This review for 1976 has been issued to analyze and synthesize research related to the teaching and learning of science completed during the year. The review is intended to provide research information for development personnel, ideas for future research, and an indication of trends in research in science education. Research has been listed in general categories of: (1) Learning and Development; (2) Teaching Methods and Procedures; (3) The Education of Teachers; (4) Evaluation in Science Education; (5) The Use of Media in Science Education; (6) The Concepts, Processes and Content of Science; and (7) College Level Research. In all, 327 separate studies are cited in the bibliography and most are mentioned in the text. Many intermediate summations and generalizations are included at the end of sections and subsections. (HM)
A SUMMARY OF RESEARCH IN SCIENCE EDUCATION—1976

John W. Renner and Michael R. Abraham
University of Oklahoma

Don G. Stafford
East Central Oklahoma State University
and the
University of Oklahoma

an Interscience® Publication
published by JOHN WILEY & SONS
New York  Chichester  Brisbane  Toronto
This publication was prepared pursuant to a contract with the National Institute of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent National Institute of Education position or policy.

Volumes prior to 1973 of the SCIENCE RESEARCH REVIEW SERIES are available from the SMEAC Information Reference Center, The Ohio State University, 1200 Chambers Road, Room 310, Columbus, Ohio 43212. Volumes from 1973 on are available from John Wiley & Sons, Inc., 605 Third Avenue, New York, N.Y. 10016.

Single copy price for this volume: $8.95
Preface

Research Reviews are being issued to analyze and synthesize research related to the teaching and learning of science completed during a one-year period of time. These reviews are developed in cooperation with the National Association for Research in Science Teaching. Appointed NARST committees work with staff of the ERIC Science, Mathematics, and Environmental Education Information Analysis Center 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
Patricia E. Blosser
ERIC/SMEAC
Foreword

When viewing the research done in science education in a year, one is first nearly overwhelmed with its diversity. If one has agreed to prepare a cogent review of that research, its diversity immediately becomes the first problem to be solved. Briefly stated, that problem is: How can the important work being reviewed be organized into a useful, coherent, meaningful pattern?

To answer the foregoing question we employed a technique not unlike that used by Mendelyeev when he developed the first periodic table of the elements. In other words we read abstracts and sorted them into groups that had common subject matter. When the abstract left any doubt regarding the subject matter of an item, the original source was consulted before a decision was made. The organization of this report, therefore, is purely empirical.

Those persons who did the research fashioned this report's organization into seven categories:

1. Learning and Development
2. Teaching Methods and Procedures
3. The Education of Teachers
4. Evaluation in Science Education
5. The Use of Media in Science Education
6. The Concepts, Processes and Content of Science
7. College Level Research

While each of the foregoing categories constitutes a chapter in this volume, each chapter has an internal organization based upon that chapter's content.

Our intent in preparing this report was to provide a general overview of the research in science education done in 1976 and to state conclusions and questions arising from the research which might lead other researchers to continue and expand what has been done. A second intention of this report's authors was to include a review of those studies that could genuinely be called research. There are, however, other types of activity in science education which will be of interest.
to science education researchers and which have been published in some form. A rather complete bibliography of the activities taking place in science education in 1976 follows the narrative of this report. That bibliography contains many entries other than the research reviewed in the narrative of the report.

The authors express their thanks to Dr. Stanley Helgeson and his staff of Ohio State University for their assistance and patience in searching for and locating the data upon which this report is based. Our thanks are also extended to Donna Abraham who typed the report's manuscript.

Norman, Oklahoma
August, 1977

John W. Renner
Michael R. Abraham
Don C. Stafford
# Contents

## Foreword

Learning and Development

- Research With the Ausubelian Theory .................................................. 1
- Research With the Piagetian Paradigm—
  - Phase I: Measurement ........................................................................... 4
- Research With the Piagetian Paradigm—
  - Phase II: Curriculum ............................................................................ 14

Teaching Methods and Procedures ........................................................ 18

- Research in the Methods and Procedures of
  - Teaching Elementary School Science ................................................. 18
- Research in the Methods and Procedures of
  - Teaching Secondary School Science .................................................. 25

The Education of Teachers .................................................................... 34

- Research in Elementary School Science Teacher Education .............. 34
- Research in Secondary School Science Teacher Education ............... 40
- Research in Science Teacher Education in K-12 ................................. 47
- Research in Inservice Teacher Education ............................................ 49

Evaluation in Science Education ......................................................... 51

- Evaluation Surveys ................................................................................ 51
- Curriculum Evaluation .......................................................................... 54
- Affective Studies ................................................................................... 60
- Instrument Development ....................................................................... 69
- Evaluation Methodology ....................................................................... 69

The Use of Media in Science Education ................................................. 74

The Concepts, Processes, and Content of Science .................................. 81

- Syllabi .................................................................................................... 81
- Specific Content, Concepts, and Processes .......................................... 83

College Level Research .......................................................................... 88

- Component Comparisons ..................................................................... 89
- Instructors and Students ...................................................................... 108

References .............................................................................................. 113
A Summary of Research in Science Education — 1976

JOHN W. RENNER and MICHAEL R. ABRAHAM
University of Oklahoma
Norman, Oklahoma 73019
and
DON G. STAFFORD
East Central Oklahoma State University
Ada, Oklahoma 74820
and
University of Oklahoma
Norman, Oklahoma 73019

Learning and Development

The research done in the category of "Learning and Development" during 1976 focused upon the theories of David Ausubel and Jean Piaget. One study was designed to compare the theories of Piaget, Vygotsky, and Ausubel, four studies considered the work of Ausubel and a total of 21 studies examined the paradigm proposed by Piaget. The four studies considering the Ausubelian theory tested the efficacy of the advance organizer concept. The work focusing on the Piagetian theory considered the measurement of levels of intellectual development or the influence of the theory upon curriculum.

Research With the Ausubelian Theory

The four studies reported which considered the advance organizer technique were made at three different educational levels. These levels are the ninth and twelfth grades, college level physical science for non-science majors, and college juniors in a teacher education program.

Murchison (203) explored the usefulness of the advance organizer concept in teaching science to ninth graders. Four groups of students were formed and each exposed to a different organizer: a pretest, a concrete organizer, an abstract organizer, and a control. The groups were brought together immediately after experiencing the organizer, given a lecture on sound waves and then administered a two-part test. The student groups were matched on IQ, sex, and teacher-determined motivation ranking.
The pretest was the same as the test on sound waves given after the lecture, and the concrete and abstract organizers involved generalized wave theory. One part of the test involved rote material and all questions which had been answered in the lecture. The second part of the test required abstract and/or problem solving ability. Murchison found that, when abstract material is to be learned, a pretest is the best organizer for high IQ boys and for high motivation students. An abstract organizer is best for high motivation boys and a concrete organizer is preferable for high IQ girls and for low motivation students. When rote material is to be learned, a pretest is the best for low IQ girls. Murchison concluded that multiple advance organizers are important aids to learning.

West and Fensham (307) tested some predictions arising directly from Ausubel's theory concerning the subsuming role of advance organizers. Specifically, the basic question studied was regarding whether the learner's existing knowledge plays a subsuming or organizing role in subsequent learning. The selected topic of "Principles of Equilibrium" was taught in twelfth grade chemistry; a period of one week (160) minutes was used to teach the topic. A total of 374 subjects was involved. The researchers stated explicitly that they were not interested in whether or not an external organizer led to more effective learning. Rather, they were interested in whether the role played by an external organizer was equivalent to the role played by the learner's prior knowledge. The results of the study did not show that the role played by the learner's prior knowledge was an organizing or subsuming role. In other words, the results did not show that the role played by the learner's prior knowledge and the role played by the advance organizers were the same. According to these researchers, this research did show that, at least for the topic taught, "there is an apparent equivalence in terms of learning outcome between whatever role is played by the learner's relevant prior knowledge and the advance organizers." The conclusion was drawn that the research supported Ausubel's proposition that prior knowledge structure plays a subsuming or organizing role in new learning.
Baylis (21) utilized two classes of advance cognitive organizers (expository and comparative) to study their effects on learner ability to conceptualize and modify cognitive structure as a readiness process prior to interacting with science content. The data were collected through a pretest-posttest multiple treatment design and the use of the "Word Association" instrument and a "Curriculum Embedded Questionnaire." The former measured the conceptualization of cognitive structure (defined as "meaningfulness of individual concepts and relatedness among concepts") and the latter measured the perception of the process (defined as conceptualization and valuing of the instructional process). This researcher concluded that the two subsets of advance cognitive organizers can provide a theoretical base for designing readiness instruction. The processes used were shown to facilitate both conceptualization and modification of a cognitive structure relevant to a specific science content area. Finally, the investigator believes the data showed that the learners both conceptualized and valued the process; in other words, the process itself was learned.

A laboratory inquiry-oriented course in college physical science for non-science majors was selected by Tavares (291) to use in evaluating the effects using advance organizers. Tavares' sample was divided into eight groups: high and low academic ability, high and low degree of previous knowledge, existence and non-existence of advance organizers before studying the content. The null hypothesis tested was: The treatment does not introduce any statistically significant difference for any of the groups. The quantitative data necessary to test that hypothesis were obtained from the (1) Test on Understanding Science, (2) Wisconsin Inventory of Science Processes, (3) an investigator-constructed content achievement test, and (4) an investigator-constructed student attitude inventory. The null hypothesis was accepted for any of the variables with all of the groups.

The research reported here neither confirms nor denies the value and/or efficacy of Ausubel's advance organizer concept. There appears to be a need to approach the problem in a more systematic manner than was evident in the research reported here. Those persons conducting research in this area should, we believe, agree on a set of priorities and procedures
Research With the Piagetian Paradigm — Phase I: Measurement

Most assuredly the intellectual development model of Jean Piaget has been recognized as researchable by science education. The research reported during 1976 centered around (1) the measurement of intellectual ability as described by Piaget and (2) the value of the Piagetian model to the general area of science curriculum.

In establishing his now famous four stages of cognitive development, Piaget utilized various tasks to determine a person's thinking patterns. Each of these tasks is individually administered and a substantial amount of time is required to administer tasks to determine a learner's reasoning patterns. For several years researchers have been interested in developing procedures and materials which will allow the reasoning levels of an entire class or group to be determined at one time. Such an undertaking suggests the use of some type of paper-and-pencil evaluation tool as a possible procedure. There remains, however, much work still being done with the clinical-interview technique designed by Piaget; that technique preserves the one-to-one task administration.

Hooper, Brainerd, and Sipple (128) conducted a four-year longitudinal analysis of concept development and, as part of that study, designed a series of logical concept tasks based upon Piagetian theory. These tasks were individually administered to students five years of age and older. The tasks were focused upon studying groupings associated with concrete operational reasoning. A sample of 180 children in kindergarten and grades three and six was used; equal numbers of males and females were included. These researchers concluded that their "concept task series" provided a generally reliable assessment of logical reasoning for this age group.
Using six Piagetian-styled tasks that correspond to the six major logical groupings of concrete operational thought, Camp (43) asked whether or not the scores on those six tasks formed a unidimensional scale. He used 102 children equally divided among grades one, two, and three, and studied differences in performances for the three grade levels, relationships between performance on each of the six tasks and IQ, and differences in performances of males and females. Camp found: (1) the tasks did form a unidimensional scale, (2) there was a significant increase in difficulty between grade levels on four of the six tasks, (3) a significant degree of association existed between IQ and only one of the six tasks, and (4) there was no significant difference between performances of males and females on the six tasks. Based upon the data, Camp made several recommendations, among which was the recommendation that early curricula should focus upon experiences with objects rather than relying on verbal transmissions about objects.

Working in Thailand, Pungah (237) studied the developmental sequence of the conservations of number, mass, weight, and volume with 80 boys and 80 girls ranging in age from 4 to 11.3 years. The results of this study support Piaget's findings that conservation reasoning with these four tasks occurs in the sequence: number, mass, weight, and volume. Thai children conserve number between 6 and 7 years, mass between 7 and 8 years, and volume after 11 years. The conclusion was also drawn that sex and socioeconomic level were "unlikely to be contributing variables" in the development of conservation reasoning by Thai children.

Michael Shayer, working at Chelsea, England, (265) studied development in thinking of middle school and early secondary school students. He used children between the ages of 9 to 14 years. Each participating child (the exact number Shayer used was not available, he did say that there was a need "to test about 2,000 children in each year of age" for his purposes) completed a spatial concepts task, a volume and heaviness task, and the pendulum task. He concluded that the major development in early adolescence is concrete operational reasoning, and believes this is the type of reasoning that science programs for this age level "should both
build on and promote." Shayer also found differences between the general population of this age group and the population found in the "above-average comprehensive schools."

A total of 96 seventh grade science students ranging in age from 11.7 to 14.4 years (a mean age of 12.6 years) was administered eight conservation tasks to investigate if successful completion of those tasks was dependent only upon one type of reasoning or if there were three components (early concrete, middle concrete, and early formal thought) required to successfully complete all the tasks. When Lawson and Nordland (162) subjected their data to a principal components analysis, they found one factor which could be identified as early concrete-operational reasoning, another that identified "something intermediate between early concrete thought and early formal thought," and a third factor that seemed to indicate early formal thought. These researchers state that their results "support Piaget's distinction between concrete and formal thinking and the tasks' ability to measure these thinking abilities. They further state that the results do not necessarily support the Piagetian theory in general nor the notion that formal reasoning constitutes a "unified stage of development."

Benefield and Capie (23) used ten of the 16 binary operations which Piaget describes as being acquired during the formal thought period to construct a test using four content types: true content, false content, neither false nor true content, and randomly related content. (It is the 16 binary operations, where acquired, that represent the whole system of integrated operations at the formal level.) Six operations were eliminated because of logical equivalence to other operations. Each of the ten operations used was defined by a set of four truth conditions. Using the 16 binary operations, each of the set of truth conditions could be judged true or false when dealing with two variables. Each test item contained a substitution instance of propositional operation corresponding to a Piagetian binary operation. The subjects were asked to check "Yes, this could be true" or "No, the statement could not be true." Four 10-item tests were constructed; the items were assigned randomly to complete a 40-item test. One hundred fifty-five students in grades 4 through 12
constituted the population for the study and completed the test. The range of scores was 0-34, with a mean of 14.9. Significant differences were found among operations involving each of the four content types as well as the composite.

Janovský (134) investigated the development of the quantification of speed in children and adolescents. The population used consisted of 69 seventh graders, 66 ninth graders, and 47 eleventh graders; all participants were currently enrolled in mathematics. Four levels for qualifying speed were defined and two sublevels were established for each level. Six test items were prepared for each sublevel; the measurement instrument, therefore, consisted of 48 items. Forty-nine subjects were also interviewed to check the accuracy of the results obtained from the 48-item test. The questions asked the students to compare speeds and predict distances. A test on using the operations involved in fractions was also administered. The investigator found that the four levels and the sublevels formed scales that could be used to assess the development of quantification of speed in children and adolescents. Several correlations with reading and mathematics and the developed instrument were also included. The investigator concluded that the results demonstrate the need for additional instruction aimed at developing an increased understanding of the equivalence of ratios within the context of speed.

Data from other research support this conclusion but at a more general level; those data—some developed by Piaget—demonstrate that students do not develop understanding of the ratio concept until early in the formal reasoning period.

A group of 105 third grade children was administered a screening test to determine if they exhibited concrete-operational reasoning with respect to the ability to make speed-comparisons and perform simple division. From that group, Boulanger (34) found that 74 had met the criteria; those were included in his study to determine the effects of training in the proportional reasoning associated with the concept of speed. From the 74 children, 51 were randomly selected and assigned to three equally sized, treatment groups. The mean age of the 51 children was 8 years, 10 months. One treatment (T & C) consisted of training the children in the proportional reasoning associated with converting distance and elapsed time into a
The children were next exposed to speed comparison problems. The second treatment consisted of using comparisons of distances traveled and elapsed time only (CO); no training was given on proportional reasoning. A control group was also used. The T & C treatment was provided during two 25-minute and one 10-minute session on three successive days. The CO treatment consisted of one 30-minute and one 15-minute session on two consecutive days. Those receiving the T & C treatment scored significantly higher on measures of immediate retention than did the control group but that significance disappeared when long-term retention was checked. Both groups performed significantly better than the control group on problems different in context but similar in format. The T & C group could perform this transfer immediately while the CO group achieved significant success on transfer problems only after a delay of three weeks. In total transfer there were no significant differences between the groups. The investigator says: "It is tempting to conclude that the conflict experienced by the latter (CO) group in not being given a method of solution had a stronger residual effect than the directly trained group."

