
This is one of several papers issued in cooperation with the National Association for Research in Science Teaching to analyze and synthesize research related to the teaching and learning of science completed over a two-year period. The majority of studies reviewed are unpublished doctoral dissertations. The research reviewed is grouped according to the standard categories used by the ERIC Science and Mathematics Education Information Analysis Center: Instructional Procedures, Teacher Education, Equipment and Materials, Curriculum, Achievement, Teacher Resource Materials, Science and Society, Evaluation and Educational Objectives, and Teacher and Student Characteristics. The research reviewed made it clear that few research workers were carrying on continuous, basic research programs.
RESEARCH REVIEW SERIES - SCIENCE
PAPER 5
A SUMMARY OF RESEARCH IN SCIENCE EDUCATION
FOR THE YEARS 1963-64, ELEMENTARY SCHOOL LEVEL

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.
A SUMMARY OF RESEARCH IN SCIENCE EDUCATION FOR THE YEARS 1963-64, ELEMENTARY SCHOOL LEVEL

by

John D. Cunningham, Keene State College, Keene, New Hampshire
David P. Butts, University of Texas, Austin, Texas
RESEARCH REVIEWS - SCIENCE

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.

Robert W. Howe
and
Stanley L. Helgeson
Editors
The Science Education Information Reports are being developed to disseminate information concerning documents analyzed at the ERIC Center for Science Education. The Reports include five types of publications. General Bibliographies are being issued to announce most documents processed by the Center for Science Education. These bibliographies are categorized by topics and indicate the availability of the document and the major ideas included in the document. Special Bibliographies are being 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 Teachers are bibliographies that identify references for the professional growth of teachers at all levels of science and mathematics teaching. This series will include six separate publications. Occasional Papers will be issued periodically to indicate implications of research for science and mathematics teaching. Research Reviews will be issued to analyze and synthesize research related to science and mathematics education over a period of several years.

The Science Education Information Reports will be announced in the ERIC Newsletter as they become available.
because of varying definitions of what constitutes research, it is difficult to compare the quantity of research conducted from period to period unless all such research is reviewed by a single team or a common definition is used. Such a comparison was not made and we cannot, therefore, compare the quantity of research conducted in elementary school science education during the 1963-1964 period with that conducted earlier.

Our overall impression, however, is of a slight improvement in the quality of research. Although further improvement is unquestionably needed, there was some trend toward greater quantification and more rigorous research designs. As in most past review periods, however, there are too many normative studies. In general, the studies reviewed had not as yet complied with the various proposals for the improvement of research in science education that appeared in Volume 1 of the Journal of Research in Science Teaching.

Several reviews of earlier research of various aspects of science education during the period (see especially Blackwood (4) and H. Smith (46) and most of Issue 3, Volume 2 of the Journal of Research in Science Teaching). For various reasons the period under review may be a turning point in the type of science education research being conducted. If anything, the reports do hint at the growing interest of psychologists (and other behavioral scientists) in science education and science educators in psychology.

The majority of studies herein reviewed were unpublished doctoral dissertations. No masters theses were directly reviewed. It is pertinent to point out that some of the dissertations have subsequently been published, sometimes with additional data and slightly changed conclusions. Similarly, some studies published during the period reviewed are not cited because of their previous review as unpublished doctoral dissertations.

The research reviewed is grouped roughly according to the standard categories used by the ERIC Information Analysis Center for Science Education.

INSTRUCTIONAL PROCEDURES

Variations in instructional procedures and their effect on student performance were studied by several individuals. In each of the studies reviewed the populations were fourth, fifth, and sixth grade students.
One strategy of instructional procedure examined was the usefulness of programmed instruction. Two studies involved constructing such materials and describing some specific outcomes in terms of the students' interest (40) or achievement (15). Comparison groups were not described in either of these studies. Two experimental studies involving programmed instruction were also reviewed, one using a post-test score only for analysis (12) and the second involving pre- and post-test change scores (5).

As a technique for individualizing instruction, Schiller (40) developed student activity booklets for sixth grade units which were used in addition to regular science. Evaluation was described in terms of how much students participated. The conclusion of the study was that students did participate.

Using audiotaped instruction concerned with the kinetic molecular theory of heat as a means by which the teacher variable could be reduced was the focus of Harris's study (15). He compared the pre- and post-test scores of the treatment group and concluded that children do learn through this approach. There was no contrast or control group described.

