The sixth paper in the science series of research reviews produced for the Science and Mathematics Education Information Analysis Center concerns research in science education at the elementary school level published in 1968-69. The following topics are reviewed: Effects of instruction in new elementary science programs (Science - A Process Approach, SCIS, ESS); the impact of instruction at pre-school and primary school levels and in intermediate grades; developmental studies without particular instructional programs; teacher education; teacher behavior and characteristics; and instrument development for subsequent research. Eight miscellaneous studies are also reported. Summarizing conclusions are made at the end of each section, and some general conclusions, mainly concerned with weaknesses of design, analysis and reporting of research are made. The bibliography contains 124 citations, mainly journal articles and unpublished doctoral dissertations. (AL)
SMAC/SCIENCE & MATHEMATICS EDUCATION INFORMATION ANALYSIS CENTER

... an information center to organize and disseminate information and materials on science and mathematics education to teachers, administrators, supervisors, researchers, and the public. A joint project of The Ohio State University and the Educational Resources Information Center of USOE.
SCIENCE EDUCATION INFORMATION

REPORTS

RESEARCH REVIEW SERIES - SCIENCE
PAPER 6
A SUMMARY OF RESEARCH IN SCIENCE EDUCATION
FOR THE YEARS OF 1968-69
ELEMENTARY SCHOOL LEVEL

by

James J. Gallagher
Governors State University
Park Forest South, Illinois

THE OHIO STATE UNIVERSITY
ERIC Information Analysis Center
for Science Education
1460 West Lane Avenue
Columbus, Ohio 43221

June, 1971
The Science and Mathematics Education Information Reports are being developed to disseminate information concerning documents analyzed at the ERIC Information Analysis Center for Science and Mathematics Education. The reports include four types of publications. Special Bibliographies are developed to announce availability of documents in selected interest areas. These bibliographies will list most significant documents that have been published in the interest area. Guides to Resource Literature for Science and Mathematics Teachers are bibliographies that identify references for the professional growth of teachers at all levels of science and mathematics teaching. Research Reviews are issued to analyze and synthesize research related to science and mathematics education over a period of several years. The Occasional Paper Series is designed to present research reviews and discussions related to specific educational topics.

The Science and Mathematics Education Information Reports will be announced in the SMAC Newsletters as they become available.
Research reviews are being issued to analyze and synthesize research related to the teaching and learning of science completed during a two-year period of time. These reviews are organized into three publications for each two-year cycle according to school levels—elementary school science, secondary school science, and college science.

The publications are developed in cooperation with the National Association for Research in Science Teaching. Appointed NARST committees work with staff of the ERIC Center for Science Education to evaluate, review, analyze, and report research results. It is hoped that these reviews will provide research information for development personnel, ideas for future research, and an indication of trends in research in science education.

Your comments and suggestions for this series are invited.

Stanley L. Helgeson
and
Patricia E. Blosser
Editors

Sponsored by the Educational Resources Information Center of the United States Offices of Education and The Ohio State University.

This publication was prepared pursuant to a contract with the Office of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgement in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>New Elementary Science Program:</td>
<td></td>
</tr>
<tr>
<td>Studies of the Effects of Instruction</td>
<td>2</td>
</tr>
<tr>
<td>Studies of the Impact of Instruction:</td>
<td></td>
</tr>
<tr>
<td>Pre School and Primary School</td>
<td>13</td>
</tr>
<tr>
<td>Studies of the Impact of Instruction:</td>
<td></td>
</tr>
<tr>
<td>Intermediate Grades</td>
<td>15</td>
</tr>
<tr>
<td>Developmental Studies Without Instruction</td>
<td>19</td>
</tr>
<tr>
<td>Studies of Cultural Influences on Intellectual</td>
<td></td>
</tr>
<tr>
<td>Development and Science Learning</td>
<td>22</td>
</tr>
<tr>
<td>Teacher Education</td>
<td>26</td>
</tr>
<tr>
<td>Teacher Behavior and Characteristics</td>
<td>42</td>
</tr>
<tr>
<td>Instrument Development</td>
<td>46</td>
</tr>
<tr>
<td>Miscellaneous Studies</td>
<td>47</td>
</tr>
<tr>
<td>Conclusion</td>
<td>51</td>
</tr>
<tr>
<td>Literature Cited</td>
<td>53</td>
</tr>
<tr>
<td>Additional References</td>
<td>63</td>
</tr>
</tbody>
</table>
INTRODUCTION

This review covers research related to science instruction at the elementary school level reported during 1968 and 1969. The literary search for the review was performed by the ERIC Information Analysis Center for Science and Mathematics Education, Columbus, Ohio. One hundred twenty-four research reports were included in the bibliography. This compares with 120 in the previous ERIC Center review, *A Summary of Research in Science Education for the Years of 1965 - 1967* (34). Although different criteria may have influenced bibliographic inclusion in this review and the earlier one, these figures provide a rough indication that the quantity of research related to science instruction at the elementary school level has increased. Since the earlier review covered a three-year period, a sizable increase in the yearly volume of research is indicated.

Not all of the reports included in the bibliography were reviewed. Criteria for inclusion were: significance of the question investigated, validity of research design, rigor of data interpretation, and adequacy of reporting. Since all of the criteria are somewhat subjective, the biases of the reviewer undoubtedly entered into judgments.
Studies Involving AAAS - Science--A Process Approach

Dasbach (20) studied the effects of Science--A Process Approach and contemporary mathematics on teachers and pupils. Three school groups were identified: (a) schools using both Science--A Process Approach and contemporary mathematics, (b) schools using only Science--A Process Approach and (c) schools using only contemporary mathematics. Four major questions were asked:

(1) What are the effects of teaching two contemporary programs?
(2) In what ways does the study of contemporary mathematics appear to strengthen, or interfere with, instruction based on Science--A Process Approach?
(3) Does concurrent study of the two programs enable students to apply mathematical competencies in science?
(4) Would reordering mathematics concepts better enable students to use mathematics in science?

The last question, investigated in detail by Kolb (46), is reviewed below. Conclusions on the other three questions are only tentative because of weaknesses in the experimental design which are pointed out by the investigator. In addition, much of the data is not given quantitative treatment. Thus, the reader is unable to reach his own conclusions and must either accept or reject the investigator's conclusions without full knowledge of the data.
Reports reviewed were categorized as follows:

New Elementary Science Programs: Studies of the Effects of Instruction

Studies of the Impact of Instruction: Pre-school and Primary School

Studies of the Impact of Instruction: Intermediate Grades

Developmental Studies without Instruction

Studies of Cultural Influences on Intellectual Development and Science Learning

Studies of Teacher Education

Studies of Teacher Behavior and Characteristics

Instrument Development

Miscellaneous Studies

This categorization, which does not coincide precisely with the ERIC classification scheme, was chosen because the reviewer felt it represented a more functional grouping of the studies reported during 1968 and 1969.

NEW ELEMENTARY SCIENCE PROGRAMS:
STUDIES OF THE EFFECTS OF INSTRUCTION

New materials for elementary science instruction developed by the government supported curriculum projects have provided a stimulus for a variety of studies of the effects and effectiveness of instruction based upon them. Fourteen of these studies are presented.
Kolb studied the effects of two instructional sequences in mathematics on fifth graders' acquisition of quantitative science behaviors. Pupils in eleven classrooms were randomly assigned to two treatments:

1. a mathematics textbook sequence, and
2. an investigator-prepared sequence patterned after Science--A Process Approach.

Both treatments included identical science lessons. Two criterion measures were developed for the study:

1. a 26-item test corresponding to the 26 tasks included in the mathematics sequence, and
2. an 11-item achievement test of the behavioral objectives of the science lessons.

Kolb found that pupils given the mathematics sequence patterned after Science--A Process Approach scored higher on the science achievement test. Thus, Kolb concluded that greater achievement in science can be expected if a carefully planned mathematics sequence precedes the science lessons.

