DOCUMENT RESUME

ED 191 683 SE 031 758

AUTHOR Blosser, Patricia E., Ed.: Steiner, Robert L., Ed.
TITLE Investigations in Science Education, Vol. 6, No. 3. Expanded Abstracts and Critical Analyses of Recent Research.
INSTITUTION Ohio State Univ., Columbus. Center for Science and Mathematics Education.
PUB DATE 80
NOTE 67p.
AVAILABLE FROM Information Reference Center (ERIC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 (subscription $6.00, $1.75 single copy).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Abstracts: Individual Differences: Inservice Teacher Education: Research: Research Methodology: *Science Education: *Science Instruction
IDENTIFIERS *Science Education Research

ABSTRACT

This issue of Investigations in Science Education (ISE) provides analytical abstracts, prepared by science educators, of research reports in the areas of individual differences, instruction, cognitive development, and inservice teacher education. Each abstract includes bibliographical data, research design and procedure, purpose, research rationale, and the abstractor's analysis of the research. (CS)

INVESTIGATIONS IN SCIENCE EDUCATION

National Association for Research in Science Teaching

ERIC Clearinghouse for Science, Mathematics, and Environmental Education

Published Quarterly by

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

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

Volume 6, Number 3, 1980

NOTES FROM THE EDITORS iii

INSTRUCTION. 1

Ben-Zvi, Ruth; Avi Hofstein; David Samuel; and Richard F. Kempa. "Modes of Instruction in High School Chemistry." Journal of Research in Science Teaching, 14(5): 433-439, 1977.
Abstracted by WILLIAM G. LAMB. 3

Bowyer, Jane; Benjamin Chen; and Herbert D. Thier. "A Free-Choice Environment: Learning Without Instruction." Science Education, 62(1): 95-107, 1978.
Abstracted by DONALD E. RIECHARD 12

Crocker, Robert K.; Kevin R. Bartlett; and Harry G. Elliot. "A Comparison of Structured and Unstructured Modes of Teaching Science Process Activities." Journal of Research in Science Teaching, 13(3): 267-274, 1976.
Abstracted by DEAN ZOLLMAN 19

Huntsberger, John. "Developing Divergent-Productive Thinking in Elementary School Children Using Attribute Games and Problems." Journal of Research in Science Teaching, 13(2): 185-191, 1976.
Abstracted by MICHAEL J. PADILLA 27

Markle, Glenn and William Capie. "The Effect of the Position of Inserted Questions on Learning from an Activity-Centered Science Module." Journal of Research in Science Teaching, 13(2): 167-1970, 1976.
Abstracted by WILLIAM G. HOLLIDAY. 32

Ryman, D. "The Relative Effectiveness of Teaching Methods on Pupils' Understanding of the Classification of Living Organisms at Two Levels of Intelligence." Journal of Biological Education, 8(2): 219-223, 1974.
Abstracted by DAVID H. OST 39

Tamir, Pinchas. "How Are the Laboratories Used?" Journal of Research in Science Teaching, 14(4): 311-316, 1977.
Abstracted by DAVID R. STRONCK 45

RESPONSES TO ANALYSES. 55

Linn, Marcia C. and David I. Levine. "Research on Logical Reasoning: Issues in Assessing the Role of Familiarity of Variables," by Anton E. Lawson. Investigations in Science Education, 4(4): 20-25, 1978.
Response by DAVID I. LEVINE and MARCIA C. LINN 57

Tamir, Pinchas. "How Are The Laboratories Used?" by David R. Stronck. Investigations in Science Education, 6(3): 45-53, 1980.
Response by PINCHAS TAMIR 61

Willson, Victor L. and Antoine M. Garibaldi. "The Association Between Teacher Participation in NSF Institutes and Student Achievement" by George G. Mallison. Investigations in Science Education, 6(2): 17-26, 1980.
Response by VICTOR L. WILLSON 65

[illegible] THE EDITOR

[illegible] issue of INVESTIGATIONS IN SCIENCE EDUCATION contains articles focu[illegible] instruction and related factors: subject matter taught, the environ[illegible] in which learning is to take place, the effect of structure, the use of games to teach science, the role of written questions, and the use of the laboratory. It also contains responses from several researchers to the critiques of their articles which have appeared in other issues of ISE. We are pleased that science educators are taking the opportunity to respond to questions raised in the "Abstractor's Analysis" portion of the critiques.

This issue also marks a change in the editorial staff. Robert L. Steiner, who has done a most capable job as associate editor, has taken a new teaching position and has moved to Tacoma, Washington. (This move took place in September of 1979 but Bob had helped to produce several issues of ISE before leaving for the West Coast.) Beginning with Volume 6, Issue 4, Victor J. Mayer will become associate editor.

Patricia E. Blosser
Editor

Robert L. Steiner
Associate Editor

INSTRUCTION

Ben-Zvi, Ruth; Avi Hofstein; David Samuel; and Richard F. Kempa, "Modes of Instruction in High School Chemistry." Journal of Research in Science Teaching, 14(5): 433-439, 1977.

Descriptors--*Achievement; *Attitudes; *Chemistry; Evaluation; Instruction; Observation; *Problem Solving; *Psychomotor Skills; Science Education; Secondary Education

Expanded abstract and analysis prepared especially for I.S.E. by William G. Lamb, Frederica Academy, St. Simons Island, Georgia.

Purpose

The purpose of the investigation was to determine the relationships among intelligence, achievement, attitude to and interest in science, and practical laboratory skill in chemistry and to examine possible differences in these variables related to sex and choice of major field (science vs. nonscience). The specific questions listed by the investigators included:

1. Are there distinct modes of performance in high school chemistry?
2. How can one determine the overall achievement in chemistry?
3. Do high school teachers consider a student's ability in practical work when they assess their students?
4. Are there differences in achievement between science and nonscience students?
5. Can laboratory-centered chemistry curricula be recommended for nonscience majors?

Rationale

The study was justified by reference to a statement in Shulman and Tamir (1973) that "research in the relations between laboratory and other learning modes remains scarce." It is related primarily to prior work on assessment of attitudes and laboratory skills by one of the investigators (Kempa). There is no evidence in the report of a theoretical base; the study appears to be what Anderson (1976) calls "direct observation."

Research Design and Procedure

A variety of tests and questionnaires were administered to 233 tenth grade students from 12 general education chemistry classes using the Rehovot curriculum "Chemistry for High School" (Ben-Avi, 1972). The 12 classes were selected as representative of 200 classes taught by 80 teachers who had received extensive inservice training in the curriculum. A multiple-choice "knowledge of science" achievement test (KR-20 = 0.75) was used as a pretest administered prior to beginning the chemistry course. After six months of instruction (administration time specified only for the chemistry achievement test, assumed for others), students completed:

1) An IQ test (split-half = 0.83) (Ortar and Morieli, 1966)
2) A multiple choice chemistry achievement test (KR-20 = 0.78)
3) A Likert-type science interest and attitude inventory (Alpha = 0.72) (Meyer, 1970)
4) A "problem-solving" lab practical with technique, procedure, manual dexterity and orderliness subscales (overall Alpha = 0.70).
5) A "routine manipulative skills" lab practical with subtests similar to the above (overall Alpha = 0.53)
6) An "observational ability" lab practical (Alpha = 0.50 or 0.63, depending on the scoring method used.)

In addition, data were collected for instructor-assigned grades in chemistry, choice of major at the end of the tenth grade year and student sex.

Once data were collected, two factor analyses with varimax rotation were conducted. The first, on the problem-solving and manipulative skills lab practicals—"showed that the four subtests could be reduced to distinct factors: '(I) higher abilities (technique and procedure) and (II) lower abilities (manual dexterity and orderliness).'"

The second analysis—apparently using total scores on the IQ, achievement, attitude and observation lab practical measures, higher ability and lower ability factor scores from the two lab practicals discussed above, and instructor grade—indicated five factors: FACTOR I—a "cognitive-intellectual ability" factor on which IQ, achievement and final grade loaded highly; FACTOR II—a "problem solving" factor on which both higher and lower ability factors from the problem solving lab practical loaded highly (0.81 and 0.84) and pretest general science achievement loaded moderately (0.39); FACTOR III—a "routine manipulative skills" factor on which both higher and lower ability factor scores from the routine manipulative skills lab practical loaded highly; FACTOR IV—an "observational ability" factor containing only the observational lab practical; and FACTOR V—an affective factor containing only the attitude and interest survey.

Once factors were identified, the sample was sorted as science (n = 102) or nonscience (n = 131) majors, descriptive statistics for the five factor scores based on each student's standardized scores on the original measures were generated for each group, and the two groups were compared using a t-test. A similar procedure was used to compare boys and girls.

Findings

Statistically significant differences included: science majors > nonscience majors and boys > girls on the "cognitive/intellectual" factor and girls > boys on the observational factor.

Interpretation

The investigators concluded that achievement on written tests and lab practicals constitute independent modes, that laboratory ability has at least three components—problem-solving ability, manipulative skills and observation, and that all of these should be used in assessing students' ability in chemistry. They found that the teachers used only written test performance for assigning grades. The lack of a difference in achievement on lab practicals indicated that the science curriculum for nonscience majors should be "laboratory-oriented and based at least partially on facing students with problem-solving situations." The superiority of girls to boys in observational abilities was hypothesized as due to different rates of color blindness between the sexes. The cognitive differences were noted but not interpreted.

ABSTRACTOR'S ANALYSIS

The relationship of the study to the matrix of other studies in this area is difficult because of the potpourri of issues addressed. The measurement of laboratory skills appears to be a continuing interest of Kempa whose approach is taxonomic. This approach is typical of others who have addressed the same problem (Doran, 1978). The philosophical, psychological and/or empirical bases of Kempa's taxonomy were not discussed in the report but were referenced. The reasons for expecting (or not expecting, as the case may be) a difference between science and nonscience majors or between boys and girls on the measured variables are not discussed. (Final grades in science courses for the majors is an exception.) That lab work is intrinsically good and lab skills should comprise measured objectives is an undiscussed and unquestioned assumption.

As nearly as can be determined from the report, the study provides no new conceptual or methodological contributions. In fact, the validity of some of the conclusions is questionable on methodological grounds.