Dutton (12) analyzed programmed instruction in the fourth grade with the topics of sound, light, and heat. Two classes, experimental with treatment, and comparison without treatment, were compared on a post-test basis only. The treatment group had a high achievement score and students and teachers appeared interested.

When he used programmed instruction as a means to teach inquiry training, Blank (5) found that children asked more questions and participated in the class more. Their performance did not seem to be related to ability but was related to arithmetic and social problem scores. Blank used the change scores of the individual as the basis for his analysis.

Another variable in the instructional procedure examined was the explicit directing of instruction toward specified goals. Mason (26) designed science units which emphasized critical thinking. Using pre- and post-test change scores of fourth, fifth, and sixth grade students, he found all groups gained information and critical thinking ability regardless of the treatment.

The reading approaches used by intermediate grade teachers in their science instruction was investigated by Hells (30). He concluded, among other things, that "good" reading practices were more frequent at successively higher grade levels and that they were found more commonly in social studies than in science teaching.

In general, it seems relevant to note that to determine the adequacy of the effectiveness of an instructional procedure, the research design should include a treatment group and a comparison group with evidence of pre- and post-test gain. Of the studies reviewed,
only one (5) incorporated this design. In each of the other studies, further evidence is needed before it can be stated that instructional procedure A indeed offers more than does instructional procedure B.

TEACHER EDUCATION

A number of survey studies of specific science courses for teachers was reviewed [Bryant (8), Moser (33), and Service (43)]. Lerner (25) added a similar survey of the science methods course. Kisner (22) approached the concern of teacher education to find out what was included in the courses. Brandou (7), Hunt (18), and Miner (32) presented studies of the content of in-service programs and in-service teacher characteristics.

Service (43) analyzed certain elementary school science programs in terms of the academic science requirements of the college. He described several common elements in the elementary school science curriculum and variations in the college science requirements. How they fit or should fit together was not identified. Bryant (8) and Michals (31) also examined the college science requirements for elementary teachers. Both found variations in course requirements.

Kisner (22) studied the content of science and science methods courses in Oklahoma colleges. He constructed a check-list of what ought to be in the courses based upon the analysis of nine authorities. Through interviews with college instructors he concluded that in general the eight Oklahoma institutions taught much the same type of program but with individual emphases dependent upon the instructor.

Moser (33) surveyed the science preparation of elementary teachers in New York. While he found that they had, on the average, ten semester hours in science, he also found that 22 percent of the teachers had no college science background. Biology was the most frequently mentioned science course, being taken by 36 percent of his sample.

Lerner (25) surveyed the status of the elementary school methods course. In nearly 25 percent of her population, she found that the course was optional.

Brandou (7) explored the instructor variable of in-service programs using secondary school physical science teachers as instructors in conducting in-service sessions. Post-tests, biographical information, and field interviews suggested that the impact of the in-service program was related to the amount of teaching experience and the number of years the in-service instructor had in a particular school system. The more experienced classroom teachers also participated more frequently in in-service programs. The relevance of the program as conducted by high school physical science teachers was greater for grades 3, 4, 5, and 6 than for the primary grades. Participation in the development of classroom study guides
was found to be related to subsequent classroom use by the teacher.

Hunt (18) developed and implemented curriculum materials in space science which included a specific in-service program for the teacher. Using a pre- and post-test change score, he found that students showing the greatest gains were those in classrooms where the teacher had had the in-service program. Miner (32) also studied the in-service education of elementary school teachers.

In retrospect, it is difficult to make generalizations based on surveys of present status. Prior to the description of new programs for the preparation of teachers, it is certainly most useful to know what is the current practice. However, there is a definite limitation in defining current practice only in terms of the numbers of credit hours. For the survey to be useful, the content of these courses also needs to be specified. Knowledge of current practice, however, does not lead to change in and of itself. Design for new practices must be constructed. Evaluative studies then must be conducted to determine that specific sequences in the sciences and their teaching lead to improved student performance. Studies of this nature were not present in those reviewed during the 1963-64 period.

EQUIPMENT AND MATERIALS

Studies of the design and/or use of instructional equipment and materials in science instruction were scarce in the research studies reviewed. However, several studies did describe the relative shortage of materials.

Ricker (37) studied the criteria for selecting the use of science equipment in Maryland elementary schools. Analysis of a panel-validated questionnaire survey indicated that teacher preparation to use equipment varied as did most equipment available to that teacher.

