaviors and strategies used by sixth-grade subjects classified by their teachers as superior problem-solvers. A secondary purpose was to determine if students classified as such were indeed superior problem-solvers.

Rationale

This investigation was conducted under the assumption that problem-solvers' strategies at different points in solution can be classified into four "phases," labeled by the author as (1) initial attack phase, (2) follow-up phase, (3) incubation phase, and (4) solution phase. These categories roughly correspond to similar categories described by Wicklegren and Wertheimer.

Research Design and Procedure

Subjects were 25 sixth-grade students selected for the study solely because they were identified by their teachers as superior problem-solvers. A detailed description of the selection process (Were any excluded? Was there a limit to the number chosen?) is not available from the report. Individual, audiotaped interview problem-solving

sessions were conducted with each subject. The audiotapes were transcribed to scripts, which were analyzed using a Problem-Solving Behavior Tally Sheet (PBTS) developed by the author. The subjects' statements were coded on the PBTS as belonging to one of the four previously mentioned "phases" and were further classified into from two to four sub-categories (types of behaviors) within each phase. Eleven types of problem-solving behaviors were identified in all. The taped session consisted of the presentation and solution of six problems, given in the same order to all subjects.

The first problem was used to familiarize the subjects with the procedure and was not coded. Subjects solved the problem, were asked how they arrived at the solution, and then were told that their "method of solution" was the information the experimenter wished to record. Throughout the rest of the session, a hand signal was used to encourage "thinking out loud".

The time for completion of each problem was recorded and the PBTS was used to compute an Efficiency Index (EI) for each subject on each problem. The EI was defined as the average number of statements per second required to solve the problem.

For each problem subjects were divided into successful problem solvers (SPS group) and nonsuccessful problem solvers (NPS group). Using only the data from the SPS group for each problem, Spearman ρ correlation coefficients were calculated for each of the eleven problem-solving behaviors. These ρ's compared subjects' EI ranks within the SPS group with a percentage ranking of the subjects' employment of a particular problem-solving behavior. For example, a student who used a particular behavior most often in his/her solution of problem 4 would be ranked first for the purpose of calculating the ρ between that behavior and the EI rankings for problem 4. The correlations thus gave an indication of the degree to which a particular problem solving behavior was associated with a low or high Efficiency Index.

Findings

Forty eight percent of the sample successfully solved one-half or more of the six problems, and these subjects were labeled the SPS group. The other subjects comprised the NPS group. The means of the SPS group on IQ (Kuhlman-Anderson Test: Form B), SRA Math and SRA Science subtest scores were significantly ($\alpha = .05$) higher than the means of the NPS group on the same measures. All ten of the multiply-nominated (selected as superior problem-solvers by two or three teachers) students were in the SPS group. Members of the SPS group tended to have lower EI values (implying greater efficiency), were not dependent of physical manipulations or calculations in solving the problems, used rough tables and matrices if calculations were needed, and expressed their reasoning and procedures with ease.

Spearman ρ values as described above were reported with significant ($p < .05$) correlations indicated.

Interpretations

The author outlined, by analyzing the correlation coefficients, which behaviors and strategies were effective and which were counterproductive for each problem. Large, positive ρ's indicated an effective behavior as a result of high correlation with EI rankings for that problem. Large, negative ρ's likewise indicated counterproductive behavior. Mandell concluded that (1) audiotaping appears to be an effective method for discovering how subjects solve problems, (2) use of pre-designed behavior classification schemes makes possible the discovery of problem-solving behaviors and strategies of individuals and groups, and (3) statistical analysis of these data can provide evidence of common problem-solving strategies. He also concluded that, since all ten multiply-nominated subjects were in the SPS group, at least the three teachers involved in this study can recognize superior problem solvers among their students.

ABSTRACTOR'S ANALYSIS

Mandell's attempt to determine sixth-graders' problem-solving strategies is admirable, but perhaps too ambitious. One gets the impression that he was trying to perform an all-encompassing study of problem-solving behavior without sufficient thought as to what it was he might be measuring. The reader is asked to accept an extensive classification scheme without much justification for the categories or even explanations as to the meaning of the cagegories. A large volume of psychological problem-solving literature suggests that a less ambitious study of particular behaviors and the implications of those behaviors might prove more fruitful.