Lawson and Blake (160) addressed themselves to the criticism that some Piaget-designed tasks "such as the pendulum task and the balance-beam task...appear to require knowledge of specific physics content and therefore do not measure the underlying cognitive operation which they purport to assess." These authors used three Piagetian tasks (the pendulum, bending rods, and the balance beam) designed to measure formal thought, a 13-item pencil-and-paper biology examination designed by the investigators, and a non-science content examination designed by Longeot. The student sample consisted of 32 males and 36 females from the same high school who were enrolled in a second semester high school biology course. The age range of the students was 14 years, 7 months to 17 years, 10 months with a mean age of 15 years, 5 months. Lawson and Blake concluded that since the students "did not perform more formally on the non-science content examination and on the biology content examination than on the Piagetian tasks, the claim that students perform poorly on the Piagetian tasks because they require knowledge of specific physics content has not been supported." Lawson and Blake also concluded that "The Piagetian tasks are relatively content free and can serve as realistic indicators of concrete and formal thinking abilities."
In 1970 Robert Karplus, University of California, Berkeley, published an article which included a puzzle which he called "The Islands' Puzzle." He suggested that the puzzle might perhaps be useful in assessing the presence of formal reasoning. Blake, Lawson, and Nordland (29) administered "The Islands' Puzzle" (as designed by Karplus) to 126 secondary school students and then interviewed those students with three Piagetian tasks. Fifty-eight pharmacy majors at Purdue took the Karplus-designed version of "The Islands' Puzzle" (Form A). A rewritten form, Form B, was administered to 52 pharmacy majors and 50 completed a second rewritten form, Form C. A random sample of 24 of the 58 students who completed Form A were interviewed with the same three Piagetian tasks. The investigators concluded that "The Islands' Puzzle" may measure deductive logic but is not measuring the same abilities as the three Piagetian tasks used. These researchers do not recommend using "The Islands' Puzzle" to assess operational levels as defined by Piaget or to characterize a person's cognitive level.

Between 1970 and the present, the group working at the Lawrence Hall of Science, University of California, Berkeley, has published five studies which reveal their findings of "Intellectual Development Beyond Elementary School." The sixth of these studies is authored by Wollman (317) and is concerned with the ability of persons to control variables. Pilot studies had confirmed that, starting as early as the third grade, children are aware of the role of variables in certain contexts. A total of 1555 students in grades 4, 6-10, and 12 participated in the research. The students responded to written questions following an oral introduction during which an example was modeled. Most students finished the question in less than 10 minutes. Two scorers read and evaluated each paper and an 87 per cent agreement was achieved. The data revealed that all responses could be placed in four separate response categories. Wollman concluded that the concept of the controlled experiment probably begins near the beginning of the concrete operational stage. The root of the mature concept is probably in the child's concepts of evenness and fairness. Absent at the concrete level, however, is a clear idea of what to do in general when asked to judge a multiple-caused event. Even when variables are specifically isolated for the learners at the concrete stage they do not systematically determine each variable's role by varying it and holding all others the
same. Thinking that there is a single formal concept of controlling variables and that it belongs only to the formal stage rather than developing throughout the concrete and formal stages is probably not useful.

Karplus et al. (144) administered a proportional reasoning task and a control-of-variables task to 3600 students between 13 and 15 years of age in Denmark, Sweden, Italy, United States, Austria, Germany and Great Britain. They collected the data in order to make a determination of the distribution of concrete and formal thought at these age levels. The results of the research show that 25 per cent of the sample were rated at the formal level, 32 per cent in a transition stage, 15 per cent of the students attempted to solve the problems by addition (Karplus calls this the additive level) and 28 per cent of the sample operated at an intuitive level.

Two paper-and-pencil tasks were added to the placement tests administered to 885 students enrolled in general chemistry at the University of Nebraska, Lincoln, in the fall, 1974 (5). The purpose of the research was to improve the placement advice given by the admission officers and to test the idea that scores on written Piagetian tasks would be particularly effective as grade predictors. The results show that the Piagetian-tasks score accounted for very little variance in course performance. The specifically designed Piagetian tasks did not account for additional variance over and above that accounted for other measurement devices already used in the placement procedure.

Not infrequently, scores on Piagetian tasks are found not to correlate highly with grades received in college science courses. We hypothesize that the teaching techniques (lectures, demonstrations, show-and-tell) used in many college courses do not demand the use of formal thought. Furthermore, students are tested upon what they can memorize. Many concrete operational students can memorize enough information to receive a respectable grade and, consequently, the correlation between course performance and operational level as determined by tasks is reduced.
Dunlop and Fazio (82) used 466 students enrolled in general college chemistry or college physical science to test the assumption that the level of reasoning used by students when solving problems is substantially below their capacity. Specifically, these researchers investigated abstract preferences of science students in 18 problem-solving tasks and the relationship between those preferences and level of cognitive development. When studying the abstract ability (level of cognitive development) and abstract preference scores for science students in five different grade levels (eighth, ninth, twelfth, college freshmen, and college seniors) the finding was that the abstract preference scores did not significantly differ from grade to grade even though abstract ability did increase significantly as grade level increases. Furthermore, there was no significant difference between college freshmen science and non-science majors but a significant difference in favor of the science group did exist between the abstract preference scores for the science and non-science groups.

Lawson and Wollman (161) designed and carried out a study to find answers to four questions:

1. Can instructional procedures be designed and employed to successfully affect transition from concrete to formal cognitive functioning in fifth and seventh grade students with regard to one aspect of formal thought, i.e., the ability to isolate and control variables?

2. If training can enable concrete students to perform at a formal level on tasks which were used in the training, will that training transfer to tasks also involving the control of variables but using novel materials?

3. If training can enable concrete students to perform at a formal level on tasks requiring the control of variables, will that training transfer to tasks involving different concepts but ones which also involve formal thought?

4. What is the relationship between a student's level of intellectual development and his/her ability to profit from the training?

To secure data to answer those questions, Lawson and Wollman worked with 32 fifth grade students and 32 seventh grade students. The fifth grade sample was composed of 14 males and 18 females, had an age range of from 9.5 years to 12.1 years, and a mean age of 10.5 years. No IQ data were
available. The seventh grade sample was divided equally between males and females, had an age range from 11.9 years to 13.6 years, and a mean age of 12.6 years. The IQ range was 100-115 with a mean of 109.

The experimental design used was pretest/posttest control group model. The students in the experimental groups received four sessions of individual training on cause-and-effect relationships. The control groups attended their regular classes. Posttesting followed the training sessions and consisted of individual interviews conducted by two examiners who did not know to which group the students belonged. A female examiner conducted the testing with the females and a male examiner worked with the males. The Piagetian tasks used in the pretesting enabled the students to demonstrate early concepts (IIA), late concrete (IIB), post concrete, and early formal (IIIA) levels of thought. The tasks used in the posttesting allowed late formal (IIIB) thought to be demonstrated. In the posttesting, the bending rods task measured whether or not the training was effective in facilitating the ability to control variables with materials identical to those used in the training. The pendulum task was used to determine whether or not the training was generalizable to a problem also involving controlling variables but using novel materials. The balance beam task was also used to measure the extent to which the training led to formal thinking. Paper-and-pencil tests and oral responses were also used. The seventh grade sample also responded to a shortened version Longeot examination.

These two researchers answered the first two questions stated earlier affirmatively. Students can be trained to isolate and control variables. These authors answer question three like this: "Although the training was effective in promoting formal thought with regard to one aspect of formal reasoning, it was limited in extent." In other works, these researchers' data do not support answering question three affirmatively; training for general formal operational ability is not productive. The data collected generally show—in answer to question four—that the more formal students were more receptive to training than were the more concrete students.
The review of the foregoing research was more extensive than the others reviewed here because the Lawson-Wollman research article received the NARST award as being the most outstanding research publication in 1976.

As the studies involving the measurement within the Piagetian paradigm were reviewed, one constant factor seemed to consistently trouble us. That factor is time—duration of the experiment. If one subscribes to the Piagetian paradigm he/she understands that the endpoint is the construction and/or reconstruction of mental schema and/or structure. That process does not, we believe, happen rapidly; time is required. In addition, in all probability except for the experiment's activities, the remainder of the learner's educational experience is at the show-and-tell level. In other words, the Piagetian-designed activities of the experiment must overcome the rest of the activities the child is having and produce results that will demonstrate that the Piagetian paradigm is superior (if it is). To accomplish this in a brief period of a week or two—or an eight-day period as used by the following study—is expecting a great deal. We would urge researchers to use greater time periods which will allow thorough testing of Piaget's mental structure paradigm.

A study designed to test how the theories of Piaget, Vygotsky, and Ausubel relate to concept acquisition by young children was carried out by Billings (28). Thirty-one second grade classes were used by Billings in this research. The concepts used were "interaction" and "evidence of interaction" from Interaction and Systems (Science Curriculum Improvement Study). Ten classes received verbal instruction and concrete experiences, eleven classes received "only concrete experiences," and ten classes—the control group—received no prescribed treatment. An investigator-designed and validated test, the Concept Acquisition Test (CAT), was used as a pretest and posttest. The experiment was done over an eight-day period. A post-posttest was also administered a week following the experiment. The group who had concrete experiences with the concept scored higher on the post and post-post administration of the CAT than did the group receiving verbal instruction and concrete experiences, but not significantly so. Billings states that the findings contradict "Ausubel's theories regarding concept presentation to young children" but "do not give total support to the theories of either Piaget or Vygotsky." This investigator also reports
that the group who received only concrete experiences with concepts "used their own labels to describe 'interaction' and 'evidence of interaction'." Billings conjectures that perhaps their invention of labels for the concept accounted for the higher scores of the concrete experience group, and if so, support is added to Piaget's position that when teaching new concepts to children their own vocabulary should be used.

Research With the Piagetian Paradigm — Phase II: Curriculum

Fifty-six female and 46 male subjects furnished Enwieme (85) the data to evaluate the incidence of formal reasoning among education students in eight specific content areas. The students were completing certification in business, English, mathematics, music, modern language, science, social studies, and speech. Enwieme wished to know whether or not differences in the incidence of formal thought occurred among education students in those content areas as well as whether or not a relationship existed between formal reasoning and sex, socioeconomic level, parental education, and/or grade-point average. A specially designed test measured the achievement of formal thought. There were significant differences found among formal reasoning abilities for students from the eight content areas. No significant differences were found between formal reasoning patterns and sex, socioeconomic level, parental education, and grade-point average.

Chiappetta (53) made an exhaustive literature search to determine what research had found relative to the percentage of secondary school students who had moved into the stage of formal reasoning. Chiappetta concluded that "most adolescents and young adults do not appear to have attained the formal operational stage of cognitive development." This researcher also related research on cognitive development to science achievement and found that intellectual functioning below level is the rule rather than the exception. Chiappetta states that courses such as PSSC, CBA, CHEM Study, and BSCS are too abstract for most students. He also believes that curricula such as ISCS and IAC are "better suited to the majority of the students." The conclusion reached is that science education will profit from the research done only if the concern of teachers reaches the point "where it focuses on student learning."
Perhaps the most important conclusion of this study states that "Science teachers who are chiefly concerned about themselves in relation to their teaching role or about their adequacy as teachers will be unable to focus on the intellectual capabilities of their students..."

Using the Science-LA Process Approach (SAPA) curriculum, Johnson (138) investigated whether or not children who had reached the concrete operational stage of intellectual development were more likely to succeed on SAPA exercises which require multi-classification ability than were children identified as pre-operational. The matched pairs design was employed using one child from each developmental stage; the members of each pair had mental ages within approximately one standard error of each other. The students were taught nine sequentially arranged classification exercises as prescribed by SAPA. The concrete group performed significantly better than did the pre-operational group in three exercises requiring exhaustive sorting.

Griffiths (111) used 60 physics, chemistry, and developmental science college students to study the relationship between the cognitive level of students and their approaches and language used in solving an inclined plane problem. This researcher found that four types of responses were found to the problem. Two of the groups actively sought a general relationship among the variables and were not content until they had satisfied themselves about a relationship. These groups did show a difference in the language they used to explain the problem; Griffiths rated both groups as "formal." A third group was willing to quit at any time and further experimentation seemed to confuse them. A fourth group was passive and resorted to the question: "What else do you want me to do?" This research showed that the paradigms of science exhibit the characteristics of Piaget's formal reasoning stage. One group (12 students) exhibited characteristics that showed readiness for the formal reasoning paradigm of science. Six students (the second group) could be raised quickly to the necessary level and audio-visual materials, programmed instruction and other such procedures were useful with this student type. The last two groups, 70% per cent of the sample,
"present a totally different problem," with the last group "discriminated against the most." Much of the educational technology of the last few years has been aimed at this group with possible counterproductive results. In general, "the efforts to meet the demands of the students may be detrimental," Griffiths believes that the products of formal operational science cannot be indefinitely superimposed upon a nonformal cognitive structure. When this was repeatedly done students rejected physical inputs and retreated to their vocabulary: words not fully understood." This type of situation can stop intellectual growth. So for about 70 percent of Griffiths' sample, scientific paradigms which "exhibit the dynamic characteristics of Piaget's formal stage" are probably harmful. This led Griffiths to conclude that "the focus of higher education must be shifted from literacy to cognition."

A project to prepare 43 high school students from the inner city for engineering led McKimmon (130) to design and teach a two-hour per day laboratory experience in the logic of science. The students also experienced remedial reading, English, and mathematics. All of the 43 students were well motivated toward an engineering career and came from the upper third of their high school graduating classes. Twenty-eight of the 43 students reasoned with concrete operations and only seven had reached the highest level of formal thought. Forty percent of those students who were concrete operational moved into higher levels of thought. (Comment: A remarkable accomplishment in a six-week period.) The logic-of-science laboratory was successful in promoting formal reasoning among the sample of inner city students.

Karplus (145) and his colleagues have been studying the relation of Piaget's developmental levels concept to science teaching. A system which describes intellectual behavior which can be observed and which occurs during the concrete and formal reasoning stages has been developed and is explained here. The findings of this group of researchers have led to procedures which can be used to identify concrete and formal content concepts. If, for example, temperature is thought of as a reading on a thermometer or as warm or cold sensations, it is a concrete concept. If, however, temperature is defined as average molecular
kinetic energy, it is a formal concept. This report explains how equilibrium, or self regulation, fits the intellectual development model and how reasoning patterns are encouraged by the use of the learning cycle: exploration, concept introduction, and concept application.
Teaching Methods and Procedures

There is probably no research area in education that evokes as much criticism as does the area generally referred to as "methodology". Some critics state, for example, that genuine research in this area is impossible because the "teacher variable" cannot be controlled. If different teachers are used for two or more teaching methods, the critics take the position that the "teacher variable" is uncontrolled. If the same person teaches different groups of students using a different technique with each group, those critical of this research area use as criticism the observation that no one person can perform in an unbiased manner with more than one teaching method; surely the teacher will have a "favorite" method which will bias the results of the research.

Regardless of the controversies surrounding research into the efficacy of the different procedures used to teach science to learners from kindergarten through twelfth grade, research in that area continues year after year, and the effort made in 1976 was no exception. When the results of those efforts are studied, one is struck by their diversity. Beyond the classical experimental group-control group design, the profession of science education does not seem to subscribe to a reasonably defined paradigm or set of paradigms to use in conducting or evaluating research in the general area of science teaching procedures. The foregoing remarks are made neither as evidence of strength of the profession nor as a negative criticism, but as an observation.

Research in the Methods and Procedures of Teaching Elementary School Science

Much of the research effort in 1976 was aimed at studying the effectiveness of using the inquiry teaching method. Wall (303) studied the effects that three different teacher education experiences in inquiry teaching had on the teacher's perception of the process of science and the ability of the upper elementary grade students of those teachers to use science inquiry skills. The teachers had been in: (1) an NSF-sponsored, four and one-half weeks, Science Leadership Program; (2) an inservice program conducted by those in (1); or (3) teachers enrolled in a graduate elementary science methods course. A control group
was also involved. Teacher understanding of science processes was assessed by the Science Process Inventory (SPI). The ability of the student to use science inquiry skills was measured with the Test of Science Inquiry Skills (TSIS). The students of summer workshop teachers did not score significantly better on the TSIS than did the students of the control group teachers. On the SPI, summer workshop teachers scored significantly better than did the methods course and control teachers, but not significantly better than the inservice participants. Other comparisons made were inconclusive.

Working in Anne Arundel County, Maryland, Fertitta (1992) developed a plan for the infusion of science processes into an existing unified science program. One of Fertitta's goals was to develop a plan flexible enough to be adapted to other science programs. The plan was composed of three distinct parts. Part One led teachers to find that there were seven areas where science processes could be utilized in their existing science program. Part Two represented points of contact among concepts, sub-concepts and content objectives and component process skills. Part Three represented at least one sample activity for each of the 113 sub-concepts within the program. A total of 150 students from the program were evaluated and 54.7 percent of them showed a 20 percent gain in the ability to use science process skills; 65.2 percent gained 15 percent or more in ability; 76 percent gained 10 percent or more and 86 percent gained 5 percent or more. Guides titled Component Skills and Sample Activities were prepared. Fertitta concluded that the procedures designed and used were effective in infusing the teaching of process skills into an existing Unified Science Program.