The heart of the study revolves around the lab practicals, the best of which is of minimum reliability (Alpha = 0.70 for the overall measure). The observation measure is contaminated to an unknown extent because it requires color discrimination in the blue-green and violet ranges and color-blindness was not determined. More than 50 percent of the variance of each of the measures seems due to measurement error.

The inclusion of IQ in calculating "cognitive" factor scores for group comparisons is also suspect. The investigators seem to be saying that students who major in science do so because they make better grades in science courses because they are more intelligent. It's impossible to tell whether their data support this because, despite their promise in the introduction, IQ is not treated as an independent variable.

I suspect that IQ would be similarly related to achievement in non-science courses. My own experience as a chemistry teacher in high school is that, with very rare exceptions, students who make good grades on written chemistry tests also make good grades on written tests in math, languages, social studies, literature, typing, physical education, industrial arts, and home economics. Some unique feature of Israeli, as compared to American, schooling may create a different state of affairs, however.

With regard to the written report itself, reviewing this study was a frustrating experience because determining what went on from reading the report was quite difficult. In addition to a discrepancy between stated intent in the introduction and actual design as mentioned above with IQ, description of the lab practicals which were apparently developed especially for the study was totally inadequate. The problem-solving test involved an "open-ended" investigation of (the rate of? the stoichiometry of?) the thermal decomposition of $CdCO_3$, with both "planning and actual performance" assessed via checklist along undefined dimensions labeled as technique, procedure, manual dexterity and orderliness. The manipulative skills lab practical involved following "very specific directions" to investigate the precipitation of $PbCrO_4$ (Ksp? Effect of solvent temperature or common ion effect?) assessed using a checklist for technique, procedure, manual dexterity and orderliness.

Technique and procedure hung together and manual dexterity and orderliness hung together during one factor analysis but not another. The results of the two-factor analyses, and my own failure to understand how technique and procedure can differ very much from manual dexterity and orderliness on a test where students follow "very specific directions" puzzle me considerably. Two or three sentences of description for each test and a few items from the checklist(s?) would have helped immensely.

Other tidbits of information which should have been included are: the method used to select the "representative" sample; the time of administration of the IQ, lab practical and attitude measures; the reliabilities of the lab practical subtests; the unit of analysis for the factor analyses; the factor/instrument matrix which indicated "higher" and "lower" abilities; the method of scoring the observation lab practical used for generating data used in the analysis; the nature of the content validation procedure used for the measures developed especially for the study; the number of boys and girls; at least a reference to the algorithm used for calculating the t-test; and the probabilities associated with the boy/girl comparisons. Since the entire report required only six and a fourth pages, all of the additional information mentioned including description of the lab practicals could have been included within the JRST 10-page limit for a research report.

With the reporting of this study, what is the current state of the art of research in this area and what directions are indicated for the future? The state of the art in developing lab practicals has been recently reviewed by Doran (1978) who concluded that "research into psychomotor aspects of science laboratory objectives is woefully lacking." Because the lab practicals developed for this study are not adequately described, it is impossible to determine the contribution of this study to the state of that particular art.

The discrimination between science and nonscience majors is, I suspect, an intractable problem involving at the very least preferred learning style, cognitive style, cognitive preference, rate of development of formal operational thinking, interaction of all these with instruction in both the sciences and nonsciences, encounters with important role models, differential socialization of the sexes, parental expectations and job market structures. It is, at any rate, beside the point as long as the practical consideration is development of a single chemistry course to be taught to all students before they declare majors. I believe the problem of differences between majors and non-majors is relatively unimportant compared with the other problem implied but not discussed in the study: how may laboratory experiences be used to improve classroom instruction in chemistry? (In not qualifying that question, I am assuming that all normal humans learn in pretty much the same fashion.)

If this is the important question, and I believe that it is, then the investigators' stress at this time on psychomotor skills is misplaced. The newer science curricula are lab-centered because this approach is thought to be most effective in helping students understand and learn the concepts and intellectual processes of science. Psychomotor skill development is a means for enabling students to encounter phenomena which will help them acquire such understanding/learning, not an end in itself. My agenda as a high school chemistry teacher is helping students to learn and understand chemistry, not training them to become lab technicians.

The important research problems include, then, the following:

(1) What is the logical structure of chemistry? How do we know whether someone knows and understands the concepts and intellectual processes of chemistry? Which ideas are central and which are peripheral? The methods discussed by Herron et al. (1977) seem useful for initially addressing this problem.

(2) How do students psychologically construct the logical structure of chemistry? (This is quite a different problem because if the logical structure were the same as the psychological structure to begin with, all we'd need for perfect chemistry teaching would be a good set of notes and means of delivering the message to students.) Which ideas may be learned initially and which require prior learning/understanding of other ideas? If human learning is viewed as

information processing as described by Newell and Simon (1972), we can use computer analogies for this second (and most important?) problem. Elucidation of the data base format can probably be accomplished using methods reviewed by Preece (1978). Documentation of the program used to massage the data can probably be accomplished using methods similar to Piaget's. (See Gruen, 1978, for a recent review and a comprehensive bibliography.)

Investigation of problems (1) and (2) should lead directly to some productive hypotheses related to students encountering lab-type experiences as a route for chemistry learning. Only at this point would a consideration of discrete laboratory psychomotor skills required for chemistry learning become broadly useful.

In effect, I am criticizing the investigators' lack of a theoretical base for the research they did. Even despite the methodological problems, the data are of interest. Somehow, Ben-Zvi and her associates have developed a curriculum that does not radically diminish the attitude towards science of those who vote with their feet to not enter science courses again. The factor analyses reported surely mean something. The problem is that without a well-articulated theoretical base, no one can be sure really what the data mean, not even the investigators themselves. Research that takes the stance of direct atheoretical observation (a correlational fishing trip, so to speak) is useful in the absence of organizing theories. But science education has advanced to the point where enough such theories abound to enable us to begin using them. Let us all do so.

REFERENCES

Anderson, O. R. The Experience of Science. New York: Teachers College Press, 1976.

Ben-Zvi, R. (Ed.). Chemistry for High School. Rehovot: Science Teaching Department, Weizmann Institute of Science, 1972. (In Hebrew.)

Doran, R. L. "Assessing the Outcomes of Science Laboratory Activities." Science Education, 62(3): 401-409, 1978.

Green, G. J. "The Theoretical Ideas of Piaget and Educational Practice," in P. Suppes (Ed.) Impact of Research on Education: Some Case Studies. Washington, D.C.: National Academy of Education, 1978, pp. 267-318.

Herron, J. D.; L. L. Cantu; R. Ward; and V. Srinivasan. "Problems Associated with Concept Analysis." Science Education, 61(2): 185-199, 1977.

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.

Newell, A. and H. Simon. Human Problem Solving. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1972.

Ortar, G. and A. Morieli., "Miltha" Intelligence Group Tests. School of Education of the Hebrew University and the Ministry of Education and Culture, 1966. (In Hebrew.)

Preece, P. W. "Exploring Semantic Space: Review of Research on the Reorganization of Scientific Concepts in the Semantic Memory." Science Education, 62(4): 547-562, 1978.

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

Bowyer, Jane; Benjamin Chen; and Herbert D. Thier. "A Free-Choice Environment: Learning Without Instruction." Science Education, 62(1): 95-107, 1978.

Descriptors--*Educational Research; Elementary Education; *Elementary School Science; *Instruction; *Learning; Learning Activities; Open Education; Science Education; *Self Directed Groups

Expanded abstract and analysis prepared especially for I.S.E. by Donald E. Riechard, Emory University.

Purpose

The purposes of this study were:

1. To find out how children operate in a free-choice environment.
2. To discover what children learn about controlling variables in the absence of direct teacher instruction.

The authors stated that it was also of interest to establish baseline data for future studies.

Rationale

The research was concerned generally with student learning when student autonomy is maximized and teacher-controlled instruction is minimized. More specifically, the focus was on the ability of students to control variables. The authors feel that learning to control variables is important not only for science but because such knowledge is also fundamental to everyday decision making.

The part of the study related to controlling variables is referenced to the classic work of Inhelder and Piaget (1958). Free-choice environments are linked to studies by Linn et al. (1976a, 1976b).

Research Design and Procedure

Ninety middle-class sixth graders made up the population for this study. Forty-five children were randomly selected to form the experimental group which would participate in the Science Enrichment Center. The remaining 45 children served as the control group and had no science program during the course of the investigation. The children in the experimental group attended the Center in groups of eight for two 45-minute sessions per week for 12 weeks. There were no significant differences in IQ or age between experimental and control groups.

The Science Enrichment Center was situated in a book-storage room. The Center permitted students a great deal of free-choice in activities selected as well as mode of operation. Fifteen different activities were available at any one time. A total of 50 activities were used in the study. Each activity was accompanied by a student-activity sheet which 1) suggested an interesting activity, 2) described a solution, and 3) provided the learner with a related challenge to solve. Report forms were available for the students at the completion of work with a particular set of equipment. An adult paraprofessional staffed the Center. In general, the paraprofessional functioned as a facilitator and was not involved in initiating conversation with the children.

The following kinds of data were gathered in order to evaluate the effects of the Science Enrichment Center: 1) videotapes of the Center in operation were made during the seventh week, 2) on-site observations of the paraprofessional and the students were made in the fifth and tenth weeks, 3) an attitude survey was conducted in the sixth week, and 4) pretests and posttests were given to determine students' abilities to control variables.

The pretest consisted of two paper-pencil tests--the Spheres test and the Balloon Box test. The posttest was made up of the two paper-pencil tests and an individual interview using an adaptation of the Bending Rods task. Two subgroups were formed in both experimental and control groups. The paper-pencil pretests and posttests were then administered in a counterbalanced order so as to permit evaluation of the effects of the testing sequence both within and between groups.