Ayers (3) studied the use of Part 1 of Science--A Process Approach on a pre-primary group ranging in age from two years eight months to five years eight months. The Ayers Science Process Test, comprised of 25 items was used as the criterion instrument. Conclusions are difficult because no pre-test or control group was used. Differences found could be due to maturation and/or experiences other than instruction in Science--A Process Approach and socio-economic status, and between achievement and I.Q. were not statistically significant.
Pierce (73) studied the effects of Head Start experience on subsequent learning in Science--A Process Approach. Thirty-two kindergarten pupils, half of whom had been enrolled in a Head Start program, comprised the test population. Pupils were tested before and after instruction in Science--A Process Approach using a standardized I.Q. test and nine Competency Measures from Part A of Science--A Process Approach as instruments. A delayed post-test using the same instruments was also given three weeks after conclusion of the instructional period. Pupils with Head Start experience scored higher than those who had not been in Head Start prior to instruction. However, after instruction, as well as on the delayed post-test, no differences were found between the two groups. Both groups made significant gains in I.Q. after instruction and low I.Q. children demonstrated achievement equal to that of high I.Q. children.

McGlathery (54) contrasted the effects of instruction in Science--A Process Approach and a local science program on achievement of five and six year olds from different socio-economic backgrounds. Competency Measures served as criterion instruments. Children instructed in Science--A Process Approach scored higher than children in the control group. Five and six year olds scored about equally. Middle-class children scored higher than lower class pupils on tasks requiring verbalization, but no differences were observed on non-verbal tasks.

Partin (70) compared achievement and interest in science between fourth graders taught Science--A Process Approach and a control group
of fourth graders who used a textbook. Criterion measures were the California Achievement Test, the Sequential Tests of Educational Progress--Science, Competency Measures, and an investigator developed Informal Interest Inventory. No differences were found between experimental and control groups on the first two tests. However, boys scored higher than girls in these two tests and on the Informal Interest Inventory. The experimental group scored higher on the Competency Measures and the Informal Interest Inventory. Performance on the Competency Measures was unrelated to I.Q. scores.

Raun and Butts (77) studied the relationship between strategies of inquiry and pupils' cognitive and affective behavioral change. Pupils from grades four through six who had no other inquiry-centered instruction were exposed to five months of experience in Science--A Process Approach. After instruction, pupils were tested on achievement, attitudes, recall, and divergent thinking. Findings showed that performance in observing, classifying, using numbers, and using space/time relations is correlated with those behavior factors associated with intelligence, divergent thinking, science recall, reading, and perception of potency of science.

An apparent contradiction occurred between the results of Butts and Raun and those of Ayers, Partin, and Pierce. Butts and Raun found differences in achievement after instruction related to I.Q., whereas, the others did not observe this. This discrepancy raises a question. Can the difference in results be attributed to treatment differences
arising from the teachers whose classes were studied, or do they arise from methodological differences?

Howe and Butts (38) studied the effects of instruction based on Science--A Process Approach on children's learning of selected concepts of volume. Fourth- and sixth-grade children with approximately two years of experience in Science--A Process Approach and a similar group of students with no experience in this program were (a) pre-tested on their knowledge of displacement of volume and volume as occupied space, as well as with the Learning Hierarchies Test, (b) instructed in concepts related to volume, and (c) post-tested using the same instruments.

Results showed that fourth-grade children who experienced instruction based on Science--A Process Approach scored higher on the volume tasks in the pre-test than children who had not experienced it. All groups experienced in Science--A Process Approach out-scored the inexperienced groups on the Learning Hierarchies Test during the pre-test. No other significant differences were found.

Ransom (76) studied the effect of Science--A Process Approach on creative thinking and performance in classifying and inferring among second-grade pupils. Three groups were studied: (1) eight classes using Science--A Process Approach whose teachers received in-service training and outside support; (2) eight classes in which teachers were using Science--A Process Approach but who did not receive specialized training in the program; and (3) eight classes whose
teachers were using non-process oriented programs. Criterion instruments included the Science Process Instrument (Inferring and Classifying) and the Torrance Tests of Creative Thinking, Figural, Form B. No differences were found among groups' performance on the criterion measures. Significant correlations were found in pupils' scores on the process measures and certain subscores on the Torrance Test.

Ritz (80) compared the effects of Science--A Process Approach and the Frostig Program for the Development of Visual Perceptions on kindergarten children's attainment of reading readiness, visual perception, and process skills. Three criterion measures were used: The Frostig Developmental Test of Visual Perception, the Competency Measures for Groups from Part A of Science--A Process Approach and the Metropolitan Readiness Test. Three treatment groups of eight classes each were used, one instructed in Science--A Process Approach, the second instructed in the Frostig Program and the third instructed in both. Findings showed that instruction using these programs had no effect on the development of reading readiness of kindergarten children. However, no assessment or control group was used.

A related study was done by Sweeters (95).

Studies Involving Science Curriculum Improvement Study

Chalmer (16) studied the effects of Selected Frostig Visual Perception Units on educationally disadvantaged first graders' achievement in the Science Curriculum Improvement Study "Material Objects"
unit. Criterion measures included Frostig Developmental Test of Visual Perception and ten "Material Objects" Student Activity Pages. It was found that the Frostig program enhanced pupil achievement in the "Material Objects" unit in some classes but not in others.

Stafford (92) studied the effect of instruction based on the Science Curriculum Improvement Study first grade program on pupils' attainment of conservation. Sixty first-grade children were pre-tested on six areas of conservation ability and then divided into two groups. Both groups were given identical instructional programs except for science. The experimental group was instructed using the Science Curriculum Improvement Study first grade program for one semester; the control group received instruction based on a textbook. After instruction pupils were re-tested using the same instruments as in the pre-test. Results show that in all six areas of conservation tested, the experimental group scored higher than the control.

Neuman (63) explored the effect of instruction in selected experiences taken from the Science Curriculum Improvement Study "Material Objects" unit on first-grade pupils' attainment of conservation of weight and quantity. Pupils were pre-tested, given eighteen weeks of instruction based on "Material Objects" then post-tested using adapted Piagetian conservation tasks. A control group was pre-and post-tested and given instruction based on a textbook.

No differences were observed between the experimental group and the control group on the pre-test and on the post-test. However, on the pre-test, boys scored higher than girls.
Almy (1) studied logical thinking of second-grade pupils who had received instruction based on Science--A Process Approach, Science Curriculum Improvement Study, and the Greater Cleveland Mathematics Program during kindergarten and first grade. A control group which did not receive instruction based upon any of these programs was also included. Assessment of logical thinking was achieved through individual, "Piaget-type" interviews on conservation, class inclusion, seriation, transitivity of length and multiple classification. Some classroom observations were made to determine if the instructional materials were used in ways intended by the developers. In all, more than 1000 children were studied in six suburban school districts in metropolitan areas near New York and San Francisco.

Some of the specific questions the study was designed to answer were:

(1) Do children who have participated in any of the "new" programs reveal more advanced thinking at the second grade level than children who have not been involved in such instruction?

(2) Do children who have participated in any of the new programs in kindergarten, as well as in first grade, reveal more advanced thinking at the second grade level than children who have participated in such programs only in first grade?

(3) Do children who have participated in programs stressing the actual manipulation, as well as the labeling, of objects reveal more advanced thinking at the second grade
level than children who were in classes where the experience was predominantly of a paper-and-pencil character?

(4) Do children participating in any one program tend to be more advanced in their thinking at the second grade level than are children participating in any other program?

(5) What evidence is there to suggest that variation in the level of thinking attained by the second grade may be associated with variation in the skills of the teachers of the "new" programs? Since the number of teachers involved in teaching any one program was small, and many other factors might also contribute to the variation in the children's performance, the intent here is to look for clues, rather than anticipating definite answers.

The results of the study are not clear-cut. Second-grade children from the control group scored as well as second grade children who had had prescribed lessons in kindergarten and first grade. However, children in these two groups scored better than children who had received only one year of instruction, beginning in first grade. In spite of the equivocal nature of the findings, this study does provide evidence of the nature of young children's thinking and complexities involved in instructing children in this age group.