Is single language instruction in science any more effective than bilingual instruction when working with bilingual students? Working in New Mexico under the auspices of the University of Washington, Juarez (141) gathered data from 104 fifth grade children from four different schools. Those schools had used bilingual education for at least four years prior to the study. The students were randomly assigned to one of four treatment groups at each of the four schools and the teachers were trained in the use of the materials to be taught. The students' achievement was first measured during the initial science instruction.
and again on the final set of science activities. In addition, student language preference (Spanish) and attitude toward science were also measured. The major finding of the study is that bilingual children instructed in Spanish learned science content and process skills as well as bilingual children instructed in English. But neither group learned content and process skills as well as students instructed bilingually in science activities presented in a subordinate and superordinate order. In addition, bilingual children preferred instruction in the two languages.

The relationship between curiosity and attitude and mode of instruction (inquiry and non-inquiry) was investigated by Metz (193). He used 200 third, fourth, and fifth grade children in his study. The Classroom Observational Record (COR) was used to assist in the data collection after pilot observation had been used to place teachers in inquiry and non-inquiry categories. Metz assessed children's attitudes towards science by using the non-verbal measure called Faces which is a part of the evaluation package of the Science Curriculum Improvement Study. The curiosity of a child was determined by the number of non-repetitive manipulations made and the number of meaningful questions asked. The results of this study, although not statistically significant, show a trend that suggests that children experiencing non-inquiry science become less and less satisfied with the discipline. The results showing that a child's curiosity is a function of the teaching method used were highly significant; students of teachers using inquiry were found to be much more curious than were students of non-inquiry teachers.

Roger T. Johnson (136) also focused his research on inquiry because he believed there is a "strong relationship between inquiry learning and a cooperative learning structure." His study was designed to determine if students involved in an inquiry-based science program perceived it as a cooperative or competitive goal structure. There were 108 sixth-grade students involved. Data were collected from three teaching situations: (1) a free inquiry setting using the Batteries and Bulbs unit of the Elementary Science Study (L), (2) the "Electrons in Action" unit from Concepts in Science and accompanying materials (TM), and (3)
and again on the final set of science activities. In addition, student language preference (Spanish) and attitude toward science were also measured. The major finding of the study is that bilingual children instructed in Spanish learned science content and process skills as well as bilingual children instructed in English. But neither group learned content and process skills as well as students instructed bilingual children instructed in English. But neither group learned content and process skills as well as students instructed bilingually in science activities presented in a subordinate and superordinate order. In addition, bilingual children preferred instruction in the two languages.

The relationship between curiosity and attitude and mode of instruction (inquiry and non-inquiry) was investigated by Metz (193). He used 200 third, fourth, and fifth grade children in his study. The Classroom Observational Record (COR) was used to assist in the data collection after pilot observation had been used to place teachers in inquiry and non-inquiry categories. Metz assessed children's attitudes towards science by using the non-verbal measure-called Faces which is a part of the evaluation package of the Science Curriculum Improvement Study. The curiosity of a child was determined by the number of non-repetitive manipulations made and the number of meaningful questions asked. The results of this study, although not statistically significant, show a trend that suggests that children experiencing non-inquiry science become less and less satisfied with the discipline. The results showing that a child's curiosity is a function of the teaching method used were highly significant; students of teachers using inquiry were found to be much more curious than were students of non-inquiry teachers.

Roger T. Johnson (136) also focused his research on inquiry because he believed there is a "strong relationship between inquiry learning and a cooperative learning structure." His study was designed to determine if students involved in an inquiry-based science program perceived it as a cooperative or competitive goal structure. There were 106 sixth-grade students involved. Data were collected from three teaching situations: (1) a free inquiry setting using the Batteries and Bulbs unit of the Elementary Science Study (L), (2) the "Electrons in Action" unit from Concepts in Science and accompanying materials (TM), and (3)
the "Electrons in Action" unit without the materials (T). The same
teacher taught all groups. A six-week instructional period was used
during which the classes met for 50 minutes daily. During the last
week of the experiment, 14 students per treatment group (a total of 42
students) were randomly selected for individual interviews. The
students were shown four pairs of photographs depicting an aspect
of cooperative or competitive classroom structure and two or three
sentences describing each of the pictured situations. The photographs
dealt with a general school atmosphere as well as with a science class.
The intent was to focus attention on the cooperativeness or competitive-
ness of each picture and to compare that decision to his/her own
classroom. There were no differences in the manner in which the general school
atmosphere was perceived. Less than one-fourth (24 percent) perceived
the school as cooperative and only 14 percent perceived division of labor
as part of the school. One hundred percent of the students in the
inquiry group (L) perceived a cooperative goal structure, 86 percent of
the TM group perceived a cooperative structure, and 50 percent of the
T-group perceived a cooperative goal structure. The differences were
significant. One hundred percent of the L-group perceived themselves
as working with materials and 100 percent of them preferred it. In
the T-group, 93 percent indicated a preference for working with materials
and many recognized that they were not. In the TM-group, 93 percent
recognized they were working with materials and 100 percent preferred
it that way. One hundred percent of the L and TM group and 93 percent
of the T-group preferred cooperation to competition. Johnson concluded
that the data supported the hypothesis that "inquiry-oriented science
classes were perceived by students to be more cooperative than the text-
book classes." Furthermore, all three groups preferred cooperation.

What relationships are there between the average ability level (as
measured by the IQ score) and class size and the teaching strategies a
teacher employs? Yeany(325) used videotape and collected data from a
random sample of 64 student teachers in grades three through six. The
videotapes were analyzed for level of teaching strategy by "two trained
raters" who were using the Teaching Strategies Observational Differential
(TSOD). The Elementary Science Activities Checklist (ESAC) was used to
collect data on how the students in those elementary science classrooms perceived the teaching strategies being used. Yeany's results were not statistically significant, but the "trend seems to be toward more direct teaching as class size increases." The results involving class ability were also nonsignificant.

Using the Instructional Practices Questionnaire (IPQ), Patterson (221) classified 17 fifth-grade teachers as high-individualized or low-individualized instructors. The students of those teachers completed the Bristol Study Skills instrument (which Patterson called "a science cognitive instrument") and questionnaire which was designed to "validate teacher responses on the IPQ." The Bristol instrument contains five subtests: properties, structures, processes, explanations, and interpretations. Patterson investigated how a classroom rated as high or low individualized by the IPQ was related to the achievement of science-related cognitive skills by the children. He found that those classrooms rated as high in the individualized instructional situation produced students who ranked highest in science-related cognitive skills.

Teacher behavior patterns in the elementary science classroom was the focus of research done by Penick, Shymansky, Matthews, and Good (225). These researchers defined two contrasting teaching strategies, which were labeled "teacher-structured learning in science" (TSLS) and "student-structured learning in science" (SSLS). The difference in the two strategies rested in the amount of directiveness used by the teacher in an activity-centered science classroom. In other words, the research centered upon how student classroom behavior was influenced by teacher behavior. The population was composed of 250 students and eight teachers—two teachers and classes in grades one through three and one class and one teacher for grades four and five. Fifty students from each grade level were randomly assigned to the SSLS or the TSLS treatment; sex and race were equalizing restrictions. The "Teacher Behavior" and "Student Behavior" sections of the Science Curriculum Assessment System (SCAS) Classroom Interaction Categories were used to collect student and teacher data. The SSLS classroom produced fewer patterns containing non-lesson-related behavior and greater clustering patterns which resulted in a more predictable set of behaviors than in the TSLS classroom. Perhaps
providing directions and other restrictive behaviors provided the less-than-conforming student with a more relaxed and less anxiety-ridden frame of mind. These researchers concluded "materials, activities, and curricula, in general, cannot be made 'teacher proof!' Teachers need to be aware that their behavior patterns greatly influence the educational outcomes of a classroom.

Penick (224) extended the study of the SSLS and TSLS classroom strategies to growth in creativity and used the Torrance Tests of Creative Thinking (TTCT) to measure that attribute. The population used was 51 fifth grade students. He found that there were no differences in the verbal creativity data from students who experienced the two teaching strategies but that the figural composite score on the TTCT demonstrated that the SSLS group scored significantly higher than did the TSLS classes.

Few teachers would believe that teacher behaviors convey the same message to all children in that class. There is evidence that dyadic interactions are useful in the classroom. Shymansky (268) studied the relationship between certain aspects of one-to-one interaction between a teacher and a student in an activity-centered classroom and that student's behavior during the remainder of the lesson. The time of the interaction and whether it was verbal or non-verbal and the student's classroom behavior were the specific variables studied. During a five-week period, 78 student observations were made in fifth grade science using the "Student Behaviors" portion of the Science Curriculum Assessment System (SCAS) Classroom Interaction Categories. The data collected suggest that lengthy dyadic interaction between student and teacher may reduce productivity and learning effectiveness and be thought of as interference by the student. Some teacher-student interaction is needed but the teacher and the materials combine to provide the learning environment.

Using the evaluation materials prepared by the Science Curriculum Improvement Study (SCIS)--with some modification--to measure achievement in twenty third grade classes, Norris (211) compared the effects of individualizing the Interactions and Systems unit of the SCIS. The SICS
material had been adapted for the individualized learning system and was entitled Science Curriculum for Individualized Learning (SCIL). Norris used the pretest, posttest, post-post-test design. There were no differences in achievement between the SCIS and the SCIL groups on the post-test but the SCIL group scored significantly higher on the post-post-test. Norris also studied the students' attitudes toward science and science class using a survey instrument entitled "How I Feel About School and Science"; that instrument has five different scale factors. The results between the SCIS and the SCIL groups were mixed; one group excelling on some scales and the other group excelling on other scales.

Sixth grade students in Winona, Minnesota, were the population Fick (93) used to study the acquisition of non-verbal cognitive abilities and productive thinking skills when experiencing the ESS units Tangrams and Geo Blocks. An experimental and control group design which utilized a post-test was employed. The instruments used were the Verbal Form A and Figural Form A of the Torrance Test of Creative Thinking and the Thorndike-Hagen Nonverbal Cognitive Abilities Test. The experimental group scored significantly better than did the control group at the 0.02 level on non-verbal cognitive abilities: beyond the 0.01 level on verbal fluency, at the 0.02 level on verbal flexibility, at the 0.20 level on verbal originality, and at the 0.001 level on figural flexibility, figural originality, and on the composite scores of all divergent thinking tests.

Nelson and Abraham (210) believe that the post-laboratory discussion is being overlooked by many teachers of elementary school science. Furthermore, these researchers believe that the accommodation of mental structures—according to the Piaget model—can result from an effective post-laboratory experience. The teacher can, according to these investigators, use the laboratory experiences to move the students toward the symbolic mode of thought using what is called a probing strategy. A probing strategy is opposed to a teaching procedure which asks for student observations and then accepts or rejects those data on some notion of correctness. Nelson and Abraham studied the effect of the two strategies on student increase in observation, inference, verification, and classification. The sixth grade population of an inner city and a suburb were
used, a total of 116 students. The Classroom Observational Record (COR) and the Inquiry Skills Measure (ISM) were used to gather data. The ISM was used in a pretest, posttest manner. The investigators issue several precautions regarding the generalizability of their findings but cautiously conclude that to increase the quantity and quality of inferences, the probing strategy is better than a non-probing strategy. Furthermore, a probing strategy increases the quantity of observation made in an urban school but the non-probing strategy is better in an urban school.


Mosley and Bell (201) report a study undertaken to examine the influence of objective statement specificity on student learning resulting from the independent laboratory-based Physical Science Investigation Program. The study involved three teachers and 138 eighth grade Regents students from six intact classes. Each teacher taught one section using specific behavioral objectives and one using non-specific objectives. In both treatments, objective statements were presented prior to the instruction of the given unit. A non-randomized control group, pre-posttest design was used. Pretest scores showed no significant differences between classes. Two-way analyses of variance of the posttest generated F-values that were not significant for the teacher and interaction source of variance. The treatment effects were significant (p < .05). Higher means were obtained by the groups provided with the specific behavior objectives. A post-hoc questionnaire indicated student perception of the use of behavioral objectives as being helpful generally, helpful in achieving higher grades, and helpful in providing guidance through the unit.

How do the forms of questioning techniques teachers use influence student cognitive achievement and retention in junior high school science? Guthrie (112) identified two distinct questioning strategies: Type I questions were those formed primarily in the third person and therefore treated the student as an outsider and solver of the problem posed. Type II questions were those which were designed to place the student right at the focus of the problem and to convince the student he/she,
should assume the responsibility of responding to the question from a "quasi-experience base." Guthrie utilized an experimental group control group design. Using Type I questions, however, did not significantly improve student achievement or retention of subject matter, increase the percentage student-talk versus teacher-talk (when compared to earlier studies of Flanders and others), nor did Type I questions significantly increase the length of student responses.

Carlson (48) evaluated an individualized contract-directed high school chemistry course and compared that evaluation with a group-instructed, teacher-directed course. He found no significant differences between scores on cognitive measures or affective measures of course preference, attitudes toward science or toward a specific class, or self actualization. After one academic year, a statistically significant treatment difference favoring students in the group-instructed, teacher-directed classes were found on an investigator-designed, summative, criterion-inferenced examination. After the same period middle academic ability students from the individualized treatment produced higher scores on the Tennessee Self-Concept Scale.

Teacher-directed instruction where the students progressed together through a prescribed sequence of concepts with access to a set of behavioral objectives was compared to the technique of letting the student progress through the same set of concepts in his own time (self-paced) with access to the same set of behavioral objectives by Ritter (249). The materials used in instruction were BSCS Biology: Molecules to Man. Ritter found no significant differences between the groups in short-term, mid-term, or long-term achievement nor did he find any significant differences between the groups in long-term process skill improvement.

Monaco and Szabo (197) evaluated the team approach for teaching biology. A sample of 147 sophomore students was divided into six sections; each section was randomly assigned as control or treatment. The team consisted of three instructors. Each member taught a control group for an entire year. After an introductory session of nine weeks, team members began teaching their specialty (botany, genetics, or microbiology). Students in the treatment groups worked with a different team instructor each.
for nine weeks. The criteria used in this study consisted of T-scores from five separate objective tests. The experimental design was a 2x3x5 factorial analysis of variance with intact class sections randomly assigned to experimental or control groups. Variables included two groups, three instructors, and five content areas. Among the conclusions were that academic biology students acquired more biology subject matter knowledge when they were instructed by a team rather than by an individual instructor and that there were significant differences among each of the instructors relative to student achievement in biology.

Riban (246) distinguishes a field study from a field trip. According to Riban, a field expedition is of at least a few days in duration whereas a field trip is one or more periods on the same day. Sixteen junior high school students participated in a multi-day field study in the Grand Canyon in Arizona. A total of 85 items with geologic emphasis were selected from the preliminary tests of the Earth Science Curriculum Project and completed by the field study. Later the test was administered to "a group of advanced students of science from the same school who had not participated in the fieldwork." Neither the field studies group nor the comparison group had completed courses in the earth sciences. Riban states that the learning displayed by the field study group "excelled any reasonable expectations." The achievement of the field study group was "completely superior to that of the comparison group of science-prone students." Some statistical data are offered that substantiate Riban's claim.

Mayfield (185) studied the factors which affect rationality in a group solution to a problem. Sixty 11th and 12th graders were divided into groups of five and given the NASA "moon survival" problem. After participating in the group situation, each group member completed Bales' Interpersonal Rating Form A (revised). Mayfield cautions about the generalizability of his findings, but several are worth noting in relation to outcomes from group work. Most individuals' solutions were not as correct as later decisions made by the entire group, and an individual's second solution tended toward the group's solution. At some point in the session after some useful and some useless items had been identified, the group devalued the problem and the remaining decisions were seen as...
unimportant. Mayfield found that there were four primary factors which contributed to non-rational decisions being made. Those factors were concession, force, rules-of-thumb, and emotion.

One hundred and six chemistry students from eight classes, and involving four teachers, were the subjects Abraham used to study grouping as an instructional methodology. Specifically Abraham studied whether or not grouping by divergent thinking potential influences the amount and kind of student verbal interaction during a science inquiry. Divergent thinking potential was operationally determined by each student's composite score on the verbal portion of the Torrance Tests of Creative Thinking (TTCT). Those scores were then used to form homogeneous and heterogeneous groups. Science inquiry was used to stimulate the formulation of hypotheses and the design of experiments. The students were told they were in an experiment to see how they went about solving science problems. A film was then shown to the students and they were given two questions to discuss. Abraham summarized the results of this study as meaning that there was an interaction between grouping and divergent thinking potential; and that interaction showed a trend which favored medium divergent-thinking persons in homogeneous groups and high divergent-thinking persons in heterogeneous groups. When the mean number of statements made by persons in both homogeneous and heterogeneous groups who have high, medium and high divergent thinking potential is considered, homogeneous grouping is an effective way of encouraging greater amounts of verbal interaction.