Findings

The authors provide a rather lengthy narrative on "Results and Discussion." Their narrative can be summarized as follows:

1. Analysis of the videotape and site observations substantiated the fact that children operated in a free-choice, non-teacher-directed environment.
   a. Paraprofessional's contribution to the total amount of talk was only 5 percent.
   b. For the most part, the children assumed their autonomous role with ease.
   c. The use of a reward seemed to increase the number of students completing more autonomous and more cognitively demanding challenges.
   d. About one-third of the children elected to work in pairs or trios rather than alone.
2. The attitude survey revealed that, in general, the students liked the Enrichment Center.
   a. Two-thirds of the students indicated a desire to attend the Center more often than twice a week.
   b. The most frequently mentioned negative aspect was the necessity to complete "Activity Report Forms."
   c. The most positive aspect of the Enrichment Center was "doing the experiments."
3. Pretest and posttest results were analyzed with non-parametric statistics.

a. Pretests revealed no initial differences between the experimentals and controls.

b. Posttests of the experimental group were significantly different from the control group (alpha = 0.01) on both the Spheres and the Balloon Box activities. The experimental group students increased in their abilities to control variables. On these two tests (Spheres and Balloon Box), approximately one-third of the experimental group moved from a state in which they were unable to recognize the necessity for controlling variables to one in which they evidenced some understanding for the need to control experiments.

c. On the "Bending Rods Interview" the results indicated that well over two-thirds of the experimental group (78 percent) understood the necessity for controlling experiments.

Interpretations

Implications are summarized as follows:

1. Sixth-grade, middle-class children can work constructively, efficiently, and with enthusiasm over an extended period of time in an environment which encourages student autonomy.
2. In a free-choice environment, children can learn about controlling variables--an area which requires an advanced level of logical thinking.
3. The fact that a paraprofessional can be directly involved in children's activities is extremely important due to teachers increasingly assuming the role of supervising and directing parent-aides and paraprofessionals in their classrooms.
4. Perhaps it is possible to expose more children to science by using supplemental programs and paraprofessionals in a paradigm similar to that used in this study.

5. The authors conclude that the more interesting question concerns the amount of growth children might make if they were exposed to meaningful teacher-led classroom science and experiences like those provided by the Science Enrichment Center.

ABSTRACTOR'S ANALYSIS

Written Report. It is obvious to a reader of this study that a tremendous amount of information was collected. Among the authors' purposes was an expressed interest to establish baseline data for subsequent studies. Meeting that purpose could well be the most important outcome of the investigation.

The videotapes, on-site observations, and attitude surveys, for example, provide data which are interesting but in-and-of themselves are not always quantifiable into defendable conclusions. The observation that about one-third of the children worked in pairs or trios instead of alone might not lend much to this particular study. However, the observation could form the basis for designing an in-depth study on group interaction in a free-choice environment.

One of the basic problems in a study designed to collect baseline data is encountered with the write-up. Much information is generally collected. Manuscripts submitted to journals, however, typically have rather stringent limits on length. Thus, critical decisions must be made on what to include and what not to include. The authors of this article did a good job of providing appropriate and interesting information. Further, it appears that the study's purposes were met. Since the purposes were stated in very general terms, however, the investigation lacks the "tightness" of an empirical study designed around specifically stated research questions or hypotheses.

The diagrams on the student activity sheet, report form, and equipment used in the written tests help the reader understand the narrative on procedure. It is noted, however, that the equipment designated by the caption for the Spheres test was actually that for the Balloon box test and vice versa.

Research Design and Validity. In order to evaluate a research design, it is necessary to examine the purposes of the study to determine if those purposes can be met by the design employed. Given the general nature of the purposes of this investigation, one must conclude that the design was adequate. If there is a flaw, it is most directly related to the statement of purposes.

The use of a counterbalanced order in pretests and posttests for both control and experimental subgroups was a sound procedure. And given the nature and organization of the data, the authors' use of non-parametric statistics (chi-square and Mann-Whitney U Test) was very appropriate. A different organization of data might have allowed the use of other statistics. The reader interested in non-parametric statistics is referred to the basic text by Siegel (1956).

Efforts were made to establish some degree of validity for the study. This was done primarily through determining interrater reliability in scoring student responses on the written tests. No information is given on the tests' validity and reliability, however. Since the tests used are referenced to earlier studies, one could consult those references for further information.

Future Research. There is a relatively good body of literature on Piagetian-type tasks related to controlling variables. There seems to be very little, however, on learning science in free-choice environments.

The reader might find it difficult to isolate specific ideas for future research from this article. It seems certain, however, that numerous research questions and hypotheses lie hidden in the data collected. One question is the one posed by the authors concerning growth in science among children who have been exposed to good teacher-led science and good free-choice experiences. Interested readers might request additional information and suggestions from the authors.

REFERENCES

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

Linn, M.; B. Chen; and H. D. Thier. "Personalization in Science: Preliminary Investigation at the Middle School Level," Instructional Science, 5, 1976a.

Linn, M.; B. Chen; and H. D. Thier. "Personalization in Science: Can We Teach Children to Interpret Experiments." Paper presented at the Jean Piaget Society Meeting, Philadelphia, 1976b.

Siegal, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

Crocker, Robert K.; Kevin R. Bartlett; and Harry G. Elliot. "A Comparison of Structured and Unstructured Modes of Teaching Science Process Activities." Journal of Research in Science Teaching, 13(3): 267-274, 1976.

Descriptors--*Activity Learning; Educational Research; *Instruction; *Physical Sciences; Science Education; Secondary Education; *Secondary School Science; Teaching Methods; *Teaching Techniques

Expanded abstract and analysis prepared especially for I.S.E. by Dean Zollman, Kansas State University.

Purpose

This study investigated possible interactions between teaching strategies and pupil characteristics. Two teaching strategies, structured and unstructured, were used in an attempt to answer two research questions:

"1. Is there a difference between the two teaching modes in terms of the following:

(i) Achievement of science process objectives
(ii) Preference for teaching mode.

2. Do significant interactions occur between teaching mode and each of the following variables, again with achievement and preference as criterion variables?

(i) Sex
(ii) Intelligence
(iii) Creativity (verbal and figural)
(iv) Personality (extroversion, neuroticism, dependency)
(v) Socioeconomic status."

Rationale

Elementary science curricula developed in the 1960s and 1970s focus on pupil activities which are used to teach both the content and process of science. The term "activities" can be defined in a variety of ways. Pupils who follow a list of explicitly stated steps are performing activities, yet they are performing quite differently from pupils who are given equipment and told to explore anything about a phenomenon. In between these extremes are many variations on the theme.

In teaching the process of science two views of instruction can be clearly defined. The proponents of one view state that the process is a set of skills and operations which can be taught as many other skills and operations are. Consistent with this view are activities which are carefully controlled by the teacher and which lead to process skills by practice. On the other hand, one can view the process of science as open-ended inquiry. Many instructors believe that the most appropriate way to teach this approach to the scientific process is to allow pupils the opportunity for open-ended inquiry into natural phenomena. While undertaking these investigations pupils will, without teacher control, develop their own skills and learn first-hand about the process of science.

In investigating the merits of these two approaches one should also consider general studies of teaching strategies. Much of the research suggests that different types of pupils may prefer and/or learn more effectively from different types of instructional strategies. Variables such as those listed in the research questions stated earlier may interact with the teaching strategies. These interactions can be present in any investigation of teaching strategies.

Crocker, Bartlett and Elliot combined the above ideas and investigated "the effects of different teaching strategies in the context of process-oriented elementary science activities" and "the possible interactions between teaching strategies and the [pupil characteristics] mentioned."

Research Design and Procedure

Four classes of grade-six pupils were the subjects for this study. Each class was given two treatments--one structured(S) unit and one unstructured(U) unit. The units were "modified from a process-based elementary science curriculum developed by one of the authors" and covered the concepts of balancing (B) and density (D). The table below summarizes the treatments.

Class	Week 1	Week 2	Teacher
1	S,B	U,D	1
2	U,D	S,B	1
3	U,B	S,D	2
4	S,D	U,B	2

A 15-item test which determined achievement on understanding the process of science was constructed by the authors and was used as a pre-test and post-test for each unit. The reliability and validation of the tests were determined prior to the study. Independent variables were measured using existing inventories for socioeconomic status, personality, creative thinking and dependence proneness.

Multiple linear regression was "used in a mode essentially equivalent to a series of two-way analyses of covariance with posttest as the criterion, pretest and IQ...as covariates, and treatment of each of the other independent variables, in turn, as predictors."

To measure preference a forced-choice scale was administered to all groups following the treatment.

Frequencies of preference scores were cross-tabulated with each independent variable and chi-square tests "applied to the resulting contingency tables."

Findings

The raw scores showed higher achievement in the structured mode. However the statistical results indicate a strong interaction between the treatment and the class. One class attained a higher overall mean on the achievement tests and performed much better in the unstructured mode. The authors conclude that this class-treatment interaction complicates the answers to the first research question.

The results for preference toward the structured mode were complicated by this same type of interaction.

The examination of the interactions stated in the research questions resulted in the identification of both main effects and interactions. The main effects were IQ, socioeconomic status and sex. The first two of these variables were in the usual direction, while females scored higher than males. Neuroticism displayed an interaction while extroversion, dependency and creativity did not.

Interpretations

One class was clearly different from the other three in this study. Yet, the only measured difference was IQ, which was controlled in the analysis. The implication is that some unmeasured variable was responsible for the observed difference. The investigators speculate that some characteristics of the class as a group may have an influence on the effectiveness of certain teaching modes. These group characteristics "may have some influence in the decision to apply particular teaching strategies."

ABSTRACTOR'S ANALYSIS

The effectiveness of different teaching strategies has been the subject of many studies in science education. However, until recently most investigations have concentrated on the effectiveness of teaching the content rather than the process of science. With curriculum movement of the 1960s investigators and teachers have increased their interest in process and reasoning skills. The research by Crocker, Bartlett and

Elliot is one of several recent efforts to contrast methods of teaching science processes.

Most investigations, including this one, into teaching strategies for the process of science have focused on the amount of teacher control over the learning experience. The basic question is: Can students learn the process of science better by inventing ways of experimenting or by following and practicing steps provided by the teacher? Most investigations provide different teaching strategies and measure gains in understanding the scientific process. In addition to this research Crocker, Bartlett, and Elliot have considered interactions between various student attributes and understanding the process of science. They have also looked at relations between the attributes and preference for structured or unstructured learning experiences. This investigation has, thus, increased the body of knowledge about how teaching strategies are related to learning the process of science.

In selecting an instrument to measure understanding of the scientific process one must be concerned with a built-in bias of the instrument. For example, this investigation measured students' knowledge of the process of science with an achievement test which was created by the investigators and based on an instrument produced by AAAS. The researchers indicated that the reliability and validity of the test were established. However, they do not discuss any bias which the AAAS test may have toward the process skills taught in Science--A Process Approach.