The above comments aside, several aspects of the report made interpretation of the study difficult. Subjects were described as sixth-graders chosen by their teachers, with no mention of type of school attended, age, or socio-economic status. Without such information, generalization to a larger population is impossible.

The classification of students into SPS and NPS groups was quite ambiguous. The SPS group was initially described as successful problem-solvers within each problem, but later in the report the SPS group was described as those who solved one-half or more of all the problems. The use of this same label for two different groups made interpretation of the correlation coefficients between EI and behaviors an uncertain task. Were these the ρ's for the successful subjects in each problem or were they ρ's for the 48% of the subjects classified as SPS based on their overall performance?

The eleven behavioral categories used were explicitly stated, but the rather brief sample script with statement classifications provided in the appendix did not seem adequate for the reader to assess how the PBTS was used to categorize statements. The method used for taping subjects' protocols was also not sufficiently clear. In the first problem, subjects were asked to recall their thought processes after solving the problem, but one gets the impression that subjects "thought aloud" throughout the other problems. The distinction between these

two methods is important, for it has been found (Ericsson and Simon, 1984) that verbal reports after solution tend to be summary in nature, omitting some details and unfruitful strategies.

Much of the author's analysis of this study relies on the correlation matrix between EI ranks and behavior rankings. The EI is said to be measuring problem-solving efficiency, but this notion must be seriously questioned. One calculates the EI by dividing the total number of statements uttered by a subject by the total time for solution. Ignoring the difficulties in defining what constitutes a "single statement," consider the following situation: two students solve a given problem in the same amount of time, yet the first uses twice as many statement as the second. The first subject would have twice the EI score of the second, and would be considered less efficient in solving the problem. While realizing that the limiations of speaking speed make this situation somewhat implausible, it should be clear that the EI is measuring to a large degree how <u>verbal</u> a subject is. Less verbal subjects may be thinking through several inefficient procedures and reporting only a few, resulting in low EI scores and classifications as efficient problem-solvers.

A question relevant to this issue which was not addressed by the report is whether or not EI correlated significantly with performance. Are students with low EI scores truly good problem-solvers? Are the behaviors exhibited by low EI subjects behaviors which should be modeled and held in high esteem? Answers to such important questions rest strongly on the decision as to whether or not EI is a valuable measurement of performance. The slow and methodical, yet entirely successful, student is obviously punished by such a classification.

An issue related perhaps more to the entire classification procedure employed in the study arose in Mandell's analysis of a problem involving getting a fox, a duck, and some corn across a river without the duck or the corn being eaten. Statements such as "Can the duck swim across alone." or "He could build a pen for the fox" were classified as nonproductive. These were undoubedly not what the experimenter had in mind, but were certainly within the rules as far as

the explicit statement of the problem was concerned. Are we truly measuring problem-solving ability when practical and effective solutions are called nonproductive? This concern for solving the "textbook" problem while ignoring practical solutions is part of a larger problem of interest to all involved in science education (c.f. Simon and Simon, 1978).

The author's discussion of effective and counterproductive strategies utilized for each problem was disturbing for other reasons. Very few (10) of the reported ρ's were significant at the $\alpha = .05$ level, yet several correlations were used per problem in making inferences as to the effectiveness of behaviors. That is to say, Mandell appeared to be overstepping the limitations of his results in drawing conclusions. This aside, one is left questioning how these analyses could be expanded to apply to problem-solving in general. Were these problems chosen to reflect certain classification of problems? Of what general use is a knowledge of which strategies are effective in solving six particular problems? With the author's explicit reference of this work to science education, it would perhaps be more interesting to see problems studied which more closely resemble tasks one encounters in an elementary science course. Problems such as the fox, the duck, and the corn (a variation on the classic "hobbits and orcs" problem) have been well studied in the psychological literature, and it may be time for people concerned with science education to turn their attention away from such "puzzle" problems to more scientifically oriented situations.

The bulk of papers presented in the problem-solving literature today are similar to Mandell's in that they are concerned with describing existing problem-solving behaviors. This data gathering is an essential step in developing a general theory of problem-solving, but it is perhaps wise to pursue an alternative approach: that of developing a prescriptive model which, when instilled in subjects, allows them to become effective problem-solvers. Heller and Reif (1982) have made considerable progress along these lines in the area of elementary physics problem-solving. Mandell indicates that he is working on devising instructional materials based on his results, but that may be an arduous process given the difficulty of interpreting what he has found.