Study Involving Elementary School Science Project

Klopfer (45) studied the effectiveness and the effects of the Elementary School Science Project astronomy materials on 90 fifth-
grade pupils in the University of Chicago Laboratory Schools. Children were tested before and after a ten week instructional unit based on "Charting the Universe." Effectiveness of instruction was measured by the Charting the Universe Test Forms 207 and 208, which assessed pupil's knowledge of astronomy and specific content included in the instructional sequence. Statistically significant differences were found between pre- and post-test on both general knowledge and specific content. Effects of instruction were assessed by the Test on Understanding Science, Form Ex, and an eight concept semantic differential test. Small but significant improvement was shown in pupils' understanding of the scientific enterprise as measured by the Test on Understanding Science. Three attitudinal factors related to the instructional program and its content were identified: importance, enjoyment, and dynamism. Although pupils' attitudinal responses after instruction were still favorable, a significant decline occurred in some areas between the time of pre-test and the post-test.

Summary

The following are four important findings from these studies:

(1) Socio-economic status and intelligence do not appear to be as strongly related to achievement when instruction is based upon Science--A Process Approach as when instruction is based on text programs.

(2) Impact of these programs on primary grade children's reading readiness and attainment of conservation remains unresolved.
(3) A hierarchically structured mathematics sequence that correlates with the science sequence enhanced attainment of science and mathematics objectives.

(4) Instruction in new science programs resulted in higher scores on tests of creativity and understanding of the nature of science.

STUDIES OF THE IMPACT OF INSTRUCTION:
PRE-SCHOOL AND PRIMARY SCHOOL

Several studies emphasized the impact of instruction on the learning and behavior of pre-school and primary school children as well as on children enrolled in the intermediate grades.

Emrich (27) studied the development of generalized conservation behavior in four year olds. He found that children given feedback (confirmation of response and a trinket) were able to discriminate length and number relationships rapidly and reliably in the presence of a variety of misleading cues. Moreover, they displayed markedly more stable specific transfer than children from whom feedback was withheld. In addition, the children who received feedback were able to demonstrate non-specific transfer to a greater degree than those who had no feedback.

Oerlerick (65) examined the effects of instructing a group of kindergarten children to make observations as a way of learning about different plants and different parts of a plant. Children in the treatment and control groups were pre-tested and post-tested according to an
individual interview schedule designed by the investigator. Both verbal and non-verbal responses were recorded. A test of transfer was part of the interview. Treatment consisted of five lessons from the Iowa Television Science Education Program which focused on observation of plants. Findings: Performance on the test improved for treatment groups; that is, greater notation of properties and more verbal responses were observed after instruction.

Harper, Steffe and Van Engen (35) studied the effects of a twelve lesson instructional sequence on kindergarten and first-grade children's ability to conserve number. Experimental and control groups were pre-tested and post-tested using Steffe's Test of Conservation of Numerousness. It was found that the kindergarten treatment group scored significantly higher than the non-treatment control group on the post-test. No differences were found between boys and girls. First graders already had acquired conservation of number; consequently, the treatment brought about no observable effect.

Bridgham (11) studied the relationship between third graders' understanding of classification and seriation and the learning of electrostatics. The study consisted of a classification seriation pre-test, an 11 lesson audio-tutorial instruction sequence in electrostatics, and a post-test on electrostatics. The results suggested that an understanding of classification and seriation affects the learning of electrostatics. In addition, it was found that mean scores on the electrostatics post-test for all boys, and for girls whose fathers have a technical background were roughly equivalent;
whereas, girls whose fathers have a non-technical background scored lower on this test of learning achievement. It was also noted that girls scored better on recall items, and boys scored higher on items which required simultaneous attention to observable phenomena and the electrostatics model.

Other studies were done by Stephens and Dutton (94) Gladen and Carkin (31), O'Toole (66).

**STUDIES OF THE IMPACT OF INSTRUCTION:**

**INTERMEDIATE GRADES**

In a study of children's curiosity, Jenkins (39) examined the effects of two instructional variables: (1) number of experiences to which children were exposed, and (2) intervening experiences as measured by elapsed duration. Children in 27 fourth grade classes were presented a single lesson involving either six, two, or no demonstrations related to chemical energy transformations. Curiosity was measured with an investigator developed instrument modeled after Berlyne, in which subjects chose one question from each of several three-item pools for which they would most like to acquire answers. This instrument was administered to randomly selected subpopulations immediately after the lesson, the next day, and two weeks later. It was found that curiosity was not positively related to the number of experiences and that curiosity declined the day after the lesson but showed an increase two weeks later. The results also indicated a distinct, positive, relationship between intelligence and curiosity.
Davidoff (21) studied the impact of out-of-school science experiences on sixth-grade pupils' achievement in science and interaction in learning situations between pupils from two different socio-cultural environments. Although the content and nature of the treatment and criterion measure of achievement can only be inferred from the report, pupils provided with out-of-school experiences made greater gains on a test of achievement than a control group of pupils who did not participate in the program. However, the nature of the treatment may be immaterial; results may be due to differences in attitudes between the experimental group which was given a "special" treatment and the control group which was "left behind." Through analysis of interactions among pupils from different socio-cultural environments, Davidoff's results suggest that children will work together in pursuit of a common goal.

Price (75) investigated fifth, seventh, and eighth graders' readiness to gather data empirically by manipulating objects and materials with their hands and to transfer this skill to a problem situation. After one semester of training in the science process under investigation, no significant increase was observed in the use of self-directed manipulation as a strategy for information acquisition. Gifted children showed no higher incidence of empirical data gathering behavior than children of normal range I.Q. However, it was observed that boys tended to manipulate objects to gain information more frequently than girls.

Skinner (90) studied the effects of two instructional television formats ("direct explanation" vs "unanswered questions") and two live-
teacher follow-up methods ("typical discussion" vs "inquiry sessions" patterned after Suchman's Inquiry Training). Fifth-grade pupils were tested before and after treatment with a 46 item investigator-developed, science achievement test which assessed knowledge, comprehension, application, and analysis as described by Bloom. Pupils instructed in the "unanswered questions" format (i.e., a series of unanswered questions posed by the television instructor) demonstrated higher achievement than children who received the "direct explanation" format. No significant differences were attributable to different teacher follow-up procedures.

Spreadbury (91) studied the differential effects of students' participation in inquiry training, watching demonstrations, and performing an experiment on their ability to interpret data. The criterion instrument, developed by the investigator had a reliability coefficient of .535, which sheds doubt on the validity of findings that students who handled apparatus scored significantly poorer than those who did not handle apparatus.

Cohen (18) studied the effects of experience with natural small scale geologic features on children's concepts of fluval geology. Forty-eight fifth and sixth grade pupils were tested before and after instruction using an investigator developed interview based on premises suggested in Piaget's The Child's Conception of the World. Findings of the study indicated that children learned less from these experiences than their teachers expected.
Benson (5) compared the cognitive and attitudinal outcomes of two methods of instruction: lecture-demonstration and pupil-investigatory approaches. Fifth-grade pupils, randomly assigned to experimental and control groups, were taught by the same teachers over a seven month period. Criterion measures were the Stanford Science Achievement Test and two tests developed by Shoresman: Interests and Ideas and Activities. Results indicated the different methods of presentation produced no significant differences in content achievement. Differences in attitude toward science were significant in one school but not in the second school. Data from the interest inventory indicated no significant difference between sample groups of either school.

Boykin (8) studied the effectiveness of four different feedback modes in a programmed sequence used in teaching fourth graders about sound. Pupils were tested before, immediately following, and thirty days after instruction using an investigator developed test. A transfer test was also part of the final assessment. It was found that feedback and review variables did not compensate for ability levels in individualizing instruction. Questions were raised about the effectiveness of review and the ideal spacing of feedback.

Other studies were reported by Downs (25), Thomas and Buell (98), Rosemergy (82), Schnur (85), MacDougall (55), Donaldson (24), and Hagen (32).

Summary

Studies by Emrich, Oelerick, Harper et al., and Price provide evidence that children generally can achieve objectives of instruction if
objectives and educational experiences are consistent with one another. Bridgham's work and that of Jenkins represent areas that deserve further investigation. Generalization from these studies would be risky unless they are replicated. However, more study of relationships between cognitive development and achievement, and between instruction and curiosity is needed because of the potential impact knowledge of these relationships could have on classroom practice.

DEVELOPMENTAL STUDIES WITHOUT INSTRUCTION

Research studies have also been conducted in which children's intellectual development, uninfluenced by any particular instructional program, was appraised. Seven investigators published such studies during 1968-1969.