This point could be particularly important because as discussed above the authors divide views of the process of science into "the S-APA sense, as a set of skills and operations" and "relatively unstructured" inquiry. These two views were the basis for the different learning experiences developed especially for this investigation. One must assume that the AAAS Science Process Instrument measures the attainment of skills "in the S-APA sense." Thus, how the authors adapted this instrument for the present study is important to understanding fully the implications of their research. Unfortunately, the published research does not indicate how such a possible bias in the instrument was removed.

The authors state that a wide variety of activities have come to be called teaching by inquiry. Yet, they do not describe the two methods which were used during this investigation. The reader is left with questions concerning the nature of the learning experience. Among these questions are:

Were the structured and unstructured classes given the same amount of time to complete each activity?

How did students in the unstructured classes know the objectives of a lesson?

How did teachers in the unstructured classes respond to student questions which would have added structure to the lesson?

What limitations were placed on the classes? (For example, if students began investigating a phenomenon different from the one intended, were they guided back?)

Answers to questions such as these would provide more detail about the learning experiences and enable readers to interpret better the results and implications of the investigations.

Most investigations, including this one, show that students perform better on process of science inventories after structured rather than unstructured learning experiences. Unfortunately most researchers do not attempt to interpret their results in terms of models of cognitive development. If viewed in terms of these models, this abstractor is not surprised that students with structured experiences gain more understanding of the process of science.

To obtain knowledge of the scientific process from an unstructured activity students must:

1) develop consistent procedures for investigating phenomena, and
2) be able to understand their own thinking processes well enough to build a model of their procedures.

While the former may be possible for some students, the latter is frequently beyond their capabilities. To understand one's own reasoning patterns requires intellectual skills not available to most school children (and, from this abstractor's experiences, to many college students). Without this understanding the students are unable to realize that they have participated in a process and, thus, are not likely to assign their activities to a general model called the process of science.

On the other hand, students participating in structured activities have a model imposed on them. They probably are not told that they are performing process skills. However, these structured activities will not include false starts and seemingly random activities observed in an unstructured experience. The need to invent a process model to separate the useful procedures from the non-productive ones is not necessary. Students in a structured activity need only review how the process works because a model is implicit in everything they do.

These last two paragraphs represent this abstractor's view of the research results into the abilities of structure and unstructured activities to transfer knowledge of the process of science. Further studies are necessary to see how well models of cognitive development fit the data on process skill attainment and retention.

While the results regarding achievement on process skills are not surprising, another implication of Crocker, Bartlett and Elliot's results is. The results implied that some type of interaction among members of one group was far more important than the interaction of each individual's attributes with the treatment. The authors speculate that the group dynamics of one class resulted in much different achievements for this class. This conclusion, if verified, could have important implications for both research and teaching the process of science. It should certainly be investigated further.

Huntsberger, John. "Developing Divergent-Productive Thinking in Elementary School Children Using Attribute Games and Problems." Journal of Research in Science Teaching, 13(2): 185-191, 1976.

Descriptors--*Curriculum; *Divergent Thinking; *Educational Games; Educational Research; Elementary Education; *Elementary School Science; General Science; *Productive Thinking; Science Education

Expanded abstract and analysis prepared especially for I.S.E. by Michael J. Padilla, University of Georgia.

Purpose

The purpose of the study was to test the effects of Attribute Games and Problems on the development of divergent-productive thinking skills in elementary school children.

Rationale

Elementary school curricula are available which allow students to work with materials under a variety of circumstances and problem settings. One such unit is Attribute Games and Problems (AG&P) produced by the Elementary Science Study (ESS). This unit gives children an opportunity to work out different classificatory and organizational solutions to a set of problems which use shapes of different colors and sizes as a medium. The exercises in AG&P stress problem-solving skills and attitudes, one aspect of which is divergent-productive thinking. It is this divergent-productive thinking, defined as the ability to produce diverse solutions when faced with a problem, which is the major concern of the present study.

Research Design and Procedure

Two classrooms of fifth grade students were randomly selected from all the fifth grade classes in one selected school district. One class was randomly selected as the experimental group; the other was designated as the control. From within each class 10 subjects were randomly chosen for inclusion in the study.

Treatment was the only independent variable. Experimental training involved sessions 45 minutes in length over 15 consecutive school days. The control group received regular science instruction over the same time period. This regular instruction was not described. The experimental treatment consisted of 26 activities with AG&P materials which each student completed during the sessions. "The experimental activities were designed to develop mental strategies calling upon divergent-productive thinking modes." The author did not indicate whether these activities were the actual problem cards from the AG&P materials or newly written activities. Nor did he specify whether students worked individually or in small groups; whether the students interacted with each other about solutions; or whether there was any meaningful or important order to the training activities. The experimenter's role was to guide the students without giving answers and to encourage them in their progress.

The Torrence Tests of Creative Thinking, Verbal Form A and Figural Form A, were given to both the experimental and control groups immediately following treatment. The author states that he used a posttest only control group design.

R X O

R   O

Findings

The null hypothesis that no difference existed between the groups was rejected for two of the seven subtests on the creative thinking test. The figural flexibility ($p < .10$) and figural originality ($p < .05$) subtests proved statistically significant in favor of the experimental group using student's t-values. No differences on the other five subtests or on the composite scores were apparent. Scores for each subject on each subtest are also reported.

Interpretations

The use of contemporary science materials such as <u>Attribute Games and Problems</u> may enhance divergent-productive thinking. With a larger sample size and with treatment extended over a longer time period, more significant research results might have occurred.

The present short-term training design may not have been the optimum treatment, especially for determining latent emergence of divergent-productive thinking. Periodic use of attribute blocks over a period of three years in another situation known to the author produced obvious but unrecorded use of these thinking skills.

ABSTRACTOR'S ANALYSIS

Students' abilities to understand science facts or apply scientific principles is a most common focus of research on curricular materials. This study adds an important dimension to the literature by looking at cognitive outcomes from a creativity standpoint. Divergent-productive thinking is an important aspect of the science process skills. Observing, describing, generating hypotheses, identifying variables and interpreting results are production-oriented skills which require creative and divergent thought processes if used properly. Few studies have attempted to measure the ability of science curriculum materials to train this dimension as has this study.

The author claims to have employed a posttest-only control group design which relies heavily upon randomization for control or equivalence of groups. However, true randomization and thus theoretical equivalence of groups was not obtained or established. While randomization took place at each selection step, subjects were not chosen from the same pool. Rather experimentals were selected from one intact class and controls from another. Could any systematic effect due to teacher or the curriculum used in a classroom have affected results? This possibility cannot be ignored. Had the subjects been pretested using the Torrance Tests of Creative Thinking, Form B verbal and figural, equivalence might have been established. The Norms-Technical Manual (Torrence, 1974) suggests this procedure and states that practice effects have not proven to be a problem. Other variables such as IQ might have been used as effective covariates also, thus furthering claims of equivalence.

The stated aims of the outcome measure (Torrance Tests of Creative Thinking) appear to coincide well with this study's desired outcome, divergent-productive thinking. For readers not familiar with this instrument, a brief summary of each of the subtests is a necessity, since a basic understanding of them would be prerequisite for interpreting the two reported significant differences. The author did not provide such a summary. He also did not attempt to interpret these differences beyond a general statement that AG&P might enhance divergent-productive thinking. More important than this assertion might have been his thoughts regarding the relationship between the results and the kind of process skill thinking abilities that science teachers wish to promote. In addition, the results would have proven more meaningful had a retention test been administered. After all, training which produces results is only meaningful if those abilities can be maintained over time. Too, certain divergent-productive thinking skills might only emerge some time after initial training.

One partial (and perhaps flawed) explanation of the results follows. Assuming that students worked individually on the 26 activities and that there was little verbal interaction regarding solutions, it is not surprising that the experimental group's figural and not verbal thinking skills showed improvement. The AG&P materials are figural in nature and it makes sense that they might produce results in this area. Perhaps future work can add a verbal dimension to the materials, if this outcome is desired. Expecting an increase in verbal divergent-productive thinking while not giving students a chance for practicing this skill appears to be a long shot at best.

In summary, this study adds a significant dimension to the literature by investigating an unusual and very important cognitive outcome, divergent-productive thinking. The investigation was flawed by a serious design problem which could have been remedied quite easily. In addition, a serious lack of information and/or description is at times bothersome to the reader. What precisely was the treatment? How does the author interpret the results? What is the meaning behind each subtest on the outcome measure? This reviewer realizes the space constraints placed on authors by journals, but full descriptions of the treatments and results are absolutely necessary to reader understanding.

Lastly this reviewer would urge the author to replicate the study with a larger population, if this has not already been accomplished. In addition, a modification of the treatment so that students can work with the materials over a longer time period is necessary. This can be accomplished by giving the treatment two days per week, interspersed with regular classroom instruction. Verbal practice of divergent-productive thinking should also be included as well as some pretest measure such as the Torrance Tests of Creative Thinking, Form B.

REFERENCE

Torrance, E. Torrance Tests of Creative Thinking. Norms-Technical Manual. Princeton, New Jersey: Personnel Press, Inc., 1974.

Markle, Glenn and William Capie. "The Effect of the Position of Inserted Questions on Learning from an Activity-Centered Science Module." Journal of Research in Science Teaching, 13(2): 167-1970, 1976.

Descriptors--*College Science; *Educational Research; Higher Education; Instruction; Methods Courses; Questioning Techniques; *Science Activities; Science Education; Teaching Methods

Expanded abstract and analysis prepared especially for I.S.E. by William G. Holliday, The University of Calgary.

Purpose

The purpose of the Markle and Capie study was to assess the applicability of inserted mathemagenic questions to an activity-centered module performed by preservice teachers manipulating objects during instruction. This study evaluated the direct (relevant) and indirect (incidental) learning effects of distributing mathemagenic prequestions (questions placed before each activity) and postquestions (questions placed after each activity) among four metric-measurement activities dealing with the "process," "length and area," "mass and volume," and "accuracy and precision."

Rationale

Previous mathemagenic behavior studies have investigated inserted (or adjunct) questions placed either <u>before</u> or <u>after</u> segments of prose in terms of knowledge-recall test items. Very few investigators have reported mathemagenic-question studies using an instructional medium other than typical prose and few have evaluated students in terms of comprehension test items.