Beitzel (3) constructed a chemistry unit for sixth grade, and as a result of its use with 35 Minnesota sixth grade classes, he found that most teachers did not have adequate equipment for teaching such a chemistry unit.

As far as printed educational materials are concerned, Eaton (13) and Williams (54) found that the same materials are not necessarily suitable for all students. They hypothesized that through manipulation of appropriate factors, it should be possible to produce materials which will result in maximum learning for all students.

In summary, it seems that empirical evidence supporting the use of non-use of science teaching equipment or materials is largely lacking.
CURRICULUM

The development and evaluation of science curricula for the elementary school represented a challenge that during the 1963-64 period seemed to elude the interest of most researchers in science education. Although few exploratory or experimental studies were reported during this time period, the importance of the challenge to develop empirically-based models for curriculum design cannot be ignored.

One pattern for curriculum development was described in two studies [Stapp (48) and McCombs (28)]. In both studies key ideas were identified and then appropriate instructional materials were prepared. Although descriptions of the materials were given, no evidence of empirical analysis, such as student data, was given to measure the adequacy of the materials. For example, Stapp (48) described the development and implementation of a conservation education program. He described the procedures involved in designing the program, in preparing the materials, and in preparing teachers in using the materials. Further study is needed as to the effectiveness of each of these points.

Two other studies developed but did not empirically test curricular materials; McCombs (28) a guidebook of experiences for grades K through 3 and King (21) a sourcebook of earth science activities.

Smith (45) on the other hand, did conclude that a topic such as "Light" could be subdivided into concepts and the concepts ranked according to difficulty based on the pre- and post-test performance of students. This possibility is the beginning of a validation of a hierarchy for an individual topic. Relating this to a curriculum sequence remains yet to be done.

A further step was taken by Ashbaugh (1). He selected a series of geological concepts for intermediate grade children; constructed an instrument to measure understanding of each concept, instructed (through demonstrations and experiments) a group of children in the concepts, measured the effects of instruction, and arranged the concepts in order of difficulty. Concepts were considered appropriate for the grade level in which over 51 percent of the children correctly answered the test questions based upon the concepts.

Harris (15) also conducted a study on grade placement by studying the differential learning of selected fourth, fifth, and sixth graders. Uhlhorn (52) attempted to measure the appropriateness of the grade level placement of science concepts by means of a science experience inventory.

However, Shoresman (44) has pointed out certain methodological problems in this type of research, namely that no concept is ever really exhausted or completely learned and that the teacher variable is not really eliminated. Further he questions the long range value
of grade placement research, pointing out that we need to move from attempting to place concepts at specific conventional grade levels to experimentally deriving content gradients which are capable of serving as guidelines for the development of sequences of learning experiences.

Studies such as the above on the grade placement of science concepts have been conducted for some time. As in the earlier studies there was no attempt to relate the findings to the growing body of psychological knowledge concerning the child's cognitive skills.

One comparative education study was completed during the period reviewed; Tanveer (50) evaluated the science curriculum for primary schools in Pakistan.

Finally, as a principle to be considered in curriculum development, McGowan (29) repeats a suggestion that is not infrequently made, namely that since children's development of concepts parallels somewhat that of the historical development of the concept, increased attention should be given to this historical dimension in curricular innovations.

ACHIEVEMENT

Numerous studies during the period were on student achievement. This research followed a number of diverse patterns.

Historically little research in science education has been directed toward the gifted child. Becker (2) investigated various administrative groupings (segregated, partially segregated, and non-segregated classes) upon the academic achievement of gifted sixth grade children. He concluded that when language mental ability was held constant, gifted children placed in special groups did not achieve significantly more in science and mathematics.

Probably because of the large number of uncontrolled (uncontrollable?) variables in educational research, numerous comparisons in correlational studies were not statistically significant. For example, Downing (11) found no relationship existed among science interest level change, science achievement gain, and daily time allotment for science instruction with the sample of sixth grade students studied.

However, C. Scott (41) did find a positive correlation not attributable to intelligence between reading gain and gain in science. She concluded that the degree of concomitance between improved reading ability and improvement in science is less than many educators realize.
Elementary school teachers cannot be equally prepared in all curriculum areas (or even in all areas of science) and therefore must rely upon printed educational materials. Two dissertations from the University of Illinois investigated the effect upon achievement of rewritten materials [Eaton (13) and Williams (54)].