Raven (79) studied the development of the concept of momentum in 160 children from grades one through three. Tasks were presented which examined children's ability (1) to conserve matter, (2) to use the speed concept, (3) to demonstrate proportional use of matter and speed, each with momentum held constant, and (4) to handle the momentum concept in a situation where the masses of collision carts were visible different and in another situation where the masses of the carts were concealed. Raven explored which of two alternative developmental sequences occurred in the acquisition of momentum: (1) a logical one (conservation of matter → speed → proportional use of mass and speed → momentum), or (2) a psychological one (momentum → conservation of matter → proportional use of mass and speed → speed). A post hoc analysis of
the trend of rank mean differences between tasks indicated that the development of the momentum concept followed the psychological sequence.

A study by Merrill (59) and associates also showed that a logical hierarchy of concepts did not correspond to the sequence in which these concepts were acquired by children in grades two, three, and four.

Lowery and Allen (52) studied children's ability to sort on the basis of visual resemblance. A heterogeneous sample of 120 first grade pupils were administered a Visual Resemblance Sorting Test comprised of tasks requiring matching of two similar geometric shapes with and without clues, matching nongeometric shapes, discriminating between similar and dissimilar figures, and matching figures which are oriented differently or finely differentiated. It was found that first graders were able to distinguish shapes more easily than patterns, and patterns more easily than size. Also pupils could handle a single attribute more easily than two, and two more readily than three. Clues helped first graders in visual resemblance sorting, and geometric shapes were easier for children to sort than nongeometric ones. First graders' abilities for making fine discriminations or identifying identical figures that are oriented differently appear to be undeveloped.

Roberge (81) studied effects of negation on children's deductive reasoning. The investigator tested 263 randomly selected students from southeastern Connecticut using the Paulus-Roberge Class Reasoning Test and the Paulus Conditional Reasoning Test. Results showed the negation in the major premise had marked influence on the reasoning ability
of children. However, no information was given on reliability and validity of the instruments.

Murray (62) studied 62 second, fourth, and sixth grade children's concepts of duration when objects and events were subjected to (a) decay (apparent aging), (b) changes in height, (c) reversal, and (d) affective differences (e.g., enjoyment). Non-conservers of time were those who judged that these irrelevant transformations affected temporal aspects. The shift from non-conservation to conservation occurred between second and fourth grade for most children on most problems.

Kantz (42) studied science concepts of two matched groups of children in 1950 and 1964. Sixty children from grades one, three, and five were interviewed during both time samplings to determine the concepts held. The investigator concluded that children in the 1964 survey appeared to be more accurate in their concepts than the earlier group on 41 terms, less accurate on 20, and similar on seven.

Karplus and Karplus (43) studied abstract reasoning ability used by individuals from grade five through adulthood in solving a puzzle called the "islands puzzle." The problem was presented verbally to several groups. Responses were categorized as (a) pre-logical, (b) transition to abstract logic, or (c) abstract logic. Correct solution of the problem corresponded with use of abstract logic. Results showed a gradual shift of the group median for grade school children, junior high school students, high school students, and adults corresponding to the ordered categories of logical thinking.
Summary

More studies, like the imaginative ones in this category are needed because of their potential for increasing our understanding of aspects of children's intellectual development that here-to-fore have remained unexplored. However, this type of study assesses gross cultural influences on the development of intellect, and may have limited usefulness to educators. Once the status of a child's intellectual development is appraised, educators need to know effective ways of enhancing this development. Thus, if the ingenious questions and techniques devised by these investigators could be used to study ways of fostering intellectual development through instruction, educational research might produce data of significantly greater usefulness.

STUDIES OF CULTURAL INFLUENCES ON INTELLECTUAL DEVELOPMENT AND SCIENCE LEARNING

The assumption that the intellectual development of children is influenced by cultural factors appears to be accepted by educators. Several researchers attempted to identify relationships between some of these cultural factors and children's achievement in science.

Bozarth (9) studied the relationship between fourth grade children's attainment of conservation of liquid and selected pupil characteristics such as socio-economic background, intelligence, and achievement. Pupils were classified as conservers and non-conservers on the basis of Piaget-type conservation tasks. It was found that conservers scored higher than non-conservers on the Lorge-Thorndike Intelligence Test, on verbal and
quantitative dimensions of the Metropolitan Achievement Test, and on the science section of the Stanford Achievement Test. When adjustment was made for differences in intelligence between conservers and non-conservers, significant differences were noted on verbal, quantitative, and science achievement. When pupils were classified as high and low socio-economic status, significant differences favoring the former group were found in ability to conserve liquid, verbal, quantitative, and science achievement, and intelligence. After adjustment for intelligence between the two groups of differing socio-economic status, differences in all achievement dimensions were still observable.

Sams (83) studied differences in conservation of volume among Caucasian and Indian pupils. Tests included the Otis Quick Scoring Mental Ability Test, the Stanford Achievement Test and modifications of Piaget's conservation tasks. When pupils were classified as conservers and non-conservers no differences were found on the two standardized tests. When pupils were grouped by race, no differences were found except on quantitative achievement.

Lepper (50) studied first graders' ability to conserve continuous and discontinuous quantities, number, length, and area using five Piagetian tasks. Subjects included 30 Negro children and 30 white children matched according to the McGuire-White Index of Value of Orientation. An additional 25 white children were pooled with the sixty matched subjects for correlational analysis between success on the Piagetian tasks and the Metropolitan Reading Readiness Test as well as sub-scores from the McGuire-White Index.
Significant differences were found between Negro and white children in three conservation tasks (number, length, and area) and between groups of differing social status for conservation or continuous quality. He also noted that children could be classified into three groups (conservers, non-conservers, and transitionals) according to their performance on the criterion tasks, and that the five tasks gave a coefficient of reproducibility of .91 when subjected to scaleogram analysis with the order being that identified by Piaget as the invariant order of acquisition. Low correlation coefficients were found between success on Piaget Tasks and scores on the Metropolitan Reading Readiness Test and the McGuire-White Index.

Raven (78) studied the development of classification abilities in culturally disadvantaged and middle class children, ages six, eight, and ten. Six classification tasks of differing complexity were used similar to those devised by Inhelder and Piaget: exhaustive sorting, dual class membership, whole is the sum of its parts, conservation of hierarchy, horizontal reclassification, and quantitative inclusion. The results supported the notion that culturally disadvantaged children have reduced capability to analyze categorical relationships. The difference in classification ability increased with age and with complexity of the task.

George and Dietz (30) compared 857 urban primary grade children of low socio-economic background and 705 primary grade suburban children from higher income homes on eight tasks of basic skills: (1) identi-
fying six colors, (2) drawing four shapes, (3) counting in order, (4) identifying common characteristics of a set of objects, (5) telling time, (6) distinguishing left from right, (7) measuring with a ruler, (8) identifying hot and cold readings on a thermometer. Significant differences were found between urban and suburban children studied. The investigators concluded that third graders from lower income homes possessed skills on many tasks similar to those possessed by first graders from higher income homes.

In a study of children's understanding of science concepts and methods used to acquire information, Klein (44) found that children with low socio-economic background suggested use of experimentation and observation less frequently than middle and upper socio-economic level children. Also, low socio-economic status children gave fewer correct responses on the investigator developed criterion instrument. When adjusted for differences in intelligence among the three groups, the lowest scores were observed among the low socio-economic level children.

Summary

The results obtained by Bozarth and Sams appear to be contradictory. Bozarth found that children who demonstrated ability to conserve liquid on Piaget-type tasks scored higher on tests of intelligence and achievement than non-conservers; whereas, Sams found no differences between these two groups on similar criterion measures. The contradictory results may be attributable to differences in sample size and/or
grade level. Sams studied 64 fifth and sixth graders, whereas Rozaroth studied 209 fourth graders.

The remainder of the studies suggest differences between pupils from lower socio-economic levels and those from middle and upper socio-economic backgrounds. Differences range from basic skills studied by George and Dietz to higher order cognitive competencies examined by Raven and Klein. These status studies, like those in the previous group, represent beginning steps. Once the differences among various populations are described, plans overcoming these differences or capitalizing on them need to be developed and subjected to careful research.