Bates and Watson investigated whether the interaction of teacher, student, and classroom characteristics had a significant effect on classroom climate. Linear regression techniques were used to predict selected scales of "Anderson's Learning Environment Inventory" (LEI) from sets of orthogonal variables which were representative characteristics of the teacher, student group, and class structure. Eight scales accounted for 21 percent to 63 percent of the variance for the model group of 72 classes. Data for the study were drawn from the Harvard Project Physics Summative Evaluation Data Bank which represents a true national random sample of 54 physics teachers with 3085 students in 103 classes. One of the observations of these investigators relates to the
distribution of the sexes within classes. Classes with a relatively high proportion of females and a climate that was more democratic and intimate and contained less friction and apathy. The students in those classes were evaluated as being more satisfied than in classes where the proportion of females dropped. These investigators ask whether or not women are attracted to, or produce, that type of environment.

Which combination of lecture, discussion or direct-indirect teaching produces optimum student growth? Markell and Mayer (177) used 19 teachers, all of whom were participating in an NSF-supported College School Cooperative Science Earth Science Program, and 18 classes to gather data to answer that question. Actually this study involved the search for statistically significant linear and curvilinear (optimal) relationships between assessed student perceptions about classroom instructional procedures and pre-posttest changes in students' understanding of science concepts, attitudes toward science, and development of interests in science. The 19 teachers and 38 science classes involved in the study also participated in a research plan that included the administration of several instruments in a pre-posttest design during the academic year. Data were collected using: (1) concept-process tests, (2) the Science Classroom Checklist (SCACL), (3) the Silance and Remmers Interest Scale, and (4), an instrument to measure student attitudes toward science and scientists (BATSS). For this study, the Instructional Activities Instrument was developed, piloted, and administered near the end of the year. A direct/indirect score on this instrument was compared to corresponding pre-posttest change scores as measured by the other four instruments. A linear relationship was found between student attitude toward science as a subject and teacher direct-indirectness. Certain instructional methods were found to lose favor with students if frequently used.

A learner's attitude toward what is being studied and his/her anxiety about the course, teacher, peers and/or examinations are certainly factors in influencing the amount and kinds of learning which will be accomplished in any class. Two studies reported during 1976 addressed themselves to the general areas of attitudes and anxieties associated with different teaching methods in secondary school biology.
Clay (57) investigated attitudes toward education, critical thinking ability and specific affective behavior. She compared students who had studied the BSCS environmental module "Investigating Your Environment" (IYE), those students using one of the BSCS Biological Sciences version and those studying Modern Biology. The sample of students came from 32 high school classes throughout the United States. Sixteen classes were involved with the IYE and sixteen classes received the other two instructional methods. The IYE treatment enrolled 355 students, and 375 students received the other two treatments. Instrumentation consisted of the Attitude Toward Education Survey; Cornell Critical Thinking Test, Level X; and the Biology Students Behavioral Inventory, Form C. Clay found that the ability of students who had experienced BSCS Biological Science was significantly increased but no other differences were found among the students on any of the other variables.

The design and formative evaluation of a two-weeks minicourse dealing with the biochemical, physiological, and psychological aspects of emotional stress was completed by Betkouski (26). A total of 431 secondary biology students from four different schools (six different teachers) were the subjects for the experiment and were randomly divided among three treatments—A, B, and Control. Treatment A consisted of the two-week minicourse followed by a four-day session on coping with test anxiety (a modified version of Coping With Test Anxiety: A Guide, was used). Treatment B consisted of the coping-with-test-anxiety component only. The two treatment groups and the control took a series of pretests consisting of the A-State scale of the State-Trait Anxiety Inventory (STAI-Form X) and the Test Anxiety Scale (TAS) prior to a teacher-constructed test administered before the treatments began. An achievement test covering the material in the minicourse was also given as a pretest. The posttests, consisting of the same instruments, was administered at the conclusion of the treatments; and the STAI and TAS were given prior to a teacher-constructed test six to nine weeks after the treatments concluded. The research sought differences in the anxiety levels and knowledge levels among the groups. No significant differences in anxiety levels were detected. Significant differences in knowledge levels measured by means of the achievement test were found:
Group A scores were significantly higher than both Group B and Control scores; Group B scores were significantly higher than the Control scores.

During the last few years the type of methodological procedures usable with the mentally handicapped has, in general, received universal attention. What methodological procedures in science are effective with this particular group of learners is also receiving consideration. Bacon (16) used the BSCS curriculum Me Now to study the achievement and retention when that material was taught by "a deductive method and a discovery method." Bacon used the BSCS objective test accompanying the Me Now units, monitored teacher behavior with the Science Curriculum Assessment System, studied the reading abilities of the students, and tested the interaction effect of reading ability and treatment. He found no significant differences.

Teaching methodology to use with the educationally disadvantaged student also has received special attention recently. Kahle, Nordland, and Douglass (142) studied four teaching procedures--timed audio-tutored, self-paced audio-tutored, a traditional classroom procedure with study guides, and a traditional classroom approach with visual aids as teaching procedures. The first-year biology students of these teachers in an urban high school constituted the 160-member sample which was divided into six treatment groups. The age range of the sample was 14 years to 18+ years. The treatments were conducted over a six-week period. The groups were judged as equivalent at the start of the experiment. Self-graded, multiple-choice, formative tests were used throughout the study. A pretest with an equivalent form of posttest was also used. A significant difference in achievement was found in favor of the self-paced audio-tutored group.

Kahle, Douglass and Nordland (143) also analyzed learner efficiency when individualized and group-instructional teaching procedures were utilized with disadvantaged students. "Learning efficiency" was defined as achievement point gain per classroom minute. The subjects, procedures, and treatments described in the previous study were used in this study. The conclusion was drawn that learning efficiency "as described by net gain mean/total time mean, is increased by an individualized instructional system."
The separation of the teaching methodology employed in a classroom and the materials used is often artificial. Most certainly materials can be designed around specific teaching procedures which will lead to the achievement of specific goals. Coleman (61) designed a non-simulation game technique, Genetico, to assist students in the study of genetics. After testing and evaluating Genetico, Coleman concluded that significant cognitive learning resulted from its use, and that "race and scholastic achievement predicted cognitive achievement for the study population." Genetico did not, however, influence attendance and punctuality. Johnson (137) undertook the development of materials which would promote and permit investigation of problem recognition and subsequent research questioning ability among students. A system named Learning to Ask Research Questions (LARQ) went through four developmental phases. The experimental population was tenth grade biology students. Johnson concluded that LARQ is a useful tool for developing and investigating research questioning skills in students, taxonomic questioning is easier for students than cause-effect inquiry, relationships between pairs of skills is hierarchical in nature, and the students developed a questioning strategy in the course of several lessons.

Askham (12) designed and carried out a study to find out if nine-to-twelve-year-old children classify plants growing in a seminatural environment the same way they group non-natural objects in a classroom setting. He found that this age group used more varied and complex strategies to classify the plants than to classify the non-natural objects. In fact, says Askham, children using more and more complex classification categories suggests that there are "many latent classification schemes as yet unassessed." Since this increase in classification and its complexity was motivated by using natural objects (plants) in a seminatural environment, "further research should be moved out of the classroom or laboratory and into a more natural setting," and non-natural objects should be replaced with natural objects.

When reviewing the thirty research studies included in this section, one is struck with the many problems inherent in studying teaching methodology. Some of those problems are beyond the control of the science education researcher and he/she can, at best, only attempt to control
such variables and, at least, recognize that they exist. There is, however, one variable that is evaluated so differently in methodological research studies that comparison of those studies or attempting to apply the research results in a setting different from that of the original researcher is difficult, if not impossible. The variable being referred to is student achievement with respect to content. In the research just reviewed that variable was defined as everything from the application of methods of stating a research question to how many achievement points a student gained per minute.

There are those who believe that an entire spectrum of meanings of such a variable as achievement is acceptable and healthy for a research area. Possibly that position is true. But the fact that there is such great variability in the meaning of achievement in the profession reflects, we believe, the lack of a common theory base for the practice of and research in the profession of science education. We recognize that the establishment of one or more theory bases for science education is an undertaking of monumental proportions and will require a great deal of time. We believe however, that if science education is to continue as a viable research area, that undertaking must be started. In the interim, each researcher can make a significant contribution to the profession if he/she will carefully define how student achievement with respect to content is being viewed.
The Education of Teachers

Perhaps no topic in the field of education evokes more divergent opinion than the general area of teacher education. The laws governing teacher certification range from absolute prescription of specific courses in some states to a general statement in other states of the number of hours of college credit required. The types of educational experiences a kindergarten teacher must have are generally prescribed by law but the law is silent about what types of educational experiences the persons teaching future kindergarten teachers will have.

Has research been helpful in resolving the problems inherent in the area of teacher education? In general, the answer to that question is probably "no," but research in the past has certainly been helpful in isolating specific problems in need of solution, providing data for particular questions involved in the general teacher-education picture, and suggesting improved ways of accomplishing specific objectives in the field. A summary of the research published in the area of science teacher education during 1976 is reviewed in this chapter. The review is divided into four general categories: elementary school teacher education, secondary school teacher education, research concerning teacher education from kindergarten through grade twelve, and research into inservice education.

Research in Elementary School Science Teacher Education

The profession of education has generally agreed that both preservice and inservice education are necessary in teacher education. Which of those experiences is more effective? Does one stage in a teacher's career make that person more amenable to change than another? Stephens (278) compared the influence of a natural science course on preservice and inservice elementary teachers and on inservice primary and intermediate teachers with respect to their agreement with the philosophy to teaching science of the Elementary Science Study (ESS). He used 29 preservice and 22 inservice teachers. The eight-week treatment--two hours and fifteen minutes each week--consisted of activities from ESS materials presented as the developers intended them to be used with children. The subjects were evaluated by responding to a series of
multiple-choice, behavioral items, each presenting a teaching situation. Stephans found that the instruction did not produce any significant differences in agreement with the ESS philosophy between preservice and inservice teachers. He did find, however, that with the inservice group the treatment had a greater influence on the intermediate teachers than on the primary teachers.

Cotten and Evans (68) used an experimental group of 70 elementary teachers and a control group of 32 to study whether having teachers perform inquiry-investigative activities with a written model would lead to greater facility with process skills than could be expected from elementary school teachers who did not perform the inquiry-investigative activities. The treatment required a minimum of seven weeks. The process skills considered were observing, measuring, communicating, classifying, prediction, inferring identifying variables, and controlling variables. A 17-question, 75-item process skill instrument was developed and used in a pretest, posttest design. The investigators concluded that "the two groups can be considered equivalent prior to the experimental treatment." The experimental group showed statistically significant performances in observing, classifying, predicting, identifying variables and controlling variables. In addition, the experimental group significantly outgained the control group in inferring. These researchers used a subsample of sixteen from the experimental group, and, using the Instrument for the Analysis of Science Teaching, Version 2, found that the ratio of closed teacher questions to open teacher questions was significantly reduced. The percentage of continuous teacher lecture was also significantly reduced. Students of these teachers exhibited a significantly more positive attitude toward science instruction, and engaged in significantly more nonverbal activities and peer interactions.

Three procedures designed to encourage and educate preservice elementary teachers in the use of inductive/indirect strategies in teaching science were evaluated by Yeany (324). Procedure One, Strategy Analysis Level, consisted of educating future teachers in science teaching strategy analysis using the Teaching Strategies Observation Differential (TSOD). In the second procedure, Modeling Level, the subjects viewed video-tapes of model science lessons which used inductive/indirect
teaching strategies. A third treatment group received a combination of the first two treatments and a control group viewed films neutral to the treatments. The TSOD and the Elementary Science Activities Attitude Sort were used to collect data. The combination group adopted a more inductive/indirect teaching style than did the control group. Evidence is supplied which shows that activities that will "significantly and positively affect the science teaching style and attitude of pre-service elementary teachers can be designed."

Yeany (323) also designed a study to assess the effects of three treatment levels involving microteaching with videotape playback and strategy analyses on the teaching strategies selected by secondary science teaching methods students. Three groups of ten students each were randomly assigned to three treatment levels. All subjects taught a pretreatment and a posttreatment peer group lesson; both were videotaped. Each student viewed his own prelesson, and before the taping of the postlesson, the following instructions were given: group 1 - no instruction (control); group 2 - instruction in how to use the Teaching Strategies Observation Differential (TSOD) to assess the prelesson; group 3 - same as group 2 plus the requirement that a tape-viewing session with the instructor of the course be carried out. Trained raters used the TSOD to analyze the postlessons for the degree of directness/indirectness exhibited by each subject in their teaching style after treatment. Using ANCOVA, significant differences were found between group 3 and group 1 and between group 2 and group 3, indicating a cumulative effect of the treatment procedures. Yeany suggests that perhaps this research shows that while there may be certain advantages to individualized or self-paced teacher education programs, opportunities for the student to interact with the professor are important.

The elementary school science "methods course" is usually a college experience had by preservice teachers. A great deal is expected of the "methods course." The instructor is expected to combine the students' prior courses in educational philosophy, pedagogy and science—not infrequently very traditional "show-and-tell" science courses—and lead them to develop a thorough understanding of how to teach science and a
procedure for doing so. The number of semester hours devoted to this undertaking is not infrequently less than three. Seldom is so much output expected from so little input. Perhaps for the foregoing reasons the elementary school service methods course has attracted researchers in science education for some time. The research published in 1976 was no exception.

Barufaldi, Huntsberger and Lazarowitz (18) designed and carried out a study "to investigate changes in attitude of pre-service, elementary education majors toward inquiry teaching strategies." The "methods course" furnished the subjects for the study; 146 students were involved. The subjects, however, had elected to study the "methods course." The subject matter of the course was the "content, methodology, and processes of science that are employed in the teaching of a modern course in elementary school science." A modified form of the Inquiry Science Teaching Strategies (ISTS) was used to measure each subject's attitude toward inquiry teaching. The students were pretested (the first day of class) and posttested (the last day of class). The results demonstrated that a significant change in attitude toward inquiry teaching took place between the two administrations of the ISTS.

A total of 246 students was used by Shrigley (267) to establish the reliability of an instrument to measure the effectiveness of the instructor of the "methods course." The coefficient alpha index of reliability was 0.83. Shrigley then surveyed 286 third-year elementary education students enrolled in science methods courses at four midwestern teacher-education institutions. The data demonstrated that if a science methods course instructor is to have high credibility with college students he/she: (1) referred to practical teaching activities in class, (2) had taught science to children, (3) assumed responsibility for teaching science content, (4) modeled teaching modes similar to those proposed for children, (5) assisted science professors in designing science content courses, (6) counseled student teachers, and (7) assisted inservice teachers. Authoring science textbooks or science methods textbooks and being involved in research did not add to the instructor's credibility with college students.
Eighty-two elementary education majors constituted the group Strawitz (280) used to assess an undergraduate science methods course. The subjects involved in the study were 82 elementary education majors at Louisiana State University. Students not in the course served as a comparison group. The treatment group participated in science activities selected from ESS, SCIS, and SAPA materials. Students were encouraged to develop individual teaching styles, consider feedback from the instructors and peers, and formulate their own concepts about teaching strategies and roles appropriate for the teaching of elementary school science. The control group discussed topics such as planning for effective teaching, individualizing instruction, and evaluation. Scores on an attitude inventory were used as a premeasure as well as a postmeasure. Scores on Form E of the Rokeach Dogmatism Scale were used as a measure of the belief systems of the students. The dependent variable was teacher attitude; the independent variables were method of instruction and teacher dogmatism. Findings suggested that the instructional method was very effective in changing the attitudes about teaching science. It was noted that these changes due to the treatment were for the most part unrelated to the belief systems of the students.

Utilizing a "methods course" which had a strong field component, Piper (229) studied the changes in attitudes of the students and the cooperating teachers. Thirty-six students enrolled in an elementary science methods course were randomly selected and given an instrument using Osgood's semantic differential approach the first week of class, the sixth week on campus prior to field experiences and the thirteenth week following field experiences. The elementary teachers who had observed the university students teaching in their classrooms were given the semantic differential prior to the students' teaching and following the students' teaching in their classrooms. Five factors were used to test the change in attitudes: (1) Science in High School; (2) Science as Remembered in Elementary School; (3) Science Methods Course; (4) Science in the News; and (5) Teaching Science to Children. At the end of five weeks of on-campus activities, the university students changed their attitudes in a negative direction toward (1) and (2), but in a positive direction toward (3) and (5). Following the
reality experiences with children in the field, attitudes toward (3) and (5) were even more positive than prior to field experiences. The attitudes of classroom teachers who had university students teach science activities in their classrooms also changed in a negative direction toward (1) and in a positive direction toward (5).

The elementary school science "methods course" is used for many things. Bluhm and Hungerford (33) found that using the "methods course" as their instructional vehicle could make a statistically significant improvement in students' ability to define environmental education and in their acquisition of concepts concerning ecology. The research was conducted using a pretest-posttest design. The investigators developed their own evaluation instruments. These researchers believe that the instructional model used appeared to influence significantly undergraduates' perceptions concerning the definition of environmental education.