Theoretically (see Rothkopf, 1965), prequestions can direct the learner's attention to specific information contained in the text and needed to answer subsequent test items. As a consequence, a question-before-activity treatment should result in higher scores on a subsequent test measuring direct learning of (relevant) information previously asked during instruction. On the other hand, postquestions influence the way the learner reads by reinforcing specific inspection processes previously used by the reader. In this regard, if the student can answer questions at the end of a prose or lesson segment, the learning processes used are reinforced and the student continues using them in subsequent segments. Yet, if the student fails to answer these inserted postquestions, the inappropriate reading processes used are not reinforced and more appropriate processes are likely used during the reading of the next segment. Mathemagenic studies suggest that some postquestions are particularly effective in helping students remember textual information not covered by the inserted questions. As a consequence, a question-after-activity treatment should result in higher scores on a subsequent test measuring indirect learning of (incidental) information not asked for during instruction.

Research Design and Procedure

Twenty-eight preservice female teachers at the University of Georgia were randomly assigned to either a prequestion or postquestion treatment group. Subjects read and performed a measurement module including a rationale, a list of performance objectives and four "enabling" activities. Each activity required about 15 minutes to complete—10 minutes of which was used for such activities as "measuring the length of the room, the volume of water in a jar, the mass of a lead weight, estimating different dimensions of given objects, and estimating the accuracy and precision of measurements." Subjects were informed about the accuracy of their measurements and estimates from cards attached to each evaluated object. The two treatments were identical except for the position of the inserted mathemagenic questions related to each activity. The prequestion group was presented with these multiple-choice questions before each activity, while the postquestion group was presented with the same questions after each activity. In other words, the inserted-question position was the independent variable. The dependent variable was a 40-item multiple-choice test consisting of

20 relevant test items (evaluating the same information as the inserted pre- or postquestions) and 20 incidental items (evaluating information not covered by the inserted questions).

Findings

No significant differences were detected between the two treatment groups in terms of relevant or incidental achievement as indicated by t-test analyses of the data.

Interpretation

The investigators did not demonstrate that a prequestion treatment can result in higher achievement scores on relevant test items while a postquestion treatment can result in higher achievement scores on incidental test items. This latter finding was clearly inconsistent with Rothkopf's mathemagenic model. Perhaps this model is of limited usefulness and is more generalizable to prose and less applicable to manipulate lab-type classroom activities, as cautioned by Koran (1974). These negative results may be attributable to such extraneous variables as student interest in the instructional material and to facilitory self-generated questions about the instructional activities.

ABSTRACTOR'S ANALYSIS

This well-written study represented one of the few efforts to evaluate the positioning effects of inserted questions on science achievement from non-prose instruction. In fact, other investigators working with mathemagenic materials usually compared treatment variations of pre- and postquestions with little expressed interest in the applicability of their experimental findings to learning conditions found in the classroom. Instead, most researchers evaluated verbatim (word by word) inserted questions adjunct to prose passages in terms of verbatim test item performance. Of course, these mathemagenics studies have laid a good foundation for doing related work in applied subject areas like science education. However, few investigators have used these basic research findings as a basis for evaluating their generalizability to science learning materials. Wilson and Koran (1976) describe the few exceptions.

The present study dealt with independent and dependent variables that were operationally defined and logically related to one another. Such careful development of experimental materials by these investigators allowed readers to comprehend the nature of the treatment manipulations and the measured criteria and permitted readers to place the study's findings in perspective with other cited mathemagenic works (all of which were reported in the late sixties).

Reviews (see Anderson and Biddle, 1975; Holliday, 1979; McConkie, 1977) of more recent studies similarly suggested that inserted questions appearing before material had a facilitory effect on relevant achievement but failed to enhance incidental achievement (see Royer, 1978, who described an important theoretical exception to this generalization). On the other hand, recent studies (more recent than those cited in the present article) suggested that adjunct questions appearing after material can have both a direct (relevant) and indirect (incidental) facilitative effect on learning. Yet, Markle and Capie's review of the earlier literature revealed that postquestions usually had little direct learning effect but greater indirect effects relative to prequestion achievement. These researchers accurately predicted (from their analysis of earlier theory and research) that prequestions can result in higher scores on relevant test items, while postquestions can result in higher scores on incidental test items. Today, a more reasonable prediction would anticipate both prequestion and postquestion treatment groups outperforming a no-question control, providing the lesson segments were "properly" structured and the inserted questions were directly related to question-relevant information evaluated by the final achievement test, according to McConkie (1977). In addition, incidental test-item performance would not likely be affected by a prequestion treatment and may or may not be effected by a postquestion treatment.

Markle and Capie did make a contribution to science education by describing how a well-defined set of inserted questions differentially placed in a commonly-used medium influences students' comprehension and by describing how this potentially powerful instructional technique can be further evaluated in terms of today's theory and research. The principal shortcoming of this study was the absence of a no-question control group—the same defect identified in many other mathemagenic studies. Faw and Waller (1976) in their critique of these inserted-question studies described the consequences of this design limitation, "Without some knowledge of the base level of performance to be expected from subjects not receiving an experimental manipulation, the interpretation of performance by treatment groups is mere guesswork." Perhaps they overstated their case. Nevertheless, their criticism was particularly meaningful when evaluating studies reporting "no-differences" between experimental treatments. Therefore, an added control group incorporated into the present study's design would have established in fact whether subjects in either question treatment learned any new information as a result of inserting questions into the activity-centered science module.

Inserted-question research reported during the last decade suggests that these study aids can have a facilitative or inhibiting effect on learning from science classroom materials. Specifically to science education study questions found at the end of textbook chapters sometimes help students to: 1) focus attention on important ideas, 2) comprehend (semantically encode) information, 3) practice in reviewing and retrieving main ideas, and 4) reduce the readability load of the adjunct learning material. Unfortunately, the majority of study questions appearing in popular science books probably fails to effect any of these learning processes, according to a recent descriptive study (Holliday, 1979).

In this regard, future research efforts should take into account: 1) two inserted-question variables likely to influence achievement, and 2) two design flaws found in question studies. First, the two variables likely to facilitate questioning effectiveness are question timing and question wording, according to Anderson and Biddle (1975). Generally speaking, the shorter the time between when students read the "targeted" information and answer the relevant inserted question, the greater the likelihood of higher achievement scores on the final test. Furthermore, the more similar the wording (verbatim style) is between the relevant inserted and test questions, the greater the chance of higher achievement. At this point, a cautionary note is in order. The greater the student's reliance on temporary memory and verbatim encoding of words, the greater the chance of reduced long-term memory effects, according to Threadgill (in press).

Second, few students participating in inserted-questions studies are able to maximize their performance because of the ambiguity of the student task, including instructional directions and lack of practice in answering such questions. In addition, explicit and theoretical relationships among learner aptitudes, inserted questions, adjunct learning material, and criterion measures are seldom adequately described and logically related to one another. Correction of these two flaws in future investigations will help us better identify combinations of variables that will maximize classroom learning and reveal how learners differentially implement and alter their cognitive processes under a variety of classroom conditions.

REFERENCES

Anderson, R. C. and W. B. Biddle. "On Asking People Questions About What They are Reading," in G. Bower (Ed.) Psychology of Learning and Motivation, Volume 9. New York: Academic Press, 1975.

Holliday, W. G.; H. W. Whittaker; and K. D. Loose. "Differential Effects of Science Study Questions." Paper presented at the National Association for Research in Science Teaching meeting, Toronto, April, 1978.

Holliday, W. G. "How to Write Effective Study Questions in Science." Paper presented at the National Science Teachers Association annual meeting, Atlanta, March, 1979.

Koran, J. J., Jr. "Research on Mathemagenic Behavior: Implications for Science Instruction and Learning." Paper presented at the National Association for Research in Science Teaching meeting, Chicago, April, 1974.

McConkie, G. W. "Learning from Text," in L. S. Shulman (Ed.) Review of Research in Education, 5. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1977.

Rothkopf, E. Z. "Some Theoretical and Experimental Approaches to Problems in Written Instruction," in J. D. Krumboltz (Ed.) Learning and the Educational Process. Chicago: McNally, 1965, 193-221.

Royer, P. N. "Effects of Specificity and Position of Written Instructional Objectives on Learning from Lecture." Journal of Educational Psychology, 69: 40-45, 1978.

Threadgill, J. A. "The Interaction of Learner Aptitude with Types of Questions Accompanying a Written Lesson on Logical Implications." Journal of Research in Mathematics Education, in press.

Wilson, J. T. and J. J. Koran, Jr. "Review of Research on Mathemagenic Behavior: Implications for Teaching and Learning Science." Science Education, 60: 391-400, 1976.

Ryman, D. "The Relative Effectiveness of Teaching Methods on Pupils' Understanding of the Classification of Living Organisms at Two Levels of Intelligence." Journal of Biological Education, 8(2): 219-223, 1974.

Descriptors--*Achievement; *Biological Sciences; *Classification; Educational Research; *Instruction; Intelligence; *Learning; *Secondary School Science

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

Purpose

The author's stated purpose was to examine the relative effectiveness of a "Nuffield" approach and a "traditional" approach to the teaching of classification, for comprehensive school pupils of above and below average intelligence.

Rationale

Classification has been the focus of several types of studies over the years. Gagné (1970) suggests that children develop class concepts partly from experience and partly from verbal definitions. Inhelder and Piaget (1964) suggest that concepts of class require more abstract thinking. Kagan, Moss and Sigel (1963) have proposed that two major cognitive styles are involved in classification. Hence, if students do learn classification in different ways and if understanding of classification ranges from strict definition to an abstract appreciation, then, any single method of teaching is not likely to be successful with individuals of different ability.

Research Design and Procedure

Ryman suggests that "the data correspond to those obtained from a randomized blocks design..." and that appropriate analysis of variance techniques can be applied. Using Campbell and Stanley nomenclature it would appear that one could classify the design as either "The Post test-only Control Group" or "The Static-Group Comparison." Neither is a completely accurate label.

The primary criterion in selecting the 12 year old students for the study was that they had studied biology during their first year of studies in mixed ability groups. The selection was further narrowed in that the students had to have experienced either a "traditional" program of study or the "Nuffield" program. Four schools were identified as using the Nuffield course of studies and four which used a traditional approach; two classes were selected at random from each of the four schools.

Eight classes participated in the study. All eight classes had been taught classification by different teachers. The AH4 Group Test of General Intelligence was administered by Ryman late in the academic year. Students scoring above the mean were assigned to the upper group and students below the mean to the lower group. Thus, a double classification was established.