TEACHER EDUCATION

Studies Emphasizing Teacher Behaviors

Several studies of different approaches to teacher education were reported. Koran (48) investigated the relative effectiveness of teacher-directed instruction and modeling via video tape with a group of pre-service elementary teachers. As a criterion measure, Koran recorded the number of questions generated by subjects about the content of two Science--A Process Approach lessons used in the study. No differences were found in questioning behavior between pre-service teachers who were instructed about the lessons in the traditional classroom mode and members of a control group who received no instruction related to the lessons. However, significant differences were found in the number of questions generated by subjects who viewed a fourteen minute video tape
of a teacher conducting a science lesson with four elementary school pupils which highlighted specific questioning behaviors.

Steinbach and Butts (93) compared the effects of practice in a micro-teaching situation with elementary school children and with peers. Criterion measures were:

1. the Instrument for Analysis of Science Teaching for analyzing interactions,
2. the Teacher Performance Competencies Scale for assessing teacher competence,
3. the Classroom Observation Rating Form for identifying teaching strategies used, and
4. a Semantic Differential Test for comparing attitudes.

No differences were found between the two groups in development of pupil rapport, clarity of presentation, and attitudes. Differences favoring practice with children included more indirect teaching, higher frequency of questions, and increasing competency in assisting pupils in attaining objectives. Practice with peers, on the other hand, resulted in more overt, silent, teacher activity; more attention to student talk; and more frequent attention to pupil clarification of ideas.

In addition, part of the teachers in each treatment group were given feedback on their teaching and the remainder were not. It was found that teachers who received feedback (a) were better able to gear lessons to pupils' needs, (b) made more logical presentations to
pupils, and (c) showed improved attitudes in their evaluation of teaching.

A shortcoming of this study is the low inter-rater reliability which may limit the reproducibility and generalizability of the findings.

McCormack (53) studied the effects of training in creativity on prospective elementary teachers enrolled in a science methods course. Criterion measures were the Torrance Tests of Creative Thinking, the Science Education Achievement Test, the Self-Evaluation Inventory, and the Course Evaluation Instrument. When contrasted with a control group which received no instruction in creativity, the treatment group scored significantly higher in the fluency, flexibility, and originality dimensions of the Torrance Tests. In addition, a significant negative correlation was found between scores on the Torrance and Science Education Achievement Tests.

Fletcher (29) contrasted the effects of teacher-centered and student-centered approaches in elementary science methods classes. Groups were tested before and after treatment using the Science Education Achievement Test. Two additional post-tests were used: the Self-Evaluation Inventory and the Evaluation of the Elementary Science Methods Course. No significant differences were found between groups undergoing the different treatments.

Masla (56) studied the usefulness of instruction in interaction analysis in changing verbal inquiry patterns of elementary science methods students. Subjects were pre-tested using the Elementary Teachers Science Inventory, an instrument for assessing competency in science processes. A
randomly selected treatment group was instructed in interaction analysis, and then all subjects were videotaped while teaching a lesson in a school with a small group of pupils. When contrasted with the members of the control group, those who had been instructed in interaction analysis asked more open-ended questions, exhibited a greater proportion of unpredictable responses, and obtained higher pupil response ratios. Also, it was found that competency in the processes of science as measured by the Elementary Teachers Science Inventory does not influence the verbal inquiry pattern of prospective teachers.

In another study, Berkeley (6) considered the effects of training in interaction analysis and found that it had a positive effect on prospective teachers' verbal behavior in the classroom.

**Studies Involving Instruction in Content Areas**

Three studies were reported which dealt with various approaches to instructing prospective elementary teachers in the content of science. Siemankowski (89) compared the effects of an "auto-paced teaching process" and conventional course work on pre-service elementary teachers. The process, which included large group instruction, team teaching, programmed teaching, and audio-tutorial teaching, was used to teach introductory material in geology, astronomy, and meteorology. Criterion measures included tests of knowledge of content in each subject area, Test on Understanding Science and the Purdue Master Attitudes Scale. Results indicated that students instructed in the auto-paced teaching process mode achieved higher scores on tests of knowledge of
content and more favorable attitudes than their counterparts who were conventionally instructed. No differences were found on the Test on Understanding Science.

Using his own Test of Concept Application, Dallas (19) found no difference in pre-service elementary teachers' abilities before and after instruction based on a hierarchically structured curriculum.

Parks (69) studied the effectiveness of analysis of scientific research papers as a method for improving prospective elementary teachers' problem solving ability. In addition to their regular methods course curriculum, one group of methods students analyzed five research papers drawn from the history of science. Criterion measures included the Nelson-Denny Reading Test, the Watson-Glaser Critical Thinking Approach, the Anderson Problem Recognition and Attack Skills Test, and an investigator developed laboratory performance test. No differences were found between the experimental and control group.

Studies Involving New Elementary Science Programs

Several studies have been conducted concerning the effectiveness of various approaches to training teachers to use new science materials. Nine studies are reviewed and all deal with preparation of teachers to use Science--A Process Approach or Science Curriculum Improvement Study materials. In broad terms, the studies examine the effectiveness of specified approaches in helping teachers (1) understand the nature, aims, and organization of the elementary science program in question, and (2) conform to a model of teacher behavior that is compatible with the program.
This model involves allowing children autonomy in inquiry, asking questions to stimulate children in thought and productive activity, and shifting from a teacher-dominated to a pupil-centered mode of instruction.

Butts and Raun (15) studied the effects of an in-service program related to Science--A Process Approach. Criterion measures included a semantic differential to contrast attitudes before and after training, direct classroom observation to determine teacher behavior, and the Teacher Process Measure to assess the cognitive effects of the in-service program. Results indicated that:

(1) Scores on the Teacher Process Measure were correlated significantly with the degree of conformity to the desired model of teacher behavior implicit in Science--A Process Approach.

(2) More experienced teachers and those with fewer hours of science showed positive attitudes toward teaching the program.

In another study White, Raun, and Butts (102) found that the timing and organizational structure (released-time vs own time) strongly influenced outcomes of an in-service teacher-education program. Released-time teacher training programs resulted in more favorable attitudes; however, teachers who attended on their own time showed greater gain in knowledge of science. The Investigators suggested that the latter result may be attributable to more concentrated study.
Hall (33) compared the behavior of two groups of second-grade teachers using Science--A Process Approach and another group of second-grade teachers using a textbook program. Two different in-service formats were used for the Science--A Process Approach teachers: (1) A five day summer workshop coupled with bi-weekly visits from a specialist not on the local school staff, and (2) in-service instruction during the school year plus supervisory help from a local curriculum coordinator. Teachers using the textbook program received no special training or help. All teachers were evaluated on the investigator's Instrument for Analysis of Science Teaching.

When compared with classes of teachers who received no special training, classrooms of those initiated during the summer workshop exhibited more teacher and directional statements, more pupil overt activity, more teacher talk, more closed questions, fewer open questions, fewer open statements on the part of pupils, and less extended pupil talk. Classrooms of teachers given in-service training during the school year, when compared with those of teachers receiving no special training, exhibited more teacher and directional statements, more pupil overt activity, more frequent direct motivation and control, fewer pupil open and closed statements, and less extended pupil talk.

Hall's study shows that not all of the outcomes of teacher education are desirable in helping teachers to conform to the theoretical model of teacher behavior implicit in Science--A Process Approach. These less desired outcomes may be artifacts of an organization of instructional materials which necessitates more direction of pupils,
frequent closed teacher questions, and consequent specific answers from pupils than may be required with textbook-based instruction.

Bunsen (13) studied the effects of four different types of training to prepare teachers to use Science--A Process Approach:

1. Lead teachers trained by college educators,
2. In-service teachers trained by these lead teachers,
3. Methods course students trained in a college methods course, and

Teachers taught three processes--Classifying, Using Space/Time Relations, and Communicating--to children in grades one, two, three, and five. The Process Measure for each of the above processes were administered to pupils of teachers in the four different treatments. No significant differences were found among the four groups on Classifying and Using Space/Time Relations Process Measure. No differences were found between in-service teachers and methods course students on the Communicating Process Measure; however, on the Communicating Process Measure, lead teachers' pupils scored significantly lower than pupils of in-service teachers and methods course students.