Eaton (13) investigated the effect upon the level of pupil achievement of three factors in printed instructional materials: number of questions, activities, and incongruities. Only one factor, activity, produced significant differences in post-test scores; each group achieved highest scores when studying from materials containing the greatest number of activities.

Williams (54) rewrote sixth-grade textbook material to the third-grade level. Such rewriting led to improvement in science comprehension as well as reading abilities.

The influence of the teacher variable has confused research related to achievement in the past. Harris (15) through audiotaped material, and Dutton (12), Blank (5), and Schiller (40), through programmed instruction, offer ways to reduce this influence. Shoresman (44) believes, however, that those practices do not eliminate the teacher variable, they only control it somewhat more than usual.

The purpose of Rowe's study (38) was to investigate context-learning with physical science material used with first-grade boys. Context-learning was defined as the process of forming sets of entities in such a way that the meaning of any one of the entities is partly or wholly determined by defining the meaning of the set containing the entity. She found that the accumulation of experiences in a variety of related contexts increased the probability that a particular concept will be included in a context, decreased the redundancy, and increased the probability that a given event is perceived as a possible successful solution to a problem. Rowe suggests that cybernetics could provide powerful descriptive techniques suitable for the study of context-learning as a process.

N. Scott (42) studied the relationship of science concept achievement to inductive reasoning and cognitive styles in categorization behavior in 10- and 11-year old children. The results indicated that each age level had a unique dependency upon a particular method of assessing cognitive functions.

A number of investigators developed their own achievement tests in addition to using such standard instruments as the Sequential Test of Educational Progress and The Metropolitan Achievement Test. The necessity for this practice is probably quite real but because few studies are ever replicated and the instruments are usually not made available widely, the real significance of it is in doubt.
TEACHER RESOURCE MATERIALS

The development of several types of teacher resource materials was reported. Beitzel (3) prepared a technically-oriented, experiment-centered sixth grade chemistry unit, including a chemistry textbook and a teacher's manual. The appropriateness of the materials was determined by a test and by questionnaires.

King (21) developed a sourcebook in earth science while McCombs (28) developed a guide for science experiences for grades K through 3. In neither case was any evidence presented as to the adequacy of the instructional materials to accomplish the task for which they were designed.

SCIENCE AND SOCIETY

Two extensive histories of elementary school science education appeared in 1963 [McCloskey (27) and Smith (46)].

LEGISLATIVE ACTS AND REPORTS

The only report that falls in this category is the study by Blackwood (4).

EVALUATION AND EDUCATIONAL OBJECTIVES

As noted in previous sections of this review, many attempts at curriculum development stopped prior to the stage of evaluation. Or, alternatively, the curricular innovation was evaluated only in terms of the "attitudes" of students, teachers, and administrators toward it. However, two studies did report on the development of new tests that might be useful for curricular evaluation [Butts (10), Nelson (34), and Mason (26)].

Nelson (34) and Mason (26) developed a test of critical thinking for grades 4-6 entitled A Test of Science Comprehension. Butts (10) described a very useful instrument that might be used to evaluate problem solving abilities. Known as the X-35 Test of Problem Solving (or "tab test"), the instrument allows the student to select the kinds and amounts of information he believes would best enable him to solve specific problems. The student's chain of reasoning can thus be identified.

Another interesting study by Butts (10) casts some doubt on the widely held tenet that the provision of direct experience with the raw phenomena of science is a necessary and sufficient condition for the development of concepts.

Pollach (36) examined the relationships between tests of intelligence and reading and levels of concept development in terms of a
hierarchy of understanding defined as knowledge, comprehension, and application. The study illustrated the difficulty of interpreting the definitions of Bloom's Taxonomy of Educational Objectives and pointed to the need for clear and discrete behavioral definitions.

It is not possible to say that major breakthroughs in evaluation of learning processes, or of research methodology occurred during the period reviewed.

TEACHER AND STUDENT CHARACTERISTICS

There was clearly more research done on teacher education than there was on teacher characteristics. There was also considerably more research conducted on student than on teacher characteristics.

Because of the growing trend toward K-12 science programs, Stone (49) explored the patterns of criteria used by elementary and secondary teachers in judging the relative effectiveness of elementary science experiences in order to determine whether secondary and elementary teachers shared common goals of education and common views on the proper methodology to achieve such goals. She concluded that there were no major differences between the two groups of teachers.