Students were given a classification task which consisted of 38 numbered drawings of a variety of organisms. (Ryman points out that the drawings were similar to those found in the Nuffield Text.) An answer sheet was provided which consisted of a series of rectangles each labeled with a group name. Students responded by writing the numbers of the pictures in the appropriate rectangle. Some were to be classified into more than one category; the spider would, for example, be included in the spider rectangle, the arthropod rectangle and the invertebrate rectangle. Scores were obtained by simply tallying the number of correct responses. The maximum score was 89.

Twenty-five students were randomly identified in each of the four cells resulting from the double classification of type of program -vs- intelligence. Thus, fifty students were sampled from the Nuffield and traditional programs respectively. Similarly, fifty students were representative of the upper and lower intelligence groups.

Table 1

Composition of the Sample

	Method of Instruction		
	Traditional	Nuffield	N
Upper intelligent	25	25	50
Lower intelligent	25	25	50
Normal	50	50	100

The data were reported in three ways: 1) the raw scores of each student; 2) the mean scores of each of the four groups; and, 3) analysis of variance. These summaries as well as the analysis of the data are appropriate for this type of study.

Findings

Ryman's data suggest that the "Nuffield" approach was more effective for upper intelligent students and the "traditional" approach was the more effective for pupils of below average intelligence. The reader is reminded that a maximum score of 89 was possible. The highest mean score of 33 was still less than 40 percent correct.

Table 2

Mean Scores

	Traditional	Nuffield	N
Upper intelligent	28.72	33.00	30.86
Lower intelligent	27.44	17.36	22.40
Overall	28.08	25.18	26.63

The analysis of variance yielded an F = 4.40, with 99 degrees of freedom. Hence the ratio of the mean square for interaction and that for within cells is significant at the 0.05 level. The difference between the achievement of the upper and lower intelligence groups was found to be significant at the 0.01 level.

Interpretations

Ryman emphasizes that any generalizations from the study "can only be applied to the limited population from which the sample was drawn." He suggests that although there are differences in the effectiveness of the two teaching methods nothing specific can be stated due to the significance of interaction. Individuals may differ in how they learn to classify and therefore there is no single method of instruction that will be generally effective. As evidenced by the results of this study the "Nuffield" approach places the lower ability student at a disadvantage; the "traditional" approach does not do the higher ability student justice.

ABSTRACTOR'S ANALYSIS

The primary conclusion made by Ryman is there is "evidence of an interaction between teaching methods and intellectual levels in the development of an understanding of the classification of living organisms by first-year comprehensive school pupils." He wisely makes the point, however, that the results cannot be generalized beyond the populations sampled. Although an attempt was made to randomize the selection of the schools and teachers for comparative purposes, it does not take into consideration teacher attitudes, ability and training. Schools which opted to use "Nuffield" may very well have a different type of teacher. Hence, it may not be the method of instruction that is responding differentially to upper and lower intelligence levels but the type of teacher who selected the materials. Ryman does not characterize the teachers in his study. Similarly, data regarding the schools and the socio-economic areas they serve are lacking.

The notion of "classification" is not well defined in this investigation. It appears that children are tested on their knowledge of the classification system, not on their ability to engage in the intellectual process of classification. This problem is compounded by the fact that illustrations used in the classification task were similar to those in the Nuffield Text. Hence, it would seem probable that students in the sample had an awareness or familiarity which would affect the results.

Kagan, Moss and Sigel (1968), cited by Ryman, suggest that individuals perceive and label their environment by three modes a) the descriptive-analytic, b) the relational-contextual, and c) the categorical-inferential. Persons who employ the descriptive-analytic mode utilize overt physical attributes for classification. Individuals utilizing the relational-contextual mode essentially employ a functional approach to classification, while those with categorical-inferential approaches infer relationships. It would seem that characterizing students in this manner as opposed to I.Q. scores would be more valuable. While it is understood that grouping students according to I.Q. is not uncommon in schools, and that this is frequently interpreted as ability grouping, the weakly-developed relationship of instructional modes to intelligence levels is insufficient to justify the use of I.Q. in this study.

The use of personally developed instruments is frequently a necessity in science education research. Perhaps the tasks Ryman devised for assessing the children's ability to conceptualize are good. However, nothing is said about the development and testing prior to their use in this study. It is unclear as to whether the tasks do indeed evaluate the cognitive process of classification or knowledge of the system of classification. The extremely low scores (a mean of 26.63 out of possible 89) could indicate that the tasks are not valid measures of what students are learning. Or, it may be that the level of cognition necessary to understand classification as being presented is beyond the capability of 12 year olds. Sigel and Hooper (1968) consider the descriptive-analytic style of particular importance to cognitive functioning. A strong case might be made for either interpretations of the low scores.

Many of the criticisms and concerns expressed above may be a function of the article's brevity. The researcher may indeed have been more thorough than the reader is led to believe. Although it is doubtful that anyone would question the general conclusion of the author, "The nature of the whole teaching approach has to be chosen with considerable care, and matched to the particular needs of individuals or small groups," the study does little to move science education towards that end.

REFERENCES

Kagan, J.; H. A. Moss; and I. E. Sigel. "Psychological Significance of Styles of Conceptualization," in J. C. Wright and J. Kagan, (Eds.) Basic Cognitive Processes in Children: Report of the Second Conference on Intellective Development, Vol. 28, Monographs of the Society of Research in Child Development, No. 86, pp. 73-112, 1963.

Sigel, I. E. and F. H. Hooper (Eds.). Logical Thinking in Children: Research Based on Piaget's Theory. New York: Holt, Rinehart and Winston, Inc., 1968.

Tamir, Pinchas. "How Are the Laboratories Used?" Journal of Research in Science Teaching, 14(4): 311-316, 1977.

Descriptors--*Biology; College Science; Higher Education; *Instruction; *Interaction Process Analysis; *Laboratories; Medical Schools; *Science Education; Secondary Education; Secondary School Science

Expanded abstract and analysis prepared especially for I.S.E. by David R. Stronck, University of Victoria.

Purpose

The purpose of this investigation was to find whether or not the inquiry level of the laboratory increases as students advance from grade nine to college. Data were gathered to answer the following questions:

1. What kind of experiences do high school biology laboratories offer at different grade levels?

2. To what extent are these laboratories inquiry-oriented?

3. What are characteristic behaviors of inquiry- and noninquiry-oriented teachers in the laboratory?

4. What are the characteristics of different college laboratories?

5. How are college laboratories different from high school laboratories?

Rationale

Most science educators encourage the use of the laboratory as a place to do investigations and inquiry, not merely verifications. In a "typical inquiry laboratory" the students identify the problem and procedures and perform the investigation. In a "typical verification laboratory" the teacher identifies the problem and procedures while the students follow instructions to repeat previously studied work.

The researcher stated: "As students progress in the study of science from elementary through high school to college, their laboratory experiences may be expected to become increasingly more inquiry-oriented."

Research Design and Procedure

A single observation instrument was used in both the high-school biology laboratory lessons and in the laboratories at the college undergraduate level. This instrument consisted of three parts of the Classroom Observation Instrument relevant to the Earth Science Curriculum, i.e., (1) pre-lab, (2) actual lab work, and (3) post-lab. Smith's instrument was expanded by adding six categories to the 19 items of the pre-lab and four categories to the 19 items of the lab-work phase. The 29 items of the post-lab remained unchanged. In each of the three phases, observations of both the teacher's and the students' behaviors were recorded in 60-second intervals by two observers. The agreement level between the two trained observers was 87 percent at the high school laboratories and 82 percent at the college level.

In the Beer Sheba area of Israel 18 high-school biology teachers were randomly selected for observations in a total of 31 laboratories: 15 at the ninth-grade level, 12 at the tenth-grade level and four at the eleventh-grade level. At the Hebrew University, Jerusalem, four different laboratories were observed four times each: two in the first year of studies (chemistry and biology) and two in the second year (histology and physiology). In each laboratory, about 100 college students were working; they were divided into four or five groups, each with a different instructor. Each observation lasted 30 minutes at the University while the high-school situation was observed throughout the period.

All of the data from the observations were collected under the following categories:

Pre-lab: (a) Identifying the problem to be investigated.
(b) Instructions for carrying out the investigation.

Lab work: (a) Identifying the major components of the investigation.
(b) Teacher reacting to students' questions.
(c) Teacher evaluating students' performance.

Post-lab: (a) Analyzing the data.
(b) Interpreting the results.

A single table reported the data as percentages for the average time devoted to each phase and each component in the different laboratories, i.e., for (1) high-school biology, grade nine, (2) high-school biology, grade ten, (3) high-school biology, grade eleven, (4) college chemistry, (5) college biology, (6) college histology, and (7) college physiology.

The researcher does not define the research design of this study. But he does observe: "Only gross comparisons could be made and no attempt was made to employ statistical tests of significance of differences." If we consider the treatment of this study to be the placement in a more advanced grade level, then the study used the pre-experimental design of the Static-Group Comparison. There were no formal means of certifying that the groups would have been equivalent without the treatment. On the contrary because the groups differed widely, the researcher correctly assumed that his analysis of the data should be limited to "only gross comparisons."

Findings

The data demonstrated that in the ninth and the tenth grades, the pre-lab phase takes a quarter of the time; the actual laboratory work, about two-thirds; and the post-lab discussion, about 9 percent of the time. In the eleventh grade, the pre-lab was reduced to 6 percent of the time while the post-lab discussion expanded to 29 percent. The teacher is the dominant figure in the pre-lab phase with little input from the students in grades 9 and 10.

At the college level, there were no post-lab discussions and no student input during the pre-lab discussions. The time devoted to the pre-lab phase was positively related to the complexity of the task and negatively related to the availability of previously prepared guidelines. Chemistry and physiology required approximately twice as much time in the pre-lab as did biology and histology. While the high-school biology laboratories had patterns similar to those in college biology and histology, college chemistry and physiology placed almost two-thirds of the time in the pre-lab discussions.

In the high school laboratories an average of 11 percent of the time was devoted to verification items while 13 percent was given to investigative items. An "investigative index" was calculated by dividing the sum of the inquiry items by that of the verification items. The investigative index for the high-school sample was 1.2. Nevertheless the 18 high-school teachers provided the following range: 7 were "inquiry-oriented"; 8 were "equally using inquiry and verification"; three were "traditional." The investigative indices at the University were relatively low: biology, 0.7; physiology, 0.5; chemistry, 0.4; and histology, 0.3.