REFERENCES

Ericsson, K. A. and H. A. Simon. Protocol Analysis: Verbal Reports as Data. Cambridge, MA: MIT Press, 1984.

Heller, J. I. and F. Reif. "Knowledge Structure and Problem-Solving in Physics." Educational Psychologist, 17: 102-127, 1982.

Simon, H. A. and D. P. Simon. " A Tale of Two Protocols." In Clement, J. and Lochhead (Eds.), Cognitive Process Instruction, Franklin Press, Philadelphia, 1979.

Maloney, D. P. "Proportional Reasoning and Rule-Governed Behavior With the Balance Beam." Science Education, 67 (2): 245-254, 1983.
Descriptors--*Cognitive Development; *Cognitive Processes; *College Science; Higher Education; *Models; *Physical Sciences; *Problem Solving; Science Education; Student Characteristics

Expanded abstract and analysis prepared especially for I.S.E. by Linda Cronin and Elisabeth Charron, University of Georgia.

Purpose

The primary purpose of this study was to examine the relationship, if any, between college students' proportional reasoning patterns and rule-governed behavior with the balance beam.

Rationale

In recent years, a number of investigators have questioned the validity of Piaget's theory regarding the formal operations stage. It has been postulated that cognitive activities, such as proportional reasoning, should be viewed as a reasoning pattern. Siegler (1976) adopted a strategic, rather than a structural, approach to reasoning and determined that subjects could use one of a sequence of hierarchically ordered rules when dealing with proportional reasoning tasks using the balance beam.

Research Design and Procedure

One hundred eight college students from three different science courses constituted the sample in this exploratory study. The courses included a natural science laboratory class for nonscience majors, a survey of physics for technical majors, and an introductory physics course for science and preprofessional majors. The natural science and

physics survey students were freshmen through juniors with females in the majority. The introductory physics students were primarily junior males.

These students' proportional reasoning patterns and rule-governed behavior with the balance beam were assessed with three instruments:

1.) The fifteen item Lawson Classroom Test of Formal Operations which included four proportional reasoning tasks,
2.) The Fuller Task, a two-part paper and pencil proportionality task,
3.) A rule-assessment packet of 24 balance beam items.

All three of the instruments were administered and evaluated by the author. All of the proportionality tasks required students to calculate a value and explain their reasoning. Items were scored only if both the value and the explanation were valid. Scores on the proportionality task ranged from 0 to 6.

The 24 item rule-assessment packet was designed to classify students according to problem-solving strategy. The strategies are briefly described as: taking account of mass only (Rule 1); considering both mass and distance, but guessing in specified conflict situations (Rule 2); random guessing in all conflict situations, the "muddle" rule (Rule 3); and using the proper product of mass and distance (Rule 4).

An assumption inherent in the design of the rule-assessment instrument was that the general pattern of response to the 24 items will differ according to the rule predominantly employed by a given student. Actual student response patterns were matched against the "ideal" pattern for each rule.

In addition to completing the three instruments, students were also asked to provide written explanations of the procedures they used to complete the rule-assessment task. Scores on the proportionality tasks were computed for each student, and using the criteria previously mentioned for the rule-assessment task, students were each assigned to one of the four rule groups identified by Siegler or assigned to the "Not Consistent" category. Average proportional reasoning scores were then calculated for the students in each rule category.

Findings

The data seemed to indicate a positive relationship between proportional reasoning ability, as reflected in performance on the proportionality tasks, and sophistication of the rule primarily employed with the balance beam items. However, the relationship was not a flawless one. For example, some students with low proportional reasoning scores employed sophisticated strategies on the few items solved correctly. No statistically significant relationships between proportional reasoning scores and the balance beam rules used were reported.

When the proportionality task mean scores for students in each rule group were compared, the proportion mean scores for students using rule 3, and for those using no consistent strategy, were significantly lower than for those using rule 4.

When student explanations of reasoning used in the rule-assessment task were examined, two viable, more specific alternatives to rule 3 were found. These two intermediate rules were incorporated into Siegler's hierarchy and resulted in a new set of six rules. The data were reanalyzed using this revised set of rules. Many students originally classified as muddlers (rule 3 users) were reclassified as users of alternative nonrandom strategies in conflict situations. With this revised rule sequence, significant differences were found between proportion task mean scores for students in rule groups 3, 4, and 6 as well as those in the "Not Consistent" group.