Because of the experimental design, it is not clear if these results are attributable to the treatment. It is conceivable that the traits which caused lead teachers to be identified as such had a greater influence on pupil learnings than the training received. The design could have been strengthened if better control and sampling techniques had been employed.
Menzel (57) studied the effects of four different procedures for training 110 methods students in two processes of science—measuring and classifying. The four procedures were:

1. active laboratory-oriented work,
2. passive reading-centered study,
3. alternating, active followed by passive, and
4. intermittent, active, passive, active, etc.

A control group was also used but given no instruction in the two processes. Two criterion measures were used, the Sequential Test of Educational Progress and the Process Measure for the two processes.

No difference among treatment groups and control groups was found on the Classifying Process Measure. Some differences were found among groups on the Measuring Process Measure but these differences are difficult to account for in terms of the treatment procedures.

Wilson (103) studied the questioning behavior of 15 teachers trained in the use of the Science Curriculum Improvement Study approach and a matched group of 15 teachers who had not received training in any of the "new" science programs. The criterion instrument was the Teacher Question Inventory developed by Harris and McIntyre in which teachers' questions were categorized in a scheme corresponding, loosely to Bloom's Taxonomy. Findings showed that teachers trained in Science Curriculum Improvement Study used significantly more questions which require demonstration of skill, analysis and synthesis; whereas, the teachers who had not received this training asked more questions that caused pupils to demonstrate recall, recognition, and comprehension. Also, teachers
trained in the Science Curriculum Improvement Study asked 49% more questions than the comparison group.

Wilson and Renner (104) further studied the effects of an in-service training program on subsequent teacher behaviors. Two instruments were used: The Teacher Question Inventory and a device for recording teachers' inclusion of five "essential science experiences" (i.e., observation, measurement, experimentation, interpretation of data, and prediction). In addition to those findings reported by Wilson, the data showed (1) teachers trained in Science Curriculum Improvement Study included essential science experiences in instruction with more than twice the frequency of teachers not trained in this program; (2) they used questions which encouraged a greater development of the ability to think.

Porterfield (74) compared the questioning behavior of eight teachers trained in the Science Curriculum Improvement Study program with that of eight teachers who had not been trained in any of the new inquiry-discovery approaches to science education. Criterion instrument consisted of a modification of the Teacher Question Inventory.

Experimental group teachers used more translation, interpretation, analysis, synthesis, attitude, and value questions and fewer recognition and recall questions. However, a sizable proportion of questions asked by both groups (44% of the questions asked by the control group and 34% of the questions asked by the experimental group) were of the recall type.
Kondo (47) studied questioning behavior of teachers using the "Material Objects" unit developed by the Science Curriculum Improvement Study. Verbal behavior of four teachers was tape recorded during two "invention" and two "discovery" lessons. Tape scripts were analyzed in terms of (1) complexity of questioning behavior, (2) question types, (3) teacher reactions or responses to her own questions, and (4) the sequence of question types. Kondo found that:

1. A teacher's "style" influenced types of questions and the complexity of questioning behavior.
2. The way a lesson was approached by a teacher influenced types of questions used more than the type of lesson per se.

Other related studies were done by Breit and Butts (10), Hoffart (37) and LaShier (49).

Studies Involving Other In-service Programs

Studies of other in-service programs in science for elementary school teachers reported during 1968-1969 can be categorized primarily as descriptive or experimental. Of the nine studies reported in this section, five might be classified as experimental, concerned with the effects of various in-service programs in terms of changes in teacher behavior and/or student achievement.

Hendrickson (36) reported a survey of 241 pupils of nine teachers who had participated in an Experienced Teacher Fellowship Program and 256 pupils of nine "matched" teachers who had not participated in this
type of program. An investigator developed questionnaire was used to obtain information on pupils' perception of classroom activities, teaching techniques, and pupil and teacher attitudes toward science and mathematics. Pupils of the Experienced Teacher Fellowship Program participants indicated that there was less reading in class and that their teachers demonstrated more, let pupils do more experiments themselves, allowed more pupil autonomy, and provided answers to pupils' questions more readily than the non-participant teachers.

Pettersen (72) studied the effects of an in-service program for teachers in grades four through six. The participants along with a non-participating control group were tested on the Prouse Subject Preference Survey, the Sequential Tests of Education Progress Science Test, and the Test on Understanding Science. No differences were found among participants and non-participants on these tests. However, when the pupils of the two groups were tested using similar criterion measures, pupils of teachers who participated in the in-service program scored higher on the tests.

Dean (22) studied the effects of an intensive in-service program in science on intermediate grade teachers and their pupils. Changes in cognitive achievement of teachers were measured using the Test on Understanding Science, Form W. An investigator developed test was used to assess teachers' divergent thinking abilities. Attitudes of both teachers and pupils were measured by a semantic differential. Significant favorable changes in attitudes occurred among both teachers and
their pupils. A relationship was found between attitudinal changes and changes in understanding of science.

Lindberg (51) studied the effects of two in-service instructional modes (lecture-demonstration vs discovery) on elementary teachers' attitudes toward science. Enrollees in a university extension in-service course were randomly assigned to the two treatments. Criterion measures were a questionnaire and attitude scales developed by Dutton and Stephens and by Leake. Attitudes toward science were positive, both before and after instruction. Sixty percent of the teachers taught by the lecture demonstration method and ninety percent of the teachers taught by the discovery method reported increased use of experiments and demonstration in their classrooms. Seventy-two percent of the teachers, including all of those taught by the discovery method, preferred a discovery approach to in-service instruction.

Schmidt (84) studied the impact of an inquiry-centered science education summer institute on the teaching strategies of elementary teachers in science and in social studies. Each of sixteen teachers were observed before and after the institute. Behavioral data from observations were categorized as (1) use of rational powers except recall, (2) use of essential learning experiences, (3) divergent questions, (4) recall and convergent questions.

After the institute, teachers (1) required pupils to use rational powers excepting recall more frequently, (2) employed essential learning experiences more frequently, (3) asked more divergent questions, and (4) asked fewer recall and convergent questions. These changes were also
observed in social studies classes even though the social studies pro-
gram was not designed to be taught through methods of inquiry.

Jenks (40) studied adoption of an innovation by 72 teachers who
participated in an in-service program and 194 of their fellow teachers
in 18 schools who had not participated. He found that:

(1) lower grade teachers tended to adopt more readily,

(2) teachers from smaller schools tended to adopt more
readily, and

(3) faculties of high and low adopting schools differed
in their perception of the actual power which the
principal exerted.

Merkle (58) studied the impact of a leadership workshop dealing
with Science--A Process Approach and the Science Curriculum Improvement
Study on the knowledge, attitudes, and professional activities of
college instructors in elementary science and school supervisory per-
sonnel. Subjects were tested at the beginning and end of the workshop
and during a follow-up session several months after the conclusion of
the workshop. Significant changes were noted in knowledge of the
programs, group-process skills, change-agent skills, and attitudes
toward elementary science. Also increased in-service and pre-service
activities involving the two instructional programs were reported.

Sweetser (97) conducted a descriptive study of the role of twenty-
three Experienced Teacher Fellowship Program participants who returned
to their school systems. It was found that nineteen were promoted to
more responsible positions. All worked informally with other teachers
and shifted from a didactic to an inductive method of teaching. How-
ever, the participants felt that schools' administrators were not utilizing fully their potential for leadership.

Wailes (100) reviewed the development of National Science Foundation Institutes for elementary school teachers from 1959-1967. More than 22,000 elementary school teachers and supervisors attended during the period covered by Wailes' survey. He concluded that many more teachers were influenced by these programs than actually attended. In addition to the direct help to elementary school personnel, these programs helped to make the problems of elementary teachers known to specialists who served as institute staff members, guest speakers, and consultants.

Another study was reported by Ashley (2).

Summary

Wilson, Wilson and Renner, and Porterfield all showed that teachers instructed in Science Curriculum Improvement Study asked more questions requiring demonstration of skill, analysis, and synthesis, and asked fewer recall questions that the teachers not trained in this program.