Interpretations

In 1974 high-school biology students in Israel who studied the local BSCS adaptation program were involved in investigative laboratory experiences. The level of inquiry gradually increased from the ninth to the eleventh-grade level. On the other hand, college laboratory experiences were traditional and confirmatory.

Most of the college laboratories have a pre-lab phase which is considerably longer than in the high schools. The college histology course had a relatively short pre-lab phase because of the detailed guidelines which were written and video-taped. All of the college laboratories provided detailed directions designed to eliminate any difficulties or unexpected results.

The college laboratories placed strong reliance on the required written reports and had no post-lab discussions. The researcher condemned this practice by writing:

> The post-lab is essential for problem-solving investigative laboratories. Its absence may serve as a strong indication of the traditional verification nature of the college laboratories. ...College laboratories should lead rather than lag behind high school.

ABSTRACTOR'S ANALYSIS

The researcher wrote: "A literature survey has failed to identify studies describing what is actually happening in the laboratories in terms of students' engagement in inquiry probably because the study of the nature of the transactions (activities and interaction of teachers and students) is not an easy one to carry out in the laboratory setting." The abstractor can identify several studies of the type not found by the researcher. The abstractor suspects that the researcher did not search the literature of the late 1960's when the topic of inquiry was intensely discussed. For example, Inquiry Objectives in the Teaching of Biology, published in September, 1969, has an annotated bibliography of 161 items. One example from this list is the article by E. J. Montague and R. M. Ward: "The Development of Problem Solving Abilities in Secondary School Chemistry." The commentary on this article observes: "The unexpected results of this study suggest that students with investigative experience in the chemistry laboratory do not learn to transfer abilities in critical thinking any better than those with traditional experiences or that teachers may not use these approaches with equal effectiveness."

Another example of research studies on inquiry are the 16 articles contained within the University of Texas Publication Number 6720, October 15, 1967, Research and Curriculum Development in Science Education: 1. The New Programs in High School Biology. An article within this book is of special interest; Lehman W. Barnes, Jr. wrote

"The Development of a Student Checklist to Determine Laboratory Practices in High School Biology." The abstractor suggests that this checklist of 65 items is a more appropriate instrument than the researcher's selection of Classroom Observation Instrument relevant to the Earth Science Curriculum.

The researcher attempted to make a new conceptual contribution by contrasting the laboratory activities of high-school biology classes with those of four different college science courses. He argues that the ideal situation is approached by the eleventh-grade biology laboratories which include a relatively large amount of student input within the relatively lengthy post-lab discussions. He cites several authors who advocate the investigative laboratory.

Unfortunately he does not distinguish the level of course between ninth-grade or tenth-grade biology and eleventh-grade biology. The abstractor suspects that introductory biology is taught in grade 9 or 10 but that students in grade 11 are enrolled in a second course in biology. This second course probably attracts those advanced students who are able to do investigative laboratory work with an enthusiasm that generates lively discussions. The relatively lengthy post-lab discussions of the eleventh-grade biology course are substitutes for the written reports required by college teachers. Many teachers at both the high-school and the college level believe that the writing of these detailed, carefully reasoned reports is a necessary part of doing advanced science.

While the researcher advocates the group dynamics of the post-lab discussion, Joseph D. Novak was recommending a different ideal in the book Facilities for Secondary School Science Teaching: Evolving Patterns in Facilities and Programs. He observed: "The study team established two primary criteria for facilities to qualify as exemplary: (1) the facility must provide for easy modification or flexibility, and (2) the facility must allow for increasing individualization of instruction." The recent trend toward individualization restores an emphasis on writing reports and eliminates the post-lab discussions.

College laboratories are more individualized than high-school labs by allowing the students to leave as soon as they have completed the assignments. In most high schools the students are required to remain until the end of the period in the assigned classroom. Perhaps some high-school teachers use the post-lab discussions primarily to occupy the time at the end of a period. Some of the literature and the abstractor do not support the researcher's assertion that the post-lab is "essential for problem-solving investigative laboratories."

The researcher explained that in a typical inquiry laboratory the students identify the problem and the procedures. If this explanation is accepted, then a reasonable conclusion is that typical inquiry laboratories are found in some elementary schools and in graduate research laboratories, but rarely in high schools or colleges. The project Unified Science and Mathematics for Elementary Schools (USMES) involves children in identifying interesting problems and in developing their own plans for solving these problems. On the other hand, almost all high-school and undergraduate college teachers use the laboratories to assist students in understanding important basic concepts. An exception may be the Chemical Bond Approach (CBA), an eleventh-grade project, which gradually requires the students to take the responsibility of devising procedures and finally of defining the problem.

The abstractor suspects that the relatively high "investigative index" for the high-school laboratories is primarily a function of the relatively high level of exchange in discussion, rather than a measure of identifying problems and procedures for the laboratory. The inquiry method used by many BSCS teachers encourages discussion by posing questions which tend to lead the students toward conclusions previously determined by the teacher. Although this method involves the student in defining the problem and the procedures, the students are led to predetermined conclusions with little flexibility in the work which will be done in the laboratory. Such "inquiry" lessons do not seem to fulfill the researcher's definition of the "typical inquiry laboratory." Therefore this study may contain a confused use of the word "inquiry."

Novak, Joseph D. Facilities for Secondary School Science Teaching: Evolving Patterns in Facilities and Programs. Washington, D.C.: National Science Teachers Association, 1972.

Smith, J. P. "The Development of a Classroom Observation Instrument Relevant to the Earth Science Curriculum Project." Journal of Research in Science Teaching, 8: 231-235, 1971.

RESPONSES

TO

ANALYSES

A RESPONSE TO THE ANALYSIS OF

Linn, Marcia C. and David I. Levine, "Research on Logical Reasoning: Issues in Assessing the Role of Familiarity of Variables,"* by Anton E. Lawson. Investigations in Science Education, 4(4): 20-25, 1978.

by

David I. Levine
United States Court of Appeals
Fifth Judicial Circuit
Baton Rouge, Louisiana

and

Marcia C. Linn
Lawrence Hall of Science
University of California

Lawson's abstract and analysis of our article (Linn and Levine, 1978) contains an adequate summary of the work that was conducted, but is misleading in a number of respects.

Summary of Study

The summary of our study is adequate but Lawson's analysis of the rationale is incomplete. Lawson notes the relevance of our research to Piagetian and neo-Piagetian theory but comments, "No further rationale for the study was given except for a brief statement that a clearer understanding of how children solve control of variables is needed" (Lawson, 1978). Lawson failed to note that the sentence immediately after our "brief statement" refers the reader to our published reviews of recent work in the field (Levine and Linn, 1977) which places our empirical work alongside the work of other researchers and theoreticians. Our introduction was purposely telegraphic. Our intent was simply to place our work within the massive field of adolescent reasoning without

*This material is based upon research supported by the National Science Foundation under Grant No. SED77-18914. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

taking up valuable space in an otherwise empirical article. We do take issue with any implication that our work is not grounded in the large amount of work that took place before our study was undertaken.

Analysis of Study

We take issue with Lawson's (1979) critique of our procedure for investigating the hypothesis that familiarity with the variables in a logical reasoning task contributes to task success. Lawson makes three points in criticizing the findings: a) there was probably no real familiarity difference from problem to problem; b) not all problems involved concrete materials; and c) "as the authors themselves acknowledge, ...the method used for presenting the task gave more information about the variables for one task than another" (Lawson, 1978).

Lawson's first point, that probably no real familiarity difference existed, was, we agree, not stringently tested in our study. However, we did select tasks on the basis of our observations in elementary school classrooms in the United States and England. Lawson tries to prove his point by asking and then answering in the negative the question, "are seeds and fertilizer any more familiar than marbles and wires?" However, this is really not the proper question. As we discuss (Linn and Levine, 1978, p. 384), the question concerns familiarity of the tasks as a whole. The seeds problem concerned planting and caring for seeds. In British schools, by the time children have reached the age of our subjects (12-16 years of age), they have had experience growing plants. In contrast, the ramp task was designed to use familiar objects (marbles) in a new but not totally foreign setting. As Figure 1 in our article illustrates, the marbles were released and allowed to hit a target at the bottom of a short ramp. This task involves some variables that the students were familiar with (height of release point) but was not identical to any previous experiences, unlike the seeds task. Finally, the circuit task was intended to be the least familiar and, according to the hypothesis, the most difficult task of all. While we certainly expected children of 12-16 years of age to be familiar with wires and buzzers, our box was set up in such an unusual way that we did not expect any of the children to have encountered a similar device in the past (see our Figure 2). Six wires arranged in pairs came out of the top of a sealed box. We required the children to connect three wires, one from each of the pairs. We anticipated that the children would not have ever encountered any sort of electrical system that apparently required three wires rather than two. Our expectations were borne out, subjectively, by our observations of the children's reactions to the various apparati. The children who were given the circuit problem appeared to be far more curious about how the device worked than were the children who did the ramp problems. As we noted above, we did not test the equipment for familiarity; however, our selections were far more purposeful than Lawson's comment would suggest.

Lawson's second criticism is that the seeds problem was presented orally. Lawson's comment would have much more force if our results had been the opposite than we obtained. For our subjects, the <u>easiest</u> problem was the seeds task. This is contrary to the hypothesis that a problem presented orally, and therefore abstractly, would be more difficult than one presented concretely with materials for the child to observe directly. If anything, we made obtaining results supporting the hypothesis even more difficult than would have been the case had we used concrete materials. Of course, a methodologically perfect study would have systematically tested the abstract/concrete dimension—this must be left to future studies.

Finally, Lawson quoted us as admitting that "...the method used for presenting the task gave more information about the variables for one task than for another" (Lawson; 1979). This, he suspects, was the primary cause of the large differences in the success rates. Once again, we must point out what Lawson chose not to quote. The entire context of the passage that Lawson selectively quotes demonstrates that we were only offering a possible explanation and that we attempted to make the task instructions identical. As we went on to discuss, we do not believe, as Lawson's chosen quotation would indicate, that we have created results that are meaningless artifacts. In fact, subsequent research has verified and elaborated the hypothesis investigated in the reported study (Linn, Notes 1 and 2).