The Teacher Corps was designed to provide persons who held a college degree with "a repertoire of basic skills for teaching various school subjects, including science." Wilson, Koran and Koran (312) claim that "few if any summative evaluation reports describing the effectiveness of Teacher Corps Programs exist." These researchers worked with a multi-racial group all of whom had degrees and whose Graduate Record Examination (GRE) scores were between 650 and 1400 (a mean score of 850). The students received a M.A.T. degree if they successfully completed the two year program. These students experienced a learning procedure that can best be described as modeling, practice, feedback, reinforcement using the Science - A Process Approach elementary science program. The data used in the evaluation were collected by evaluating the students at entrance to and exit from the program on nine "target behaviors." Those target behaviors were: (1) assessing behavior, (2) evaluating instruction, (3) probing, (4) giving instructions, (5) vocabulary development, (6) establishing set, (7) reinforcement, (8) eliciting observation and classification, and (9) establishing models. Significant behavior change was found among the students in behaviors (1), (2), (3), (7), and (9). The investigators concluded that the methods of training used were successful even with students who had low GRE scores. They
content that the major advantage of the training model used is its flexibility and that the model would be advantageous with many inservice and preservice teacher programs where local personnel are used as master teachers who may not know the nature of some of the basic skills and desired outcomes of the teacher training program. Most of the concerns of the Teacher Corps are also the concerns of the preservice and inservice teacher education programs. The inference is that those programs should look at the Teacher Corps operation.

Tomera, Hungerford, and Walding (295) conducted a survey of professional scientists, science teachers, and preservice elementary education majors. They were sent a questionnaire consisting of the Perceptions of Science and the Scientist Inventory (POSSI) instrument. The instrument comprised six short-answer questions dealing with: (1) a definition of science, (2) the differences between science and technology, (3) the greatest contribution of science to man's culture, (4) the most important intellectual process to the scientist, (5) the paramount ethical responsibility of the scientist, and (6) the value of science education to human beings. Included in this report are analyses of the various types of answers given to these questions and the percentage of each sample group that responded with each answer-type. In general, the results of the study showed that scientists and science educators had similar perceptions of science but that the pre-service elementary education majors often had no response to questions or held views contrary to those of scientists and science educators.

Research in Secondary School Science Teacher Education

How can specific teaching behaviors be taught to preservice teachers? Rezba and Andersen (244) asked whether or not using a system which involved a model teacher laboratory behavior and a microteaching environment would cause preservice teachers to change their interaction with students in such a way as to complement inquiry-centered teaching. Randomly assigned groups from a sample of 56 participants were presented a printed model of laboratory teacher behavior and a perceptual model employing those behaviors—a videotape of a teacher performing the behaviors found in the printed model. The students then engaged in peer group microteaching. All content for microteaching was taken from the work of the Intermediate Science Curriculum Study (ISCS) and all teacher behavior models were based upon materials from the ISCS.
Following the instructional treatment, the participants demonstrated significant increases in the use of accepting feelings; praise or encouragement; modifying students' ideas; accepting or using students' ideas; all types of questions; and all indirect verbal behavior. The treatment also significantly decreased the use of lecture. The treatment did not show a significant impact upon acknowledging, applying, or summarizing students' ideas nor was the use of cognitive-memory questions, giving directions, or criticizing and justifying authority significantly influenced by the treatment.

The ISCS program was also used as a research vehicle by Knight (149) who investigated the value, to preservice teachers, of classroom experience with the ISCS program. A total of 31 preservice teachers was given 18 hours of preparatory instruction on simple teaching skills and science content. They then had 20 hours of experience as a teacher would have in an ISCS classroom. The Nature of Science Scale, a Word Association Scale, and a Teacher Profession Questionnaire were used in evaluation. Measures were also taken to establish that the preservice teachers understood the philosophy of the ISCS classroom and that they demonstrated they knew the philosophy while working in the classroom. After the ISCS experience, the preservice teachers expressed interest in a broader science background, an increased preference for junior high school teaching, student reading problems, individualized instruction and the evaluation of student progress. A decreased interest in the use of blackboards, teacher aides in the classroom, and writing relevant objectives for students was measured.

There is no doubt that the exact competencies needed by secondary school science teachers are not known. Qureshi (239) designed and conducted a study to identify a set of validated competencies that could be used in educational programs for that group. Interviews with 21 randomly-selected high school teachers led to the identification of twelve competencies. Those competencies were: (1) A teacher must know his subject and keep striving to update his knowledge; (2) A teacher must have a good rapport with his students; (3) A teacher must recognize the individual academic abilities of his students and try to help and encourage each of these students; (4) A teacher must make his lessons interesting without wasting time; (5) A teacher must be able to control discipline problems to protect the learning experiences of his students; (6) A teacher must plan lessons
in advance with the idea of presenting the scientific concepts and ideas in an organized and clear manner; (7) A teacher must foster unbiased, independent and critical thinking in his students; (8) A teacher, especially the one who is teaching low and average ability students, must relate the scientific concepts and ideas he is teaching to the daily life experiences and needs of the students; (9) A teacher must evaluate, in a traditional or any other practical sense, the academic progress of his students and make the results available to them as soon as possible; (10) A teacher must, appropriately, respond to sudden diversions of student thoughts; (11) A teacher must make appropriate actions to instruct the students about laboratory safety practices; and (12) A teacher must fulfill his professional responsibilities. Validation of the competencies was accomplished by using the panel-of-judges technique. Through the use of inter-judge reliabilities, inter-teacher reliabilities, and the Chi-square statistic, the conclusion was drawn that only competencies 1, 2, 4, 5 and 6 were validated.

When a person engaged in the education of science teachers looks at a class in which they are enrolled, the question regarding who will finally achieve certification is intriguing. What criteria could be used to predict success? Research has not yet produced a definite answer, but Neff (209) has demonstrated that the Scholastic Aptitude Test of the College Entrance Examination Board (SAT) is not capable of making that determination. All students at Ball State University seeking provisional certification to teach biology between 1965 and 1974 comprised the sample, which was divided into two groups for the study. Group one included all persons who achieved certification (331 students) and group two was made up of those students who did not receive certification (244 students). When the discriminant analysis technique was used with the SAT verbal scores and mathematics scores, no significant differences were found between the two groups. The conclusion was drawn that SAT scores were not reliable predictors of "provisional certification of biology teachers at Ball State University during the period between September of 1965 and June of 1974."

Do the actual experiences which a student has in college really influence the attitudes of a future teacher? Perry (226) designed and carried out a study to test whether or not a learning module which presented a laboratory-oriented, student-centered inquiry method of science teaching,
presented in a positive way, could affect preservice secondary science teachers' attitudes toward the teaching module. This researcher also studied what other factors might affect the change in attitude toward the module. Perry used 50 perservice secondary science teachers enrolled in a "science methods" course. The instrumentation used in the research included the Semantic Differential Test of Teacher Attitudes (SDTTA), the Sixteen Personality Factor Questionnaire (16PF), and the Ames Philosophical Belief Inventory (APBI). The 50 students were divided into two groups on the basis of performance on the SDTTA. The sample was ranked from greatest positive attitude toward student-directed learning to greatest negative attitude toward student-directed learning; in other words the better group was oriented toward teacher-directed learning. The sample was divided at the median. The study's results demonstrated that a significant improvement of attitudes toward positively viewing student-centered, inquiry science teaching occurred during the completion of the learning module. The two groups differed significantly on four personality factors. The student directed group was more enthusiastic, forthright, apprehensive, and tense. No differences were found based upon sex, philosophy or the number of education credits earned. The major predictors of positive change in attitude, in order of importance, were the personality factors of tenseness, maturity, stability, forthrightness, intelligence, and laxity (carelessness of social rules).

Results of a study reported by Stronck (281) involved lessons taught in biology by 58 student teachers which were evaluated by other student teachers and by the students being taught. The data were gathered by using an evaluation instrument which had nine categories: (1) knowledge of subject matter, (2) attitude toward subject, (3) explanation, (4) speaking ability, (5) attitude toward students, (6) personality, (7) evidence of planning in procedures and materials, (8) students' attention, and (9) objectives. A total of 2,399 questionnaires was completed by the students (an average of approximately 41 per lesson) and a total of 314 questionnaires was completed by other student teachers (an average of more than five evaluations per lesson). The students gave the student teachers significantly lower ratings on eight of the nine categories than did the other student teachers (peers). The exception was category 2; no significant differences were found for that category. The suggestion is made that preservice teachers should not
become overconfident when studying peer evaluation from microteaching. Secondary school students will probably be significantly more critical. When teachers taught at two separate times and were evaluated by the students, there were no significant differences between the evaluations of the first and second lessons.

Farrell, Farmer and Clark (90) began with previous research they had done to identify teaching competencies used in competency-based teacher education. They had identified 74 such competencies. There were seven competencies identified which "would not likely be discarded in a longer term study". In this research a systems analysis model was developed using the competencies as a conceptual framework for a program for teacher education in science and mathematics. The researchers sought to determine the interrelationships among the competencies comprising the model and to test the theoretical links between the model and the instrument designed for testing. Data were collected on 122 student teachers over a period of five semesters. Seven factors emerging from the analysis accounted for 46 percent of the total variance. Six of those factors loaded on the same general criterion which can best be described as dealing with the affective domain. Those six factors are: nature of the content to be learned, student intellectual development, objectives specified, how humans learn various categories of content, instructional strategy design, and feedback resulting from the implementation of plans.

Many different positions in the educational hierarchy from both the college and precollege level have an impact upon the profession of secondary school science teaching. Do all persons in those positions view the tasks to be accomplished from the same frame of reference? Conner (65) compared the views of science teachers with each other and the views of science teachers with those of "methods course" instructors on selected theories and practices in science education. The questionnaire technique was used and 42 methods course instructors from 35 Ohio teacher education institutions and 216 Ohio secondary school science teachers participated. A total of 100 items, 50 on theory and 50 on practices in secondary-school science, was rated on a five-point, Likert-type scale that indicated degree of agreement with the theory statements and frequency of use of the practices. The data were analyzed by the principal components method, t-tests and the
The principal findings were: (1) Methods instructors were significantly higher than were the teachers on the theory factor related to nontraditionalism and significantly lower on the theory factor related to content mastery. (2) Methods instructors were significantly higher than were the teachers on the practices factors related to humanism in science, nontraditional approaches, and diversity of classroom procedures. (3) The factor structures on the theory portion of the questionnaire were different from the factor structures on the practices portion of the questionnaire. (4) Analysis of variance performed by experience on the teacher responses resulted in significant F ratios for all six theory factors and on only one practices factor. (5) Analysis of variance performed by science courses on the teacher responses resulted in significant F ratios for three theory factors and five practice factors. (6) Analyses of variance performed by school size and by teacher's major field resulted in no significant differences on either theory or practices factors. (7) Analysis of variance performed by amount of the teacher's education resulted in a significant F ratio for one theory factor and no practices factors. (8) Analysis of variance performed by curriculum, either traditional or one of the national programs, resulted in significant F ratios for one theory factor and three practices factors.

The results of the study indicated that the theoretical orientation of practicing teachers differed significantly from the theoretical orientation of methods course instructors. Furthermore, science teachers used practices which differed in many significant respects from the practices advocated by methods instructors. A teacher's teaching practices did not seem to have a relationship with that teacher's agreement with given theories.

How well do teacher education graduates function in the schools? Acquiring data to answer that question requires that follow-up studies be made of those completing teacher education graduates. Most assuredly not enough follow-up studies have been made in education. Science education literature for 1976 contains such a study. Swami (286) made a follow-up study of teacher education graduates from a field-based program which was designed to prepare science teachers to implement inquiry-oriented activities in their classrooms. This researcher wished to find if teachers had changed their views of inquiry-oriented activities for science teaching and how activities implemented by teachers with differing amounts of experience
might be different. Data were gathered from 86 teacher education graduates having one to four years of experience and teaching fulltime science in Ohio during the 1974-75 academic year. A wide variety of instrumentation was used. Using Analysis of Variance, Swami found no significant difference in changes of science teachers' views regarding the appropriate types of classroom activities after one to four years of teaching experience. Furthermore, no significant differences were found in the types of activities these teachers implement in the classroom. Several variables which influence the types of activities implemented in the classroom were cited. The data demonstrated that students generally liked inquiry-oriented teaching and teachers considered administrative support for inquiry-oriented activities essential. Swami concluded that an outcome of a field-based program was the stability of views toward inquiry-oriented activities of teachers completing a field-based program.

Follow-up studies can also reveal trends, fads, and/or stability in science education. Mayer (1974) reports on a study conducted in the spring, 1974, which was a follow-up of an original study conducted in 1964-65 and a follow-up of that study conducted four years later. As a part of the 1964-65 study, a criterion model of an earth science education program was developed. The 1968-69 study identified 210 institutions as having earth science teacher education programs, 38 of those institutions had provided information for the 1964-65 study. The questionnaire used in the 1974 study went through several stages and evaluation during its development. It used data and procedures from other science education activities. The questionnaire was sent to each of the 71 institutions which had provided information in the 1964-65 study; 68 of those institutions responded and all 38 from the 1968-69 study responded. Data from the study suggest that the academic components of programs of institutions involved in earth science teacher education were "resistant to change even when it means adopting nationally developed recommendations." Earth science as an interdisciplinary study was not being implemented by the institutions in the study. The procedures for presenting the subject have seen very little change. Very little field work was found; that fact was first isolated in the 1964-65 study. The requirement of an earth science teaching methods course has been added since the 1964-65 survey. The Earth Science Teacher Preparation Project has had little or no effect upon programs offered by the institutions in the study.
Research in Science Teacher Education in K-12

Since 1958 the National Science Foundation has spent nearly 150 million dollars for teacher education and upgrading. Is there any evidence that preservice student cognitive achievement has been increased because of teacher participation in NSF-sponsored institutes? Willson and Garibaldi (310) believe that the answer to that question is generally positive but the significance to students of teachers attending institutes varies by subject matter and grade level. Furthermore, that significance was influenced by factors outside the scope of the study.

Forty-four sophomore teacher education students at the University of Iowa represented the population Rizzini (230) used to study whether or not the Iowa UPSTEP preservice teacher education program made a positive contribution to the students' self-concept. Twenty-two of the 44 students served as a control group and 22 of the students participated in UPSTEP. Twelve variables concerning self-concept were measured with Tennessee Self-Concept Scale. The UPSTEP group scored significantly better on ten of those twelve variables than did the control group. Since the concept of self is believed to have an impact upon teacher-student interactions, the results are important to teacher education programs. Several recommendations are made for use of these findings. The investigator concludes that perhaps "the most important change is putting the person back into the processes of education."

May and Crawley (183) sought to develop an instrument capable of assessing the effect of the classroom teacher's model of teaching upon the model acquired by the student. An instrument was designed which contained statements grouped into one of three categories depending upon the particular skill in question. These skills were instructional, interpersonal, and managerial in nature. A Likert-type scale was used to indicate the extent to which an individual's model was teacher-, class-, or student-centered. Data were gathered from Level I interns, student teachers, and cooperating teachers in the competency based teacher education program at the University of Georgia. Pre- and post-data were obtained from Level I interns; a single sampling of responses was collected from student teachers and cooperating teachers. A correlational analysis was conducted which examined intra- and inter-group agreements. Numerous findings were made that indicated
that the instrument was sensitive enough to measure changes. Among these results was that student teachers viewed the relationship between instructional and interpersonal models of teaching as similar whereas Level I interns saw them as unrelated, and cooperating teachers were uncertain as to the relationship between the two teaching models.

Most certainly when a student enters a course which is to give him/her a perspective on teaching science, that student should expect to develop a perspective which integrates the nature of science, the nature of the learner, and the nature of teaching. College students should, in other words, develop a theory base for teaching science. Does this happen? What would such a learning experience look like? Could an analytical scheme be derived that would be useful in examining teaching materials (namely books) used in the college methods courses and which also might furnish a perspective to look at teachers' views of science and teaching? Russell (253) developed such an analytical scheme from selected theoretical perspectives of the nature of science and the concept of teaching. The divergent interpretations of science of several science philosophers were examined and used to develop five dimensions of the analytical scheme. Selected philosophical analyses of the concept of teaching were described and interpreted, yielding six more dimensions. An initial assessment of the applicability of the analytical scheme was made by using it to examine arguments in eight passages selected from a sample of textbooks which discuss methods of teaching science. As developed, the analytical scheme may be used by science teacher educators in the design and evaluation of various aspects of their programs; several possible applications were noted. The theoretical perspectives developed in the study provide a sound conceptual basis for research concerned with views of science and teaching actually held by teachers, views implied by teachers' teaching behaviors, and processes by which views or teaching behaviors actually do change.
Research in Inservice Teacher Education

A total of 24 rural elementary teachers used a home study instructional program designed by Futrell (102) and which utilized materials from the Elementary Science Study (ESS). The 24 participants analyzed the investigations that they performed from a teacher's point and then used the activities with the students in their own classroom. The study was made in order to test the effectiveness of providing teacher education with curriculum developments in an area of low population density. Specifically, the study was to test the effectiveness of the home instructional program in: (1) teaching a new elementary science curriculum's philosophical approach and methodology; (2) developing an attitude toward science and the teaching of science more consistent with the philosophy and goals of the new science curricula; and (3) developing a positive attitude toward the program of inservice instruction itself. The results of the study showed that knowledge of the ESS program was acquired and that a positive attitude toward science teaching was developed. However, significant changes were produced in the teachers' attitude toward science. Positive attitudes were developed toward home study.