Hall found that teachers trained in Science--A Process Approach demonstrated behaviors that were quite different from those reported above.

Butts and associates found (a) significant relationship between score on the Teacher Process Measure and confirmity to the desired model of teacher behavior when using Science--A Process Approach in the classroom; (b) positive attitudes toward teaching Science--A Process Approach were associated with released time for in-service activity, more years of experience, and fewer science courses.
A weakness of many studies which contrast the behavior of teachers using a new experimental curriculum with the behavior of teachers using a traditional approach lies in the fact that prior to any treatment teachers using the experimental curriculum may have demonstrated behavior quite different from teachers using the traditional approach. The mere fact that teachers in the former group either chose, or were chosen, to enter the treatment and teach a new program suggests that they might have different attitudes toward children, teaching, and/or the curriculum. To overcome this weakness, pre-treatment observations of teacher behavior would provide base-line data for comparison of treatment effects. This would give added strength to the conclusions one could draw from the data generated.

Studies of in-service programs are, at the same time, impressive and discouraging. The studies demonstrate that behavioral changes do result from in-service activity. Yet two questions emerge: First, are the studies replicable, or do the outcomes of these studies depend on undefined variables such as the enthusiasm and personality of the in-service workshop leader? In studies of the effects of instruction in science on elementary school pupils, teacher variables are either accounted for or minimized through the research design. However, instructor variables frequently are not controlled or accounted for in studies of the effects of in-service programs on elementary teachers. Therefore, the generalizability of data from these studies is questionable.
The second question concerns the cost-effectiveness of the changes in teacher behavior resulting from these studies. In most studies, treatment involved a highly skilled professional working with a relatively small group of teachers over a rather lengthy duration. In some cases, summer institutes, lasting several weeks, comprised the treatment. If one considers the magnitude of the task in bringing about changes in elementary science teaching in even a fraction of the classrooms of this country, most approaches to in-service teacher education reviewed in this paper appear unrealistic. Current methods require too great an investment of human resources in return for behavioral change of a few individuals. As a profession can we not find ways of training teachers which are more effective and efficient than the familiar "specialist-run workshop"? More imaginative use of media and personnel in local schools should be explored.

TEACHER BEHAVIOR AND CHARACTERISTICS

Seven investigators conducted studies emphasizing teacher behavior and characteristics. The characteristics surveyed ranged from science content courses completed to attitudes toward science, mathematics and teaching. The behaviors analyzed were primarily those demonstrated as teachers operated in classrooms.

A study by O'Toole and Chesin (68) surveyed courses in science taken by practicing and prospective elementary teachers. Survey questionnaires were sent to a random sample of 500 teachers employed in elementary schools in Pennsylvania and to 500 senior elementary education
majors in Pennsylvania state colleges. It was found that prospective teachers outranked present teachers in (1) the number of science courses taken, and (2) in the breadth of courses taken in physical, biological, and earth sciences. However, both prospective and practicing teachers fall short of the AAAS-AACTE recommendations of at least 16 hours of college science preparation.

Schwirian (86) studied relationships between personal and professional characteristics of elementary teachers and their attitudes toward science using the Schwirian Science Support Scale. Characteristics examined were: age, religious preference, amount of higher education, amount of college work in science, type of undergraduate institution, number of years of teaching experience, grade level taught, and continuous vs interrupted teaching experience. Results showed age to be strongly related to attitude, with teachers age 40 and under demonstrating more favorable attitudes toward science. Age-related factors such as years in teaching and continuous teaching experience were also related to attitude. Teachers who took more science and teachers who graduated from state supported institutions also demonstrated more positive attitudes on the criterion instrument.

Kane (41) studied the attitudes of prospective elementary teachers using a semantic differential. He found that attitudes toward science and mathematics, and toward teaching science and mathematics were no less positive than attitudes toward other areas of instruction such as language and social studies.
Miller (60) contrasted knowledge of the scientific process of six groups: prospective elementary teachers, prospective secondary science teachers, prospective secondary social studies teachers, high school science teachers, and research scientists. The criterion instrument used was the Welch Science Process Inventory (form D). Prospective elementary teachers scored higher than the high school students, about the same as prospective secondary social studies teachers, and lower than the other groups.

Fischler and Zimmer (28) developed an instrument for observing science teachers which permits an observer to record three major categories of teacher activity:

(1) techniques teachers use to promote learning,

(2) teachers' questions, and

(3) lesson characteristics (e.g., concrete vs. abstract communication, theoretical vs. practical content, and teacher directed or non-directed student activity).

Eccles (26) formulated a strategy for measuring lesson effectiveness and applied it to the study of twelve teachers. Each of the teachers taught the same lesson, under similar circumstances, to four different groups of sixth-grade children. Children were pre-tested given instruction related to the principle of floatation, and then post-tested using instruments developed by the investigator. Overt behavior was recorded by a combination of audio-tape and 35mm still photographs taken at nine second intervals. A Teacher's Class Mastery Score was determined for each of the classes taught. This score was based on the improvement in
pupils' scores from pre-test to post-test. Eccles' findings suggest considerable differences among teachers in helping pupils learn about floatation as well as considerable variation in Teacher's Class Mastery Scores for a teacher with different groups.

Oberlin (64) compared science achievement of 62 college seniors majoring in elementary education and 22 eighth graders enrolled in a university laboratory school using the Standard Achievement Test for grades seven through nine as the criterion instrument. A small difference between the scores for the two groups, and a small correlation between prospective teachers' achievement scores and the number of hours credit in science, led the investigator to conclude that many elementary education graduates are not adequately prepared to teach science.

Summary

Oberlin's findings are somewhat depressing. However, since the study is narrow in scope, one is prompted to ask, first, does the Stanford Achievement Test assess educational outcomes that are of prime importance for prospective elementary teachers? Second, what results would have been obtained if prospective elementary teachers and eighth graders had been compared using other criterion instruments such as those used by Kane and by Miller.

The technique formulated by Eccles and the instrument developed by Fischler and Zimmer offer potential for more detailed study of science teaching in classroom situations.
INSTRUMENT DEVELOPMENT

Instrument development is another area of research in elementary science which needs more careful attention. The following studies place primary emphasis on the development of instruments to be used in subsequent research.

Dietz (23) developed an instrument to measure problem solving skills of first, second, and third graders. This test was designed to measure:

(1) ability to recognize the problem presented;

(2) understanding of science principles needed to solve the problem;

(3) ability to collect data; and

(4) ability to reason "if-then" statements.

Reliability as determined by Kuder-Richardson Formula 20 was .82.

Brown (12) described the Concept Prerequisite and Development Test as being comprised of five sub-tests which are designed to measure early primary children's ability to (1) detect differences, (2) detect similarities, (3) remember logically, (4) attend to a task, and (5) use all these skills in the conceptualizing. The test was standardized on 1184 children in six western and mountain states. Reliability was determined as .78 by the split half method. Results show no significant differences between boys and girls nor between urban and rural children.

Story (95) developed a 47 item non-reading science test and compared it with the Stanford Science Test. Story found that the two tests
measured different areas of achievement; also he observed that the Stanford Science Test was far above the average reading ability of sixth-grade pupils and was measuring reading ability rather than science learnings.

Uhlhorn (99) developed a Science Experience Inventory based on a random sample of science concepts taken from a city curriculum guide for intermediate grades. The instrument was tested on children from grades four through six as well as teachers and student teachers. Initial studies showed that boys have had a significantly greater number of science.

Shah (88) investigated two tests related to the Conceptually Oriented Program in Elementary Science (COPPS): (1) COPES Test of Critical Terms and (2) COPES Test of Science Concepts. Data on reliability and validity were obtained.

Summary

The studies cited in this section provide a beginning in the area of instrument but they are only that: a beginning. More research needs to be done in this area. These studies emphasized problem solving skills, concept development, and achievement. Attitudes of elementary school children as they experience science should also be investigated.

MISCELLANEOUS STUDIES

Eight studies are reported in the following section and are categorized as "miscellaneous studies". All share the common characteris-
tic of involving elementary school science. None was judged to be similar enough to studies cited in previous sections to be placed in any of those portions of the paper.