Interpretations

The tenous relationship observed between propo .ional reasoning task scores and the level of sophistication of the balance beam rules used apparently contradicts the relationship indicated by Piagetian theory, i.e., that conflict balance beam problems can only be solved by persons possessing proportional reasoning ability. A possible

explanation for the results obtained may be that students need both knowledge of the behavior of the balance beam and proportional reasoning ability in order to correctly solve balance beam problems. In any event, the results indicate that the subjects studied definitely used some set cf rules to analyze the balance beam. Even those students who did not employ the correct strategies came into the study with some previously formed set of strategies.

This finding carries some implications for physics instruction regarding the balance beam. First of all, since students possess a variety of strategies prior to instruction, they will not all benefit equally from the same instruction. Secondly, students already possessing the correct strategy could benefit from dealing with more complex situations or other matters.

Assuming that the findings of this and other related studies apply to other subjects, it can be concluded that students approach many, if not all, tasks with a set pattern of procedure. If this assumption is true, then the rule-assessment technique has potential as a tool for identifying the pre-existing patterns students bring into courses. This knowledge can be used to enhance student learning, since other research has indicated that m. h of the difficulty experienced with learning new concepts may be associated with interference from old, inappropriate ideas.

ABSTRACTORS' ANALYSIS

This study represents a positive attempt toward investigating alternative systems of evaluating and predicting problem solving ability. However, since the stage of Piagetian development for the students studied was not reported, comparisons could not be made regarding the relationship between rule use and level of problem solving ability as classified by Piaget. Since the author questioned the validity of Piaget's system of classification, it would have been interesting to determine if the rule-assessment system used by the author was better at predicting problem solving ability than were traditional Piagetian classification measures.

The use of student explanations as a supplement to responses on the rule-assessment task proved to be a beneficial additional source of data for analysis of the rule system use. The new six rule system developed as a result of the analysis of these explanations appears to be a more refined and workable system than Siegler's original four rule system. However, the author was rather ambiguous about the meaning of the category "Not Consistent" as it applied to rule use. Since close to one-fourth of the students were classified in that group when the four rule system was used, its meaning should have been clarified.

The information regarding the instruments used for evaluation was somewhat confusing and sketchy in some areas. For example, no information was provided regarding the method used to obtain the "ideal" patterns used for comparison on the rule-assessment task. In addition, no information on the reliability of the three measures was provided. The reader is also left with a significant data gap, i.e., students completed the entire 15-item Lawson test, but only the results of the four proportional reasoning items were reported.

Although the author indicated that the results were not generalizable since the subjects were above the national norm for ACT scores, some of the conclusions generated for the sample group even seem questionable. For example, visual examination of the data does reveal a pattern between increased proportion task socres and more sophisticated rule use, but no statistically significant relationships were noted. In addition, when the data were grouped into the three classes tested, the author concluded that a trend existed representing more sophisticated rule use in more advanced classes. The data presented to support this conclusion should be cautiously interpreted. When the new six rule system was applied to the rule-assessment task results, 27 percent of the students in the "more advanced" science class were in the "Not Consistent" category. Does this mean that it is more difficult to apply the scheme to higher level rule users? Or, does it mean that some of the more advanced students use a mixture of higher and lower order rules?

It would be interesting to explore a number of additional related questions in further research. For example, can the notion of rule use be easily applied to tasks in other areas of physics and other subjects? Can a rule-assessment system be developed for a broad spectrum of problem solving activities and can this broad rule system be used to accurately predict general problem solving ability?

It would also be informative to post-test the students in the two physics classes after they have received instruction in balance beam problems. This information could be used to determine whether or not their reasoning or rule use patterns change as a result of instruction, and if so, how significant or major are the changes?

Finally, this study leaves us with one major question. Can such problem solving rules be taught? If they can, then a strong case exists for trying to develop such schemes and instructing students with them rather than waiting for students to reach some appropriate developmental stage.

REFERENCE

Siegler, R. S. "Three Aspects of Cognitive Development." Cognitive Psychology, 8: 481-520, 1976.