Fifty-three elementary teachers in the downtown area of Panama City were pre-tested by Barnett (17) to establish their knowledge of science process skills and attitudes toward science, scientists, and scientific careers. The teachers were randomly assigned to four workshops according to their process skill scores; those workshops were of varying length and approaches. All workshops used the Science-A Process Approach curriculum, materials and equipment. Some of the groups, however, participated in microteaching. Due to the statistical technique used, data from only 40 teachers were used in the final analysis. Pre- and post-workshop observations were made of the teachers' classrooms to ascertain the use of equipment and materials, as well as overall behavior, by the children in those classrooms. Barnett concluded that workshops that are organized for teachers' active participation can facilitate the transfer of that experience to classroom situations and the students of those teachers have their process skills and use of equipment enhanced. No interaction was noted between workshop length and the way the workshop was conducted.
A study by LaShier, Hall, and Colbert (156) involved the participants of a four-week, National Science Foundation-sponsored workshop who were interested in piloting the life science phase of Science Curriculum Improvement Study (SCIS) materials in six school-districts in Northeast Kansas. The purpose of the study was to investigate the progress of participant usage of SCIS and to identify concerns of the participants so that help could be provided. For this purpose, a "Concerns Questionnaire" was administered to the participants before and after the workshop, and was also administered to teachers who had attended the workshop the previous year and had been using SCIS for one year. Among the results of the study were: (1) the teachers who had been using SCIS for one year were less concerned about the management of SCIS than were the inexperienced teachers; (2) the workshop participants were more concerned about management of SCIS after the workshop than before the workshop; and (3) there was no difference between the scores of the participants on the instrument "Science and Science Teaching Attitude Scales" before and after the workshop.

Thirty-six teachers from eight different schools and their students formed the experimental group used by Hounshell and Liggett (130) in evaluating whether or not bringing about cognitive and affective change in students could be accomplished by bringing about cognitive and affective change in their teachers. These researchers used environmental education as their vehicle in gathering data which would lead to the answer to their inservice education question. They designed a three-phase treatment of the teachers in the experimental group (a control group was also used) that began by studying particular areas and resources in those areas, through environmental awareness activities and ended with implementation in the teachers' classrooms. A fourth phase, evaluation, was also included. Knowledge of attitudes toward, and feelings about, the environment were evaluated with appropriate instruments. In other words, the evaluations were made on the cognitive and affective levels; almost 2000 tests were administered in the pre-, post-experimental/control testing cycles. The answer to every question, but one, asked of the data was "yes" and all those answers were statistically significant. They did show that the time of exposure of the teachers to the project had a significant effect on the attitude of students but no significant influence on knowledge scores.
Evaluation in Science Education

A large number of the studies reviewed in this summary could be classified as evaluation studies. Many of these studies (for example, evaluation of teacher education programs and evaluation of instructional methods) have been or will be summarized in other sections of this review. This chapter will limit itself to five aspects of evaluation:

1. Evaluation surveys - those studies ascertaining the state of learning science;
2. Curriculum evaluations;
3. Affective studies - those looking at the parameter of attitudes, beliefs, self-concepts, values, and interests;
4. Instrument development;
5. Evaluation methodology - those studies looking at the techniques and procedures of evaluation in science education.

There is naturally considerable overlap between these various categories of evaluation. A study assessing the effect of a student's attitudes toward science may also involve the development of an instrument to measure that attitude. These studies might be reviewed in either or both of the relevant sections of this chapter, depending upon how important the contribution to those areas is.

Evaluation Surveys

Quite a large number of studies appear to be trying to assess the state of science learning by students. Some of these surveys are very general in nature; others are more specific. Some survey on a large scale, some on a smaller scale. There are several foreign studies which give insights into the workings of science education in emerging countries. The most important of these surveys, however, is the National Assessment of Educational Progress (NAEP), which released several studies during the time considered by this review. The National Assessment studies (205, 206, 207, 257, 258, and 263) interpreted various aspects of the data collected during the 1969-70 and 1972-73 science assessments.

There were 400 questions used in the science assessment and they were asked of four age groups: 9 years, 13 years, 17 years, and young adults (age 26-35). Sample groups were compared by geographical region, sex, race, and type of community. The National Assessment objective for science were: (1) to assess the knowledge of scientific fact and
principles, (2) to assess process skills, (3) to measure an understanding of the investigative nature of science, and (4) to measure the appreciation of science's role in culture. One of the reports (205) included sample questions and an analysis breakdown of the question by groups.

Questions from the NAEP exam which were concerned with knowledge of and attitudes toward energy were studies in one of the reports (206). The results suggest an overall decline in knowledge, although some groups saw slight gains. This is in keeping with the general decline in science achievement noticed between the 1969 and 1973 surveys. Attitudes towards conservation and environmental concerns appear very strong as measured in the 1973 survey. Selected results from the attitude questions of the NAEP are studied in another report (207). Interest in, and attitude toward, science appears to have undergone a small overall decline between 1969 and 1973. The majority still, however, have positive attitudes towards science. Interest seems to be high but is declining. Other results from the survey indicated that few high school students (approximately 10 percent) see science as a career in their future. Almost all of the subjects see that research is important and beneficial. However, selected types of research get different levels of support from these students.

Information on selected scientific skills and knowledge were summarized in another NAEP study (257). The breakdown of results in the 9, 13, and 17 year old age groups appear to be consistent. Students did not do as well in these areas in 1973 as they did in the 1969 survey. There was an average two to three percent drop on the individual questions which were given in both surveys. Regional results indicated that the Southeastern region was the lowest in overall knowledge and skills in science, but had improved its position with respect to other regions from 1969 to 1973. Although boys appeared to out-perform girls in the area of science knowledge and skills, both showed about an equal decline during this period. Scores for blacks were below whites, but did not decline quite as much during the study period. Students in rural areas showed gains in the 9 and 13 year old age group and smaller than average declines in the 17 year old age group. This group was next to lowest with regard to other types of communities in 1969.
Further analysis of the racial and regional data was carried out in two NAEP Studies (258, 263). Overall, blacks paralleled whites in their declining scores in science from 1969 to 1973. There was no improvement in the relative performance in sciences for blacks. Black males apparently scored better than black females in the 13 and 17-year-old age groups. These differences increased with age. One noticeable bright spot was that blacks in the Southeastern part of the country, previously the lowest scoring group, showed increased scores compared to blacks in other regions. In general, declines were smaller in the Southeast for both whites and blacks. Although the scores of blacks were lower than whites in this region, they closed the gap slightly from the 1969 to 1973 surveys. Outside the Southeast, scores of blacks declined at rates even greater than scores of whites. One of these reports (263) also compared the racial composition of schools in the Southeast between 1969 and 1972. During 1969, most schools in the South were highly segregated. As a result of legal and social pressures, there was a quick move toward integration of the schools, and this region found its schools much more integrated in 1972. The rest of the nation saw relatively little change from 1969 to 1972 concerning the racial composition of its schools. It was cautiously suggested by the authors of these reports that there was a connection between the relative gain in science scores by blacks in the Southeast and the change toward more integrated schools in this region. It was suggested that blacks benefit in achievement from the integration of schools to a measurable extent, while whites are not harmed by integration.

In trying to account for the decrease in NAEP scores from 1969 to 1973, it has been suggested that the 1969 attitudes and knowledge scores might be inflated because of the Sputnik aftermath and the great emphasis on science that resulted. This argument suggests that the 1973 data are more realistic in measuring the actual state of where science attitudes and knowledge ought to be. It is encouraging when observing what is sometimes interpreted as the public's negative attitude towards science, to see these young people give fairly high attitudinal measurements towards science and science-related ideas. At the same time it is disappointing to see the small percentage of students who are considering science as a career.
In a *systemwide* study of science achievement trends, Renner and Coulter (243) tried to explain why students' achievement in science in Norman, Oklahoma, was higher than nationwide statistics based on the California Comprehensive Tests of Basic Skills. The science scores of these students were higher than norms and higher than would be predicted from IQ scores. It was suggested that a well-defined *systemwide* philosophy of education was compared with the systemwide choice of a K-8 science curriculum, and that these were compatible. This compatibility was said to account for the unexpectedly high science achievement.

One of the most disturbing continuing phenomena in science achievement from kindergarten through college science is the group differences in achievement according to sex and in all learning areas according to race. Many reasons have been proposed for these differences including; sex and race role stereotyping, bias in testing, techniques and instruments used in measurement, and cultural differences. Continuing study of these problems needs to be done, and if these individual differences hold up, effective ways of dealing with these differences in the classroom must be proposed and tested.

**Curriculum Evaluation**

This section is concerned with the study of specific curriculum projects. Most of these projects are supported by foundations in the United States, most often by the National Science Foundation. Also included, however, are some foreign curriculum projects like the Nuffield Project, and curricula inspired by NSF-supported curricula but modified in other countries. These studies look at a variety of topics, from measuring the effect of the curricula on attitudes, to concept attainment. The projects studied run the gamut from elementary through high school and take into account all the sciences including biology, chemistry, physics, and earth science. The studies reviewed are organized into three sections: (1) the study of specific curriculum projects, (2) a comparison of traditional with new curricula, and (3) implementation studies.
A group of studies looked at specific curriculum materials to ascertain the characteristics of these materials, or the learning requirements of these materials, or the effect of these materials on students. Charles (51), for example, reported the results of an evaluation of the Nuffield Combined Science Projects utilized in English schools for biology, chemistry, and physics. Using staff discussion, a staff questionnaire, a pupil questionnaire, and coded response tests, he gathered opinions and ideas from 18 teachers and 420 students about the effectiveness of the project in teaching certain skills. Consensus of the opinions indicated that the desired content was being taught, that some of the skills and concepts being taught were difficult for students, but that, in general, teachers felt that most of the materials were appropriate for the levels at which they were tried.

Ulens and Merrifield (298) studied the mental abilities needed as prerequisites for learning the content of the Conceptually Oriented Program in Elementary Science (COPES). Guilford’s "Structure of the Intellect" model was used as a theoretical base for studying the COPES unit covering mechanical energy concepts. Batteries of tests assessing chosen "Structure of the Intellect" abilities were developed. Pre- and posttests on the concepts to be taught were also developed. One hundred fifty-eight sixth grade subjects were given the "Structure of the Intellect" battery and the COPES pretest. They were then taught the COPES mechanical energy sequence and given the COPES posttest. The results showed that, when combined with the pretest score, three convergent and one divergent operations were good predictors of the students' attainment of the mechanical energy concepts being taught. The authors concluded that certain mental operations were necessary as prerequisites to understand this unit, and that these prerequisites included these three convergent and one divergent thinking operations. If students do not have these thinking operations as part of their intellectual repertoire, then their attainment of the chosen concepts will be much more difficult, if not impossible.

Most of the nationally supported curriculum projects developed through foundation support stated what they considered to be the important
goals to be obtained by students using the project's materials. Scientific literacy was one such goal. The Science Curriculum Improvement Study, (SCIS), was one project for which scientific literacy was considered an important goal. Bowyer (35) studied the effect of SCIS materials on the development of scientific literacy. She developed an instrument which was used with 521 rural sixth grade children. This instrument was based on the development of nine Piaget-type literacy tasks. Points were given for different kinds of answers depending on the level of sophistication implied. These literacy tasks were designed to be class administered and consisted of both activities and pencil-and-paper tasks. The activities were demonstrated for the students and they responded by writing down answers on a prepared answer sheet. Sixty-three percent of the children in the study had been exposed to SCIS for six years. The nine literacy tasks were: (1) the ability to recognize and describe variables, (2) interpret and criticize experiments, (3) interpret data from histograms and probability questions, (4) interpret data from experiments and make predictions, (5) determine relative position, (6) identify the source and receiver of energy, (7) predict and explain temperatures in energy transfer situations, (8) identify necessities for minimal survival, and (9) understand the concepts of solution and evaporation. Bowyer concluded that these tasks provided evidence that indicated that SCIS contributes to the overall development of scientific literacy. The specific concepts which most accounted for this evidence were Numbers 1, 5, 7, and 9 listed previously so far as these four tasks are a reasonable definition of scientific literacy, this goal of the SCIS project seems to be an outcome of using SCIS.

Research examining specific curriculum materials without comparing these materials to other curricula has special problems. Foremost among these is the development of the criteria for evaluation. Operational definitions need to be carefully developed. One also needs to look at the type of data that are being collected, and specifically tie these data to the operational definitions whether these data be opinions, performance data, or a combination of the two. It is especially important in these cases that a theoretical background of some sort be developed and tested. More research looking at the mental requirements of curriculum
materials is needed. Studies of these types are valuable on two counts. First of all they test the flexibility of the curriculum materials for different types and ages of students. Secondly, this approach is a way of testing our theoretical constructs to see how useful they are in defining real world problems.

A second type of study evaluating curricula concerns the comparison of traditional curriculum materials with the "new" curriculum materials. Most of these studies operationally defined text-oriented non-laboratory-centered and fact-oriented curricula as "traditional". The "new" science curriculum materials are by contrast inquiry-oriented and laboratory-centered. Most of the "traditional" materials are developed by publishers, according to this research, and are sometimes used as control groups. Foundation-supported curriculum projects are identified as being the non-traditional approach and are the experimental groups in a comparison.

TaFoya (288) compared the inquiry potential of a textbook approach versus the SCIS approach to teaching elementary science. Using a content analysis, it was determined that the SCIS approach was superior in developing inquiry skills in children. In another content analysis, Kent and Simpson (147) studied the content of the biological aspects of sex education as defined by SIECUS (Sex Information and Education Council of the United States) to see how these topics were covered by five biology textbooks. These included the three Biological Sciences Curriculum Study (BSCS) texts (Yellow version, Blue version, Green version). The conclusion was that there was no edge for any one of the text books and that all would need supplements to cover the SIECUS defined topics thoroughly.

Ashmead (11) compared the BSCS human sciences program to an existing text to see how these two approaches affected the achievement, attitudes, and process skills of their students. The sixth grade students in the study showed no difference on these three measures.

Davis et al. (73) compared the achievement and creativity of elementary school students using Science—A Process Approach (SAPA) versus a traditional textbook program. Students in grades one through six were given
the Metropolitan and SRA Achievement Tests in science, reading and mathematics. In addition they were given the "asking" and "just suppose" activities from the Torrence Tests of Creative Thinking, Verbal Form A (TTCT). Students who had attended the school from grades one through six were included. The number of subjects varied between 72 and 131 at the different grade levels. For the creativity measure, a random sample of 25 students was drawn from each grade level. Intelligence was used as a covariant in the statistical analysis. No differences in achievement, science, reading, or mathematics were found between the control and experimental groups. The SAFA students scored higher on the TTCT in verbal fluency and verbal flexibility. The authors' conclusions were that SAFA seemed superior in helping students develop divergent thinking processes which are important in problem solving.

Sherwood and Herron (266) compared an individualized version of Interdisciplinary Approaches to Chemistry (IAC) with a conventional high school chemistry approach to see what their relative effects on students' attitudes were. A class of high school chemistry II students was chosen as the subjects of the study. Students used a conventional college text during the first semester, were given an attitude measure, and, after the second semester using IAC materials, were given a second attitude measure. Two attitude measures were used. The first was the Student Opinion Survey in Chemistry, a 20 Likert-like item test designed by IAC. The second was the Scale to Measure Attitude Toward Any School Subject, which is a Likert-like 17-item test. A significant difference favoring the IAC approach was found using the IAC opinion survey. No significant difference was found with the other measure. The authors commented that they could not tell whether noticed differences resulted from change in approach, content instructional style, or other confounding influences.

Comparison research continues to have several problems; among these include the setting up of the traditional curriculum as a straw man. Oftentimes, by carefully choosing criteria or testing instruments, the outcome can be controlled to favor the experimental group. There is a need to very strictly identify what the comparison really is. Confounding influences abound in this kind of research. The matching of control and
experimental groups needs to be carefully done, and the matching of treatments to be compared must also be looked into very carefully, so that confounding influences do not mask the tested-for effects. Comparison studies often end up comparing apples with oranges. The temptation is to criticize comparison studies to the extent of eliminating them as legitimate curriculum evaluation methods and stick completely with the study of the characteristics of specific curricula. Comparison research, however, can be valuable if it is used to identify real differences between different curricula and thereby help school personnel make choices. However, when this research is used as a propaganda device, it loses credibility.

A third class of studies concerned with curriculum evaluation examines the implementation of curricula. Most of these studies point out the difficulties and problems with implementing educational materials as they were intended by their developers. Flores (97) studied the verbal and nonverbal behavior of teachers to determine whether these behaviors were compatible with the implied philosophy of the Intermediate Science Curriculum Study (ISCS) program. Twenty junior high school science teachers in Puerto Rico were observed; ten were from urban and ten were from rural schools. A questionnaire was submitted to the schools of the teachers involved to determine biographical information. Five classes were observed for each teacher. These classes were also taped, and an observation instrument called the Science Teacher Behavior Inventory (STVI) was used to categorize observed behaviors. It was found that the teachers behaved similarly regardless of their experience, age, sex, academic preparation, or the type of community in which they taught. Their behaviors were neither consistent with the instructional nor logical behaviors proposed by the ISCS. Although this study was done in Puerto Rico and a lack of effective inservice training was noted, these implementation problems do not seem to be much different from those in this country.

Jarvis (135) studied the implementation of Foundation-supported secondary school curricula. He studied curricula associated with biology, earth science, physics, and chemistry. He submitted a questionnaire to science teachers, college professors, science consultants, and state department personnel to survey these groups' opinions about NSF-supported
programs. Philosophy; the relationship of the program with teachers, students, and administrators; and the content and implied method of the program were the areas surveyed. The conclusions reached by the investigators were: (1) the programs showed a tendency to be directed toward science-oriented students, (2) unstructured laboratory activities resulted in a feeling of frustration by students, (3) teachers were all too often not involved in program selection, (4) teachers were asked to teach new programs without sufficient training in their presentation, and (5) programs were usually not used in a well-coordinated K-12 program, but instead were grouped together at random.

Implementation problems seem similar in foreign countries as in this country. Implementation problems continue to be the major roadblock to the success of these curricula. Inservice training is necessary for this effort. The kind of inservice training which will result in successful implementation has yet to be determined.

Affective Studies

The trend of the last number of years toward studying affective parameters such as attitudes, beliefs, self-concepts, values, and interests continues to influence educational research in general and science education research in particular. Most of this research is directed toward attitudes concerning either science as a discipline, science as a school subject; scientists, or science instructional procedures. The subjects of these attitude surveys are usually science students or their teachers. Many of these studies also try to relate attitudes to grades, aptitude, ability, personality factors. Many of these studies discuss the development of instruments used to measure these affective parameters. Some of these instruments are of general interest and use and will be discussed in a later subsection of this chapter.

Schrier (262) attempted to identify the areas of science in which most elementary school boys and girls were interested. Using a forced-choice questionnaire, he polled 2200 elementary school children in grades one through six. The questions in the questionnaire were categorized into
nine areas of science. The areas of interest from greatest to least were: (1) ecology and conservation, (2) chemistry, (3) health and the body, (4) biology, (5) weather, (6) physics, (7) geology, rocks and minerals, (8) astronomy; space and air travel, (9) fossils and ancient life. In general, questions from the physical science area ranked higher than the life science area which in turn ranked higher than the earth science areas. Primary grades (1-3) ranked ecology and conservation as their areas of greatest interest. Children in the upper grades (4-6) most often selected chemistry questions in preference over other science areas. The author of this study pointed out that the greatest concerns of children lie in areas where maximum direct experience is possible.

Three of the affective studies reviewed were concerned with self-concept (175, 274, and 301). Malcolm (175) tried to determine if the use of SCIS instructional materials could help children to maintain or develop a positive self-concept as measured by the Piers-Harris Children's Self-Concept Scale. He concluded that students using SCIS science materials had higher self-concepts in the areas of intellect and school status than did students not using SCIS materials. Sellers (264) examined self-concept in science by studying 320 sophomore students at a central public high school located in an urban community. This researcher investigated the relationship between students' self-concept in science and their mental abilities, sex and achievement in science. The researcher developed a test for self-concept called the Self-Concept in Science Scale (SCSS) and validated it with two other established self-concept instruments. Mental ability was measured using the Otis-Lennon Mental Abilities Test. Science achievement was measured using the Nelson Biology Test and the Test of Science Processes. The finding of this research showed that high achievers in processes, biology, knowledge and high mental abilities also had a high self-concept in science.

Vargo (259 and 301) studied the relationship between science attitudes, self-concept, and science teacher/pupil compatibility. Using 12 classrooms which consisted of 204 students and six science teachers, this researcher pre- and posttested students on the Science Attitude Scale (SAS) and the Self-Concept in Science Semantic Differential (SCSSD). Students also completed the Fundamental Interpersonal Relations
Orientation-Behavior (FIRO-B) inventory. Sex and the student's final grade in eighth grade science were used as predictor variables along with six compatibility variables from the FIRO-B in a stepwise multiple regression design. The results of the study include: (1) Pupil/science teacher interpersonal compatibility did not account for variation in self-concept in science nor attitudes toward science; (2) both science grades in previous courses and the self-concept pretest were significant predictors of self-concept in science; (3) the Science Attitude Scale pretest was a good predictor of the attitude toward science of these students but was the only one; and (4) boys possessed a more positive attitude and self-concept toward science than did girls.

Santiesteban (256) assessed the attitude of high school students toward various science instructional processes and procedures. Three hundred twenty-two tenth, eleventh and twelfth graders enrolled in various science courses in four high schools were given a test which measured their attitude toward several variables in science instruction. These included: the structure and function of laboratory, teacher questioning behaviors, textbooks, library reports and independent projects, testing and grading, and types of instruction. The items in the questionnaire were factor analyzed and differences in attitudes by sex were studied. It was found that females considered the use of projects and oral reports more important than did males, whereas males preferred small groups and performance on a particular task.

By far the largest number of affective studies concern attitudes toward science. In a survey of primary school teachers in Southwest England, Harvey (117) investigated the amount and nature of science taught, the teacher's awareness of modern curriculum developments in primary school sciences, the effect of sex differences, and the effect of science training on the attitudes of primary school teachers. He found that male teachers were more likely to teach more science, were more aware of modern curriculum developments, and were more interested in primary school science than were female teachers. He also found that the length of
Training in science did not affect the amount of science taught but did seem to affect the quality of the science taught.

Symington and Fensham (287) investigated teacher attitudes toward science with respect to dogmatism. They studied 72 teachers of fifth and sixth grade science in Australia to determine how dogmatism, attitude toward science in congruence with new curriculum influenced the adoption of "new" science programs. They used Schwirian's Science Support Scale (Tri-S) to test these teachers' attitudes toward science. They found that teachers who felt compatible with the new curriculum materials and teachers who had a good attitude toward science were measured low in dogmatism.

Buckley (40) studied teacher and pupil attitudes toward science as a function of whether or not a science specialist worked in the district. The study involved 96 teachers and 2277 students from 4 different towns. Two of the towns had a kindergarten through sixth grade science specialist, while the other two towns did not. Teacher attitudes toward science were determined by using Moore and Sutman's Science Attitude Inventory (SAI). Student attitudes toward science were measured by the semantic differential Science and Me test. Student science achievement was measured by the Stanford Achievement Test at the primary levels and Science Research Associates' Science Achievement Test at the intermediate levels. Results of the study showed that the teacher sample of the specialist towns had significantly more positive attitudes toward science than did the comparative sample from the non-specialist towns. Also, the total student sample from the specialist towns had significantly more positive attitudes toward science than did the comparative sample from the non-specialist towns. No significant differences were found between the groups concerning science achievement.

The remaining studies summarized here are concerned with students' attitudes toward science. Ward (304) studied high school biology, chemistry, and physics students in 12 states to assess their attitude toward science as a function of class size. Attitude toward science was measured by Moore's Science Attitude Inventory (SAI). In addition to this, achievement
tests were given to these students. The results showed that there was no evidence of a direct relationship between small class size and good attitudes toward science. There did, however, seem to be strong associations both between achievement and attitude and between achievement and class size. The author of this study suggests that achievement serves an intermediary role between class size and attitude. He hypothesizes that class size affects achievement and that achievement then affects attitude.

In two studies, Novick and Duvdevani (213 and 214) studied tenth grade students in Israel to assess their attitude toward science as related to school and student variables. The attitudes of students at agricultural schools were less positive than were those of students at either academic or vocational schools. The type of curriculum ("new" science or traditional) had no effect on attitudes. Students from Western extraction cultural backgrounds had more positive attitudes than did those students of Eastern extraction. These researchers used Moore's Science Attitude Inventory (SAI) and also did a cross cultural comparison to identify the relationship between Israeli students and those in the United States. Scores of a stratified sample of 684 tenth grade students in Israeli schools were compared with scores generated by a similar study done in the United States. The results of this study show remarkable similar attitude patterns between the two cultures.

In another foreign study, this one done in Australia, Gardner (104) studied physics students' attitudes toward physics as a function of pupil and teacher personality characteristics. One thousand-fourteen high school students using PSSC materials were studied using three instruments developed in previous studies by the researcher. The Physics Attitude Index (PAI) assessed students' views toward physics learning on four categories: (1) authoritarian/non-authoritarian, (2) physics as an open/closed process, (3) views of scientists as normal/eccentric people, (4) physics as enjoyable/not enjoyable. This test was given both as a pre- and posttest for the study. In addition to this instrument, the Personal Preference Index (PPI) measured personality characteristics. Several scales were developed from this instrument including: achievement,
conjunctivity, deference, play, understanding, order, nurturance, and energy. A third instrument, The Physics Classroom Index (PCI), provided scales concerning student needs as they occur in a classroom. These scales included competitiveness, organization, compliance, pleasure, intellectualization, compulsiveness, warmth, and stimulation. The author found that students high on nurturance had small but significantly more favorable attitudes toward scientists. Students high on achievement and understanding were more favorable toward non-authoritarian modes and enjoyed the subject more. Students who described the teacher as well organized and intellectually stimulating also enjoyed physics more. Achievement-motivated, intellectual students tended to hold more open views of physics. However, achievement-motivated teachers promoted a more closed view. Pupils who were warm and friendly, and also those who were submissive and conformist, were more likely to regard scientists as normal people than as eccentric people. Overall, this study showed a sharp decline in enjoyment of physics by students who took this course using PSSC materials.

An overall summary of the attitude data collected by the preceding studies can be summed up in several statements. Males have better attitudes toward science than do females. Higher achieving students have better attitudes toward science than do lower achieving students. These results are not particularly surprising. The first result has been part of the popular belief for quite a long time. Perhaps this will change in time and be reported differently by research summaries of the future. The second result is a reflection of one of the weaknesses of correlation type studies, a weakness which might be called the chicken and egg syndrome. It is difficult to tell whether attitudes cause achievement or vice versa. It is difficult to come to any sort of conclusion about this from the studies that are reported here. One final result of summarizing these attitude studies is that Moore's Science Attitude Inventory seems to be the most widely used instrument for assessing attitudes towards science.
Instrument Development

Many of the evaluation researches discussed in this chapter involved the development of instruments in order to collect pertinent data necessary for the proposed research. Many of these are specific to the research being carried out or are too limited to be of general interest and application. Some, however, show promise of value outside of the research setting in which they were developed. Of these, some are interesting only because of the techniques or procedures used in development. These will be discussed in the next sub-section. The following is a brief description of new evaluation instruments which might have some general use in science education research.

Since most of the new foundation-supported curricula developed in the last decade involve teachers in using inquiry teaching techniques, there is a great deal of concern about secondary science teachers being able to exhibit appropriate behaviors. Lazarowitz and Lee (163) have developed an instrument for determining the inquiry attitudes of secondary science teachers called The Inquiry Science Teaching Strategies Instrument (ISTS). The instrument consists of forty Likert-like items which are positively and negatively related to the inquiry approach. Sample of items positively related to the inquiry approach are: "students are often capable of designing valid experiments", and "it is desirable to present to students science questions to which answers are not necessarily known." Examples of items negatively related to the inquiry approach are: "questions which are integrated in the text are confusing to students and should be omitted," and "a primary role of secondary science teachers is to design the investigations to be done". Validity was judged by a panel of expert judges. An item analysis lent further support to the instrument. The alpha-coefficient reliability of internal consistency averaged between 0.48 and 0.85.

Another instrument used to measure attitudes, this time of eleventh-grade chemistry and physics students, is the Test On Scientific Attitudes (TOSA). The developers of this instrument, Kozlow and Nay (152), criticized what they identified as four shortcomings of other attitude measures.
These researchers say that they: (1) are too general, (2) lump several dimensions under the caption of attitudes, (3) show no discrimination between cognitive and affective components of attitude, and (4) do not adequately represent classroom situations and experiences. Taking the point of view that attitudes must be inferred from the behavior of students, these researchers developed a multiple choice format test with the stem of each multiple choice question describing a situation relative to an attitude. The distractors of the questions defined different courses of action that a student might take which represent different scientific attitudes. In developing the examination, the researchers developed behavioral definitions of eight attitudes: (1) critical-mindedness, (2) suspended judgement, (3) respect for evidence, (4) honesty, (5) objectivity, (6) willingness to change opinions, (7) open-mindedness, and (8) questioning attitude. The behavioral definitions of these eight attitudes were used to develop items. Forty items survived to form the final test.

Twenty of these items make up the "cognitive component" subtest of the total examination and describe a situation which a scientist might encounter during his work. The student is then asked to select from four courses of action the one which is most appropriate for the scientist. Twenty items make up the "intent component" subtest which presents a situation which the student may encounter in the science classroom or in everyday activities. The student again is asked to select one from four alternative courses of action which represent his reaction to the situation. Test-retest reliability judged by the KR-20 was 0.71. Content validity was determined by a panel of judges and structural validity by factor analysis.

Molitor and George (196) have developed a science process skills examination to assess the inquiry skills of inference and verification of fourth, fifth and sixth graders. In an attempt to be content free, the items were based on common, everyday experiences (for example, a window breaking). Items and item distractors were in the form of illustration and test instructions were read to the students. The test was administered in a group format. Items are in a multiple choice format. There are nine items for each skill. Validity was determined by a panel of judges. KR-20 reliability average for the three grades for the inference test was only 0.56. Verification reliability was high at 0.75.
More emphasis is being put on the role of science instruction in helping students develop self-concepts; both generally and in science. Sellers (264) has developed the Self-Concept in Science Scale (SCSS). Items for the scale are Likert-like items divided into two categories, the "operations of learning in a science classroom", and a general self-concept scale. The "operations of learning in a science classroom" subscale consists of questions concerning science processes such as observing, comparing, classifying, etc., and methods and techniques of learning, such as taking notes, testing, reading and others. The general self-concept scale consists of three subparts, identity, self-satisfaction and behavior. Content validity was determined by nine judges. Total agreement resulted in keeping 63 statements in the Likert-like five-dimension scale. Students were to use the item statements as self descriptions. Validity with an existing self-concept scale, the Tennessee Self-Concept Scale, were 0.43 for the identity section, 0.44 for self-satisfaction, 0.42 for behavior, and 0.48 for total composite. Test-retest reliability with 142 tenth grade biology students as subjects was measured at 0.82. The test can be completed in 15 minutes.

Anderson and Herrera (8) attempted to translate an existing attitude scale into Spanish. According to these authors, the problem in transferring from one language to another is not just a translation problem, but also involves the evaluation of culture-bound items. They developed a Spanish version of the Allen Inventory of Attitudes Toward Science and Scientific Careers, a 95 item Likert-like scale. The Spanish version consists of 38 Likert-like items. It was used with college-age students. Its reliability coefficient alpha was 0.89 and 0.80 in two different uses. The authors argued that a battery of such Spanish and other foreign language instruments is needed.

Science education needs well standardized systems for all phases of evaluation in science education. These "tools" are necessary for the progress of research in this field. The continued proliferation of instruments is a necessary evil until a battery of well developed, reliable, and valid instruments can be developed. A system for categorizing and storing these instruments for retrieval by researchers in the field is needed. Some sort of critical evaluation of existing instruments needs to be made and continually updated.
Evaluation Methodology

This discussion is concerned with the techniques and procedures of evaluation presently being used to do science education research. Some of these studies are concerned with formative evaluation techniques. Ciesla (55), for example, tried to determine what types of feedback had the most influence on the revisions of instructional materials developed by writers in the Individualized Science Instruction System (ISIS) project. The most influential types of feedback which resulted in revision were verbal or audiotaped feedback from teachers. Feedback that involved information obtained from students who used the instructional materials received the lowest rating. The author concluded that informal and verbal, that is non-quantitative, feedback was rated as being more influential than was feedback that was more formal, less verbal, and more quantitative.

Interaction analysis instruments also continue to be commonly used in evaluation research. Platts (231) suggests a new technique of using time-lapse photography with a moving picture camera to study classroom movements utilizing non-verbal interaction analysis instruments. Statistically, multi-variate statistical analysis seems to be used more and more commonly in science education research.

Three techniques or procedures, however, seem to be much more commonly used than in the past, and therefore require discussion. These are the Q-sort, ethnomethodological research, and content analysis.

The Q-sort is a fairly old technique, or method, of rank-ordering a large number of items into categories. It is used as an evaluation instrument, usually for some sort of affective study (attitudes, values, opinions). Typically, subjects are asked to sort items into categories according to specific criteria. These items are sometimes printed on separate cards, and students are asked to sort the cards into a forced normal distribution. Analysis can be carried out by a large number of techniques including what is called a Q-technique, a version of factor analysis. Q-sort and Q-sort-type instruments are again becoming popular. Toews (294) used a version of Q-sort called a free-sort to measure
Ethnographic evaluation methodologies are also gaining popularity and greater use in science education. These types of studies, which consist mostly of case studies or what might be called anthropological field studies, can be very useful in studying curriculum implementation and school procedures. Harding (115) studied the use of the case study method of inquiry to assess the implementation and evaluation of Nuffield curriculum materials in Great Britain. Questionnaires, structured and unstructured interview techniques were used. The case study method, Harding thinks, is useful for generating hypotheses and, because it is more open than techniques like questionnaires, is not as limited. Insight can be gained when combinations of these instruments and techniques are used. Harding used the case study techniques to study the communication and support for change for school science education. Harding concluded, among other things, that change which is initiated from outside is most successful: when it builds on or creates dissatisfaction with the present situation, when it builds on an acceptability for change, when it is adaptive to a local situation, and when resources insure the feasibility of the change. She further concluded that there are problems if the change requires the teacher to assume a changed identity or if the change is in conflict with vocational aims of science education. Teacher attitudes, an exposure to communication, openness to change, initiative, independence, and leadership are all factors affecting the success of curriculum implementation.

School procedures affecting decision-making and participants most important to the decision-making process in schools were studied by Werner (306). Using case study techniques such as review of records of important meetings, interviews with staff members, and other formal and informal techniques and instruments, the came to the conclusion that the most important internal participants in the decision-making process were, in order of importance, elementary school science consultants, superintendent of schools, elementary school principals, the director of elementary instruction, the assistant superintendent of schools in charge of instruction, classroom teachers, and lastly, students.

Davidman (72) carried on a rather extensive anthropological field study to evaluate the effectiveness of a curriculum project's training.
Ethnographic evaluation methodologies are also gaining popularity and greater use in science education. These types of studies, which consist mostly of case studies or what might be called anthropological field studies, can be very useful in studying curriculum implementation and school procedures. Harding (115) studied the use of the case study method of inquiry to assess the implementation and evaluation of Nuffield curriculum materials in Great Britain. Questionnaires, structured and unstructured interview techniques were used. The case study method, Harding thinks, is useful for generating hypotheses and, because it is more open than techniques like questionnaires, is not as limited. Insight can be gained when combinations of these instruments and techniques are used. Harding used the case study techniques to study the communication and support for change for school science education. Harding concluded, among other things, that change which is initiated from outside is most successful: when it builds on or creates dissatisfaction with the present situation, when it builds on an acceptability for change, when it is adaptive to a local situation, and when resources insure the feasibility of the change. She further concluded that there are problems if the change requires the teacher to assume a changed identity or if the change is in conflict with vocational aims of science education. Teacher attitudes, an exposure to communication, openness to change, initiative, independence, and leadership are all factors affecting the success of curriculum implementation.

School procedures affecting decision-making and participants most important to the decision-making process in schools were studied by Werner (306). Using case study techniques such as review of records of important meetings, interviews with staff members, and other formal and informal techniques and instruments, the came to the conclusion that the most important internal participants in the decision-making process were, in order of importance, elementary school science consultants, superintendent of schools, elementary school principals, the director of elementary instruction, the assistant superintendent of schools in charge of instruction, classroom teachers, and lastly, students.

Davidman (72) carried on a rather extensive anthropological field study to evaluate the effectiveness of a curriculum project's training.
model for teachers. Showing how the problems to be studied and the techniques for studying them evolve as a result of a personal interaction and involvement with the curriculum project was the object of this research. While studying the curriculum materials, he developed doubts about the project's implementation and diffusion strategies. Using hybrid methodologies, investigative strategies closer to that of ethnography and the conceptualization of art criticism than to standardized research methods, the design of this research study evolved. The techniques used varied from traditional to more innovative and informal approaches. Tape recordings of lessons, super-8 filmings, structured interviews, written observation notes of class events, a diary, attitude inventories, questionnaires, an analysis of curriculum documents, and oral dialogues with key personnel were techniques which became part of the study procedures. The unique role of the participant observer and the process of getting involved in the curriculum project, gives the researcher a special vantage point.

There are several differences between anthropological research and more traditional research: inquiry is naturalistic rather than dependent on experimental techniques, research is not necessarily started or ended with hypotheses, methods of study evolve as the researcher proceeds rather than being planned out in advance, there is more emphasis on intuitive observation and eclectic and hybridized methods, there is more personal involvement of the researcher in the project, and the approach is more informal than in more traditional methods.

Even though ethnographic research techniques can be criticized for being non-objective, it is obvious from studying these reports that much valuable information can be developed from techniques of this sort that would not come out of more traditional research strategies. In the hands of a competent and self-aware researcher these techniques can be very valuable for assessing the impact of curriculum materials.

Content analysis is another research methodology which seems to be gaining influence in assessment and evaluation in science education. The key to this type of research seems to be the development of objective checklists for observation instruments to apply to the content of the...
curriculum materials. Kent and Simpson (147), for example, examined five biology textbooks to analyze their coverage of six related topics as defined by SIECUS, using index and page counting techniques. They compared the coverage of these textbooks with topics suggested in their criteria.

A much more extensive content analysis was performed by TaFoya (288) who developed an inquiry potential analysis instrument to do content analysis comparisons between a textbook versus the SCIS elementary science curricula materials. An operational definition of inquiry potential was used to develop a set of standards and a category system was developed. The category system was established as the first process in the analysis. Randomly selected sentences of the two curricula examples were categorized and knowledge claims which were imbedded in the programs were identified and quantified. The second process of analysis involved examining the knowledge claims and the manner in which they were to be verified by students. By examining the sentences from the teachers' guides as units, it was found that 48 percent of the sample of the textbook program contained either pseudo-scientific assertions, which were confusing or non-verifiable, or theoretical knowledge claims. Thirteen percent of the statements were found to be experimental knowledge claims that should require children to observe some natural phenomenon for their verification. Yet examination of procedures provided for verification revealed only confirmation activities in which results were known beforehand. In contrast, the SCIS program contained no pseudo-scientific statements, only 4 percent of the sample was found to consist of theoretical knowledge claims, and 47 which provide information only about how we have agreed to use words, or experimental knowledge claims. Examination determined that experimental claims in this program required empirical verification by students, principally through guided inquiry, active investigation, and collection and analysis of data with no prior knowledge of expected results.

There seems to be a high potential for formal and informal evaluation methodologies to be used in science education. All of those described, and many others which are presently being used, should become part of the repertoire of techniques which are used by researchers in appropriate contexts.
The Use of Media in Science Education

The use of media in instruction has concerned educational researchers for a long time. This review of research into the use of media in science education is concerned with the use and effectiveness of certain hardware such as television, movies, and computers; and the study of strategies of instructional presentation, such as program instruction, audio-tutorial instruction, and simulation. Although science as a discipline has made common use of many of the traditional instructional media, such as film, for the most part it continues to depend on the laboratory as the major source of variety in instructional presentation. Recently, however, simulation games, audio-visual strategies, and combinations of media have interested science education researchers. To some extent, media have been used as substitutes for direct experience. There are three possible reasons for this substitution. First, the direct experience could not practically be provided in the school setting, therefore some simulation or representation of the real experience had to be provided. Secondly, the use of simulation was more efficient. Usually, this meant simulation had a cost benefit, a time benefit, or a manpower benefit which made it more reasonable to use under the constraints of a classroom situation. Finally, simulation was considered to have an educational efficiency. It was possible to do a better job using media than it was without its use. This sometimes meant that the real experience was too complicated to be understood.

The study of astronomy is one area where simulation has recently played a major role in instructional strategies. Because of the difficulty of obtaining data and information from the heavens directly, and because of the inconvenience of studying the heavens during daylight hours, simulation has played a heavy role in the study of astronomy concepts. Lang (154) compared students using computer graphic simulations with those who used the multi-media material developed by the Project Physics Course during the study of the Project Physics unit "Motion in the Heavens". Thirty students were randomly selected to serve as subjects. The simulations presented sections from the text as well as laboratory experiments. The control group attended discussions and did laboratory experiments from the Project Physics handbook. Students were tested at the conclusion of the units by...
using a Project Physics test and responded to an attitude questionnaire about the computer. Ninety days later both groups were retested on the material in the unit. The findings of this study indicated that there was no significant difference between the experimental and control group gain scores on the achievement test. Using IQ as a covariant, the experimental group scored significantly higher than did the control group on the retest given ninety days after the conclusion of the unit. The experimental group seemed to have a higher positive attitude toward the computer than did the control group.

Several research studies looked at the role of the planetarium as a simulation device in teaching astronomy. For example, Smith (272) studied three methods of teaching constellation recognition to differing age groups. The three treatments were: (1) teaching constellations in the classroom by using 35 mm slides of hand drawn constellation star fields; (2) teaching constellations in the planetarium by means of 35 mm slides of hand-drawn constellation star fields; and (3) teaching constellations by means of a planetarium sky. The three age categories were children, teenagers, and adults. One hundred-three subjects completed the study. The results of the study indicated that: (1) regardless of treatment, all age groups performed approximately the same when evaluated under the real sky; (2) when evaluated by a paper and pencil instrument, those subjects who were instructed under the planetarium sky scored significantly lower than those receiving treatments involving slides; and (3) the subjects receiving the slide treatment in the planetarium reported more positive responses to the study of constellations than did the subjects receiving the other treatments.

Etheridge (88) also studied simulation by comparing planetarium instruction versus two dimensional slides representing the constellation sphere. The experiment was a posttest-only control group design. The experimental group received instruction using a planetarium. The control group was treated identically in location and oral presentation via a recording but substituted two dimensional representation of the celestial sphere projected by slides for the visual component. Four elementary astronomy classes from two California community colleges participated in the experiment.
Each class received each treatment for each of two topics; the sun and the seasons, and the moon and its motions. Each treatment period lasted 30 minutes and was followed by a 15-minute posttest. The posttest consisted of three types of multiple choice items. One set had a visual component similar to the simulation treatment, and the other had a visual component similar to the representation treatment. The third set of items was entirely verbal and had no visual component. All subjects were additionally administered a verbal aptitude and a spatial aptitude test. The results of the study indicated a definite positive correlation between performance on the posttest and spatial ability, regardless of all other variables. No such relationship was noticed with the verbal aptitude scores. Post facto analysis indicated that this relationship was colinear. The posttest versus spatial aptitude score regression slope was significantly greater for subjects scoring above the spatial ability grand mean than for those scoring below. Etheridge made several recommendations based on the study, including: (1) that spatial aptitude tests would be useful in identifying students who have a high probability of experiencing difficulty with highly visual materials in elementary astronomy, and (2) any decisions regarding the feasibility of installing a planetarium in an institution should be based on considerations far broader than just cognitive factors. Effective areas such as motivation and satisfaction are important, though little studied at this time.

It seems that even in places where substitution for the real thing seems necessary, that there is little evidence that these simulations are effective on a cognitive level. It might be that some topics, such as those concepts found in astronomy, are so abstract that those students who can deal with abstractions easily do not have a particular cognitive preference for simulations such as those provided by the planetarium. While on the other hand, students who are less able to handle abstractions find learning the concepts taught through the use of a planetarium just as difficult as they would by viewing the sky. Even though they are dealing with the real world, their access to that world requires abstract thinking abilities they do not possess.
Several studies explored the use of media as an economic factor. Piper and Butt (228) studied the effect of a televised inservice program for elementary science teachers. They explored three questions. First of all, can a TV inservice program adequately prepare teachers to teach science? That is, can they give the competencies of the lessons to be taught? Secondly, can a TV inservice program help inservice teachers gain a more positive attitude toward teaching science? Finally, will there be an increase in the science activities taught in the classroom as the result of a TV inservice program? Seventy-six elementary science teachers who chose to participate in implementing Science—A Process Approach (S-APA) were the subjects of the study. Fifteen meetings lasting 30 to 40 minutes were held once a week and the teachers were instructed with a television component. Materials were used with a teaching assistant at the local level. The results of a science competency measure indicated that the subjects did, for the most part, gain the competencies of the lesson to be taught. A pretest-posttest semantic differential attitude instrument indicated that attitudes were more positive after the TV programs than before. At the beginning of the study there were no science lessons being taught. This increased to a steady average of two activities per week per teacher after just a few sessions. This trend continued for the period of training. Piper and Butt concluded that science inservice programs via television is a reasonable method to efficiently make use of time, place, and resources for schools with access to television.

These researchers also indicated that the study was limited to the success of initial implementation. A study of the long range effect of materials such as these is also necessary. Will these teachers continue to teach two activities per teacher per week after the training period is completed? What factors are necessary to keep the level of enthusiasm engendered by this program? Several other questions are raised by this research. How does the TV inservice program compare with other methods of implementing the same material? Finally, would this method work with less highly structured elementary science programs?

Ben-Zvi and others (24, 25) studied the effect of experimentation in high school chemistry versus the use of films to teach the same material.
covered in the experiment. Two groups of tenth grade Israeli students were instructed in chemistry using either films or laboratory experiments. Effectiveness was measured using four measures. An achievement test in chemistry consisting of 25 multiple choice items was used to measure general achievement. Specific knowledge was judged by using a written test measuring experimental techniques and the underlying knowledge for the chemical experiments. Two practical tests, one assessing manipulative skills and the other assessing planning of experimental procedures, were also carried out. And finally, an observation test was carried out to assess skills performed in the laboratory. Three hundred thirty tenth grade students; 150 in the film group and 180 in the experimental group, received a five-month treatment consisting of 11 key experiments. The experiments, either taken directly or vicariously, depending on which group the students were in, consisted mostly of two activities: (1) the collection of qualitative or quantitative data and other observations, and (2) the interpretation and evaluation of the collected data. The results indicated that the experiments were better only at helping students gain routine manipulative skills. A questionnaire consisting of 49 Likert-type items was also administered to these groups, which indicated that both groups found personalized laboratory more effective than films in promoting their interest in chemistry.

Results of the preceding study indicated that laboratory activities presently being used in chemistry are no better than films from a cognitive point of view. Obviously, these laboratories are not designed either to teach students to learn to design and carry out experiments or to teach specific knowledge. From a cognitive point of view, we are wasting our time and money with the presently available laboratory materials. They seem to be good only for helping students gain routine manipulative skills. Several possible recommendations might be based on this kind of research: (1) One might say that we are wasting our time and money on laboratory material, and we should seek other ways of helping our students learn manipulative skills; (2) One might take the approach that the motivational factor by itself justifies the expense and time which our students spend doing laboratory work; and (3) Presently available materials are not worthwhile and we should spend time and energy in developing different types of laboratory materials which in fact meet the objectives which we feel are worthwhile.
One of the techniques of using instructional media as a presentation device for teaching a science is the audio-tutorial approach. Nussbaum and Novak (215) studied the effect of audio-tutorial lessons to teach second grade elementary school students. Twenty-six students were given six audio-tutorial lessons on earth concepts, such as "the earth is round," "the globe can be used as a model for the earth," and "gravity." Each of these audio cassette tape lessons, used in conjunction with manipulative materials, lasted between 15 and 20 minutes. The audio-tutorial cassette guided students through experiences with objects and materials. Piaget-type interviews designed to assess understanding of earth concepts were used to judge the effect of these lessons. Results of this study indicated that there were no significant advances in the students' concept of the earth as a result of using these audio-tutorial lessons. According to these researchers, these results stand in contrast to other experiments done on audio-tutorial materials at Cornell. They add that this may be the result of the necessity of using models and other abstractions rather than direct experience in the instructional strategy.

In another study of the audio-tutorial materials with elementary school students, Dech (74) compared the effect of immediate reinforcement in feedback on the cognitive gain of elementary students using a commercial audio-tutorial metric unit. The 50 students in the experimental group consisted of 22 third graders and 28 fifth graders. The control group consisted of 23 third graders and 27 fifth graders. Both the experimental and control groups were given metric instruction using the same audio-tutorial material. Audiotapes given to the students in the experimental group contained immediate reinforcement as feedback, while those given to students in the control group did not. Both groups were pretested and posttested to determine cognitive gain. The results of this study indicated that immediate reinforcement and feedback, as presented in the study using audio-tutorial mode, was not a significant method for increasing elementary students' cognitive understanding of the metric system.

Rastovac (240) studied the effectiveness of a mastery learning strategy in promoting cognitive growth of secondary students. Three high school biology classes from both innercity and rural backgrounds were used for this
study. Piaget-type tasks were used as pre- and posttests in a written summative examination based on the objectives of four audio-tutorial units. Classes at both schools received three treatments: (1) audio-tutorial approach without mastery learning strategy, (2) audio-tutorial approach with mastery learning strategy, and (3) teacher directed inquiry with mastery learning strategy. Four individualized audio-tutorial units were utilized. Each unit presented behavioral objectives; study guides, additional experiments, and a final summative evaluation. The results showed no significant differences in posttest total scores due to the treatment.

Direct concrete experience is, to a large extent, possible in science to a greater extent than in other disciplines. In some cases, astronomy