Quite a number of studies dealt with student characteristics. Until very recently the development of science curricula was based to a large degree upon the logical nature of the science concepts. Quite frequently little regard was given to the psychological nature of children. One of the major influences of Piagetian research has been to force researchers to pay more attention to such characteristics of children. It is noteworthy that most new large-scale curriculum projects are now staffed by science teachers, scientists, and psychologists.

Attention to student characteristics has taken a number of forms. Two studies [Bowen (6) and Hardin (14)] studied aspects of interest in science. For example, Bowen (6) compared the topic preferences of boys and girls in school science exhibitions. As many science educators might have predicted, boys overwhelmingly chose topics in the physical sciences. In most comparisons, more girls than boys chose biological topics. The pattern was fairly uniform at the upper elementary, junior high, and high school levels. Bowen conjectured that causes were multiple, complex, and probably culturally induced. He further speculated that the sex-topic difference is one price we pay for emphasizing familiar facets of applied science as a motivational device.

Hardin (14) measuring science interest by the frequency of pupils' voluntary participation in science activities, concluded that the science interest of intermediate-grade pupils is multidimensional.
Ashbaugh (1) selected a number of concepts (geological) for fourth, fifth, and sixth grade students. He taught the concepts and gave a post-test. Each test question answered correctly by more than 51 percent of the pupils was considered suitable for that grade level.

Through use of a science experience inventory with fourth, fifth, and sixth graders, Uhlhorn (52) tested the grade placement of various science concepts. He found that for one large midwestern city the science experience background of such children was not related to the placement of science concepts in the curriculum guide.

Other comments about similar grade-placement research were made in the section on "Curriculum."

Eaton's results (13) showed that the same printed educational materials may not be appropriate for all students. He rewrote astronomical material, either increasing or decreasing the number of questions asked, the incongruities present, or the activities. Students with mental ages in the upper quartile responded best to questions; the fewer the number of questions, the higher was the post-test score. Students with mental ages in the lower quartile responded only to activity; the greater the number of activities, the higher was the post-test score.

Similarly, programmed materials are apparently not for everyone. Taylor (51) reports a method by which success with programmed science materials can be predicted from achievement and intelligence measures.

A number of investigators have suspected, although not all have verified the belief, that Piagetian interview techniques and categorization systems are tied quite closely to Piaget (and his associates) personally and, therefore, might not have wide utility. This possibility was apparently upheld by R. Smith (47). He concluded that answers given by first, second, fourth, and sixth grade children to questions about natural phenomena were influenced by the child's vocabulary, the nature and wording of the questions, and the child's experiential background rather than age-related stages.

The development of concepts of living and non-living among five preschool children was the topic of Rush's study (39). He concluded that the five preschool children studied were intellectually different and that they formed concepts either with or without instruction.

Several investigations were concerned with the science knowledge and science vocabulary of kindergarten children (Helfrich (16), Inbody (19) and Kolson (23) et al.). In general, these studies concluded that kindergarten children "are quite knowledgeable in the area of science."
Somewhat similar to the above studies but involving older children were investigations by Olmsted (35), Kuse (24), and Weaver and Coleman (53). Kuse (24) concluded that first, second, and third graders knew many, but not all, concepts of astronomy, with third graders more knowledgeable in this area than were first graders. Because first graders also had a considerable science background, especially those who had had kindergarten experience, Olmsted (35) concluded that more difficult science concepts should be included in the first grade curriculum than is presently the case. Similarly, Weaver and Coleman (53) found that first-grade children of average or below normal mental ability can be taught to develop the major concepts of science.

Despite the increasingly concrete nature of elementary science education, much learning is still verbal. Therefore, descriptive studies such as those of Howards (17) and Kerns (20) are quite important. The basic question investigated by Howards (17) was, "How easy are the 'easy' words?" He measured children's understanding, at grades four, five, and six, of various meanings of selected high-frequency, mono-syllabic, multiple-meaning words which appear in scientific word lists. He concluded that older children know more meanings for multiple-meaning words than did younger children and that so-called "easy" words were not always so easy. Kerns (20) also found variation in the meaning and significance that elementary school children attribute to selected science words.

An abundance of other information about pupil characteristics is, of course, embedded in many of the studies reviewed elsewhere. Many of the comparisons made in those studies involved sex, age, intelligence, and socioeconomic differences.

Because of the diverse nature of the above research, it is again difficult to summarize it. It was clear that during the period reviewed, there were few if any research workers carrying on continuous, basic research programs.
BIBLIOGRAPHY

Documents with ED information are available from EDRS, National Cash Register Company, Bethesda, Maryland 20014.