We have no other objections to Lawson's comments about our study. Allow one comment on the adequacy of his abstract and analysis. Lawson criticized our liberal use of abbreviations, sometimes overly concise style and the mixing of the results, discussion, and conclusion sections. These may be valid criticisms. However, as we have pointed out, Lawson himself was guilty of analogous failings. Although writing which would eliminate some of his shortcomings, such as the failure to quote fully and to analyze his objections completely, would have slightly increased the length of the manuscript, they would assist considerably in clarity and fairness to the authors and to the readers of his abstract.

REFERENCES

Inhelder, B and J. Piaget. The Growth of Logical Thinking: From Childhood to Adolescence. A. Parsons and S. Milgram (Trans.). New York: Basic Books, 1958.

Lawson, A. E. Critique of "Adolescent Reasoning: Influence of Questions Format and Type of Variables on Ability to Control Variables." Investigations in Science Education, 4(4), 1978.

Levine, D. I. and M. C. Linn. "Scientific Reasoning Ability in Adolescence: Theoretical Viewpoints and Educational Implications." Journal of Research in Science Teaching, 14(4): 371-384, 1977.

Linn, M. C. and D. I. Levine. "Adolescent Reasoning: Influence of Question Format and Type of Variables on Ability to Control Variables." Science Education, 62(3): 377-388, 1978.

Notes

1. Linn, M. C. "When do Adolescent and College Students Reason?" Paper presented at Sibyl Walcott Terman Conference on Symbol Processing and Elementary Education, Stanford University, October, 1978, ARP-13.

2. Linn, M. C. "Theoretical and Practical Significance of Formal Thought." Paper presented at Society for Research in Child Development Meeting, San Francisco, March, 1979. ARP-15.

IN RESPONSE TO THE ANALYSIS OF

Tamir, Pinchas. "How Are The Laboratories Used?" by David R. Stronck. Investigations in Science Education, 6(3): 45-53, 1980.

by

Pinchas Tamir
The Hebrew University

I am glad that my paper dealing with direct observations of laboratory transactions was chosen for a critical review by David R. Stronck. Together with the authors of I.S.E. I hope that this channel of communication will help to clarify issues related to research and practice in science education. It appears that the role of the laboratory remains a debatable issue. While I find the review interesting, I wish to offer a number of comments. I shall begin with the literature survey.

The abstractor identifies studies which deal with outcomes of inquiry-oriented laboratories (Montague and Ward, 1969) or with results obtained from questionnaires (Barnes, 1967). He suggests that the checklist of 65 items developed by Barnes "is a more appropriate instrument than the researcher's selection of Classroom Observation Instrument relevant to the Earth Science Curriculum." It appears that, for some unknown reason, the abstractor fails to realize that all the studies cited by him, as well as the Barnes checklist, are not dealing with direct observations of classroom interaction. It is this kind of research which has been almost non-existent. Only recently, following the study under consideration, such observation studies began to appear (e.g., Shymansky, 1978). I contend that many more studies of laboratory transactions in a variety of settings are still necessary in order to find out what is actually happening in laboratory lessons and what variables are affecting the learning of students in the laboratory.

Another issue which needs clarification is the role of the post-lab discussion. The abstractor assumes that "the relatively lengthy post-lab discussions of the eleventh grade biology course are substitutes for the written reports required by college teachers." He also suggests that "college laboratories are more individualized" and that "the recent trend toward individualization restores an emphasis on writing reports and eliminates the post-lab discussion." He further claims that "some of the literature and the abstractor do not support the researcher's assertion that post-lab discussion is essential for problem-solving investigative laboratories." It will be nice to obtain specific reference to this literature.

I agree with the abstractor's observation that "many teachers at both high school and college level believe that the writing of these carefully reasoned reports is a necessary part of doing advanced science." Moreover, I believe that laboratory reports are essential even in less advanced science classes. In fact, they are important in any instructional laboratory. Nevertheless, I cannot agree with the contention that laboratory reports are substitutes for post-lab discussions. Post-lab discussions may not be essential for verification laboratories provided that the teacher makes sure, during the pre-lab discussion and the actual work, that the students understand what they are doing. However, in inquiry-oriented laboratories where often unexpected things happen and where students, by definition, do not know for sure what results they will get, naturally the meaning of the results and their interpretation will benefit substantially by class discussion.

Moreover, in an inquiry-oriented laboratory different teams may be engaged in the study of different aspects of the problem under investigation which they are expected to report to the whole class before the sought-for generalizations emerge. Even when the different teams work on a common task, pooling the results may lead to a more valid analysis taking into consideration differences and similarities of replications as well as employing statistical analysis in pertinent events. Post-lab discussion also provides opportunities to both teacher and students to raise questions, to examine conflicts, to relate previous knowledge to the new findings, etc. We may note, in passing, that individualization of learning does not necessarily imply an undisturbed study of the individual student with no interaction with peers. While working with J. D. Novak (1968) running a highly structured audio-tutorial science program, we both realized that a class or small group discussion, which would accompany the individual audio-tutorial lesson, can add significantly to the consolidation of the presented concepts. This realization is even more pertinent with regard to the less structured inquiry-oriented laboratories.

The abstractor suggests that "perhaps some high school teachers use the post-lab discussions primarily to occupy the time at the end of the period." Perhaps some do. However, this argument cannot be seriously accepted as a support for those who are against post-lab discussion. Rather it may point at the need to study carefully what is going on in post-lab discussions and find ways to improve these discussions so that their potential be more fully realized.

I agree with the abstractor that in most schools the laboratories have not been sufficiently used to develop inquiry skills and that problem identification is highly neglected. One reason for this is that most laboratory manuals, even those of presumably inquiry-oriented curricula such as the BSCS or the PSSC, do not require the students to practice these skills (see, for example, Tamir and Lunetta, 1978). Yet, it is again my contention that as students progress in their study of science they should entertain more opportunities to develop their inquiry, problem identification and problem solving skills. Such opportunities need not result in less understanding of important basic concepts as evidenced repeatedly in the evaluation of the Israeli BSCS Adaptation Project (e.g., Tamir and Jungwirth, 1975; Tamir, 1976).

Although I agree that there exists certain dependence on the nature of the laboratory on the specific disciplines, nevertheless the data presented in the original article, as well as informal discussion with students and instructors in the university, clearly substantiate the basic conclusion that many college laboratories are highly traditional, emphasizing verification and often promote a rather routine technical work with little learning of either concepts or inquiry skills. Similar results and similar conclusions are reported by a recent study in the United States (Kyle, 1977).

REFERENCES

Barnes, L. W., Jr. "The Development of Student Checklist to Determine Laboratory Practices in High School Biology," in A. E. Lee (Ed.) Research and Curriculum Development in Science Education, The University of Texas Publication No. 6720, 90-96, 1967.

Kyle, W. C., Jr. "An Analysis of Instructor and Student Behaviors in Introductory and Advanced Laboratories of Five Science Disciplines at the University Level." Unpublished Master's thesis, University of Iowa, 1977.

Montague., E. J. and R. M. Ward. "The Development of Problem Solving Abilities in Secondary School Chemistry." Journal of Research in Science Teaching, 5: 354-356, 1968.

Novak, J. D. Private communication, 1968.

Shymansky, J. A. "Assessing Teacher Performance in the Classroom: Pattern Analysis Applied to Interaction Data." Studies in Educational Evaluation, 4(2): 99-106, 1978.

Tamir, P. "Factors Which Influence Achievement in High School Biology." Journal of Research in Science Teaching, 13: 539-546, 1976.

Tamir, P. and E. Jungwirth. "Students Growth and Trends Developed as a Result of Studying BSCS Biology for Several Years." Journal of Research in Science Teaching, 12: 263-280, 1975.

Tamir, P. and V. Lunetta. "An Analysis of Laboratory Activities in the BSCS Yellow Version." The American Biology Teacher, 40: 353-357, 1978.

IN RESPONSE TO THE ANALYSIS OF

Willson, Victor L. and Antoine M. Garibaldi. "The Association Between Teacher Participation in NSF Institutes and Student Achievement" by George G. Mallison. *Investigations in Science Education*, 6(2): 17-26, 1980.

by

Victor L. Willson
University of South Dakota

The abstract and comments are generally well developed and insightful. Several comments deserve reply, however.

1. In the matter of sampling schools the comment that the sample is not "parametric" is quite obscure. The sample may be biased with respect to urban systems, but it certainly has parameters to define it and was shown by Gullickson and Welch (1972) to be comparable to other samples for the regions considered, which were mainly rural or small town.

2. The use of random elimination of subjects to produce proportionality is well-documented by statisticians as a desirable action. (cf. Glass and Stanley, 1970; p. 439) The principle is that the effects will be confounded in the interactions but that main effects tests will be exact. In the one case where dummy scores were added, Junior High School Science, the result was not changed as the F-test for institute attendance was nonsignificant. The effect of adding the dummy scores (all means) was to reduce Mean Square Error, so that it should have been even easier to detect significance if it were there. There are better estimators for dummy scores using the whole design matrix, but nothing of importance would be changed in any case.

That degrees of freedom were lost by randomly discarding observations is irrelevant since the error term had well over 100 degrees of freedom in three of four analyses and 91 in another, yielding plenty of power.

3. The number of teachers in this study (or any other) does not determine usefulness of the generalization. The representativeness of the sample is the major determiner. How else would Gallup presume to speak for 200 million Americans with a sample of 1200?

4. The fact that teacher achievement is not directly related to student achievement should certainly not surprise readers. Many years of study have been invested in trying to predict student success from teacher characteristics, and teacher knowledge has not been a successful predictor. It was included in this study as the most likely covariate based on the premises of the NSF programs.

   Dr. Mallison has misinterpreted the study when he suggests that teachers' knowledge is not increased. We did not even address that issue, although there is research to indicate teacher knowledge does increase with institute attendance. We are in agreement, however, that the likely effect of the institutes is in pedagogy, as was stated in the original article.

5. Finally, the comment that the differences are insignificant was addressed in detail in an earlier response (to Edward Davis' review, I.S.E. 5(2), 1979). The essential points are:
   a. The scores represent class means, not student means;
   b. The differences are comparable to large standarized test class mean differences;
   c. In absolute percentage a gain of five to ten percent in achievement might well be worth the expenditure of one to three in-service institutes for a teacher when amortized over the teacher's career and over thousands of students.