Berkheimer (7) studied perceived differences between supervisors and teachers who use curriculum materials developed by NSF-sponsored projects and those who use commercially developed materials. Investigator-developed questionnaires were sent to three groups: classroom teachers, science supervisors, and college science educators. It was found that supervisors perceived themselves as executing roles that are supportive of curriculum materials used in their district (i.e., they encourage individual laboratory experiences when using NSF-sponsored materials). It was also found that teachers perceive the supervisor quite differently from the way the supervisor views himself. Also, college science educators perceived less a directive role for supervisors than was perceived by either teachers or supervisors themselves. Supervisors using NSF-sponsored materials and those using commercial materials did not differ in age, training, or professional experience. However, the two groups did differ in their recommendation of individual laboratory experience and use of teacher demonstrations.

O'Toole and Bedford (67) questioned the limitations of the "Dale List of 3000 Words" and the Dale-Chall readability formula for assessing the readability of Elementary science texts. To expand the List of 3000 Words to include common "science" words, the investigators compiled a list of 150 science words appearing in at least three of six selected elementary science textbook series. From this list, three 50-item tests
were constructed. The tests were administered to 720 fourth-grade pupils and it was found that 36 of the words were correctly identified by 60 percent of the children. These 36 words were used to supplement the Dale word list in evaluating the readability of five fifth-grade science texts from series not previously used. The reading difficulty of three of the five texts dropped by one year, one dropped two years and one was unaffected.

Mitchell (61) developed an instrument for observing student behavior and assessing its relationships to interest and teacher-given grades. Findings showed strong relationships between pupils' demonstration of attentive behavior and teacher grades. Also, a strong relationship was found between pupils' attentive behavior and a measure of interest.

Perkes (71) examined the relationship between teachers' techniques of evaluation and student achievement. Of 165 teaching plans examined, it was found that no teacher engaged in pre-teaching assessment of pupil competencies.

The American Association for the Advancement of Science Commission on Science Education developed An Evaluation Model and Its Application: Second Report (87). This monograph describes a model for curriculum evaluation and reports data obtained from assessment of Science--A Process Approach using this model. Section 1 presents the assumptions underlying the theoretical framework upon which the model has been constructed. This is followed by a translation of the model into operations. Included are descriptions of the development of the instrumentation and the fundamental role of behavioral objectives in the devel-
The development of curriculum scope and sequence. The evaluation instruments constructed include immediate assessment measures and longitudinal assessment measures. Section 2 of this document provides a report of the 1964-65 field tryouts which includes (1) the location of the experimental centers, (2) class and school size, (3) pupil mobilities, (4) the data and analysis of data, and (5) a summary of data from the teacher feedback forms. Selected instruments and forms are appended.

Wallace (101) reported results from a study of children's responses to Science--A Process Approach Competency Measure for the commercial edition of Parts A - E. Comparisons are made between results from 1967-68 and the 1968-69 school years. Data on teacher attitudes are also given.

Butts (14) made a comparison of the cost of materials for eight federally supported course content improvement projects and four city school system science program expenditures.

Bailey (4) studied the development of a sequential curriculum in anthropology for grades one through seven. Cline and Hardy (17) studied the integration of elementary mathematics and science curriculums.

Summary

Summarizing a diverse group of studies is a difficult task. More research needs to be done in curriculum evaluation, in pre-teaching assessment of pupil competencies, in the relation of student behavior, student interest, and achievement. These few studies may serve to interest would-be investigators in these and other topics identified in this portion of the research review.
CONCLUSION

In concluding the Summary of Research in Science Education for the Years 1965-67, Elementary School Level, Haney et al. (34) pointed to four weaknesses of the studies they examined. These weaknesses were:

1. lack of precise terminology,
2. study of global variables that are poorly defined,
3. brief treatment periods on unrepresentative populations, and
4. inadequate instruction.

Although studies included in the current review and in the earlier one were not compared directly, the weaknesses identified by Haney and his associates are still observable. In addition, two other faults were noted which are not new or unique to the studies reported during 1968-69. First, quality of reporting is frequently poor. In many cases, reports are incomplete in both logical development and data presentation. More attention needs to be given to technical writing and data reporting by journal editors, article referees, and those who train researchers. The second fault concerns the broad area of experimental design. This was mentioned indirectly by Haney et al. In some instances, rigorous logic has not been applied during research planning phases. Experimental procedures sometimes were productive of data that were not logically consistent with the hypotheses being tested. A frequent error involved the use of inappropriate statistical treatments of data with consequent deduction of inappropriate conclusions. Here, too, greater rigor needs to be applied by editors, referees, and trainers of research personnel.
In addition, people who conduct research must attend to research design with greater care. Many weaknesses in design could be overcome if researchers would solicit assistance from a statistician when they are planning a study. Also, design weaknesses could be reduced if researchers would seek evaluation of the logical consistency of hypotheses and data gathering instruments prior to the execution of the study, and among hypotheses, data, and conclusions on completion of an early draft of the report. If these two simple steps were followed, improvement in research quality would be highly probably.

Finally, sharper focus needs to be given to research in elementary science education. The period covered by this review yielded some evidence that groups of researchers are focusing on specific issues. Although still not a general trend, several studies were done on new instructional programs, on training teachers and the like. Moreover, some individuals, in conjunction with their respective doctoral students in elementary science, generated clusters of studies in reasonably limited investigative domains. Although definitive findings influencing practice have not been obtained as yet, the promise of systematic studies of carefully defined research areas, coupled with continued improvement in research methodology, is encouraging.
LITERATURE CITED

References including an ED number may be obtained from:
ERIC Document Reproduction Service, P.O. Drawer O, Bethesda,
Maryland, 20014.

1. Almy, Millie. Logical Thinking in Second Grade. New York:
   Teachers College Press, Teachers College, Columbia University,
   1970.

2. Ashley, James P. "A Study of the Impact of an Inservice Education
   Program on Teacher Behavior". Paper presented at the annual
   meeting of the National Association for Research in Science
   Teaching, Pasadena; February, 1969.

3. Ayers, Jerry B. "Evaluation of the Use of Science--\"A Process
   Approach with Pre-School Age Children\". Science Education 53

4. Bailey, Wilfrid C. "Development of a Sequential Curriculum in
   Anthropology for Grades 1-7". University of Georgia, Athens,
   Georgia.

5. Benson, Keith S. "A Comparison of Two Methods of Teaching Fifth
   Grade Science". Unpublished doctoral dissertation, Oklahoma
   State University, Norman, Oklahoma, 1968.

6. Berkeley, Keith D. "A Comparison of Student Teachers' Verbal
   Interactions in Elementary Classrooms". Unpublished doctoral
   dissertation, Wayne State University, Detroit, Michigan, 1968.

7. Berkheimer, Glenn D. "The Science Supervisor and Utilization of
   Curriculum Materials". Journal of Research in Science Teaching

8. Boykin, Wilfred E. "The Effect of Achievement, Feedback, and Re-
   view on Retention and Transfer of Fourth Grade Science Materials
   Presented by Computer Assisted Instruction". Unpublished doc-
   toral dissertation, University of Virginia, Charlottesville,
   Virginia, 1968.


55. MacDougall, Mary A. "Methods of Presenting Programmed Science Materials to Fourth Grade Pupils of Varying Ability and Achievement". USOE Contract, University of Virginia, Charlottesville, Virginia. ED 041 764 MF $0.65 HC $6.58 118 p.


   ED 035 564 MF $0.65 HC $6.58 137 p.

60. Miller, Donald C. "Knowledge of the Scientific Process Attained by Prospective Elementary School Teachers". Unpublished doctoral dissertation, The Ohio State University, Columbus, Ohio, 1969.
   ED 045 383 MF $0.65 HC $3.29 91 p.


60.


72. Pettersen, Mary Jean "The Effects of An In-Service Program in Science For Teachers of Grades 4-6 in Hammond, Indiana". Unpublished doctoral dissertation, University of Iowa, Iowa City, Iowa, 1968.


ED 026 253 MF $0.65 HC $6.58 135 p.


ADDITIONAL REFERENCES


ED 023 613 MF $0.65 HC $3.29 89 p.


ED 027 227 MF $0.65 HC $3.29 91 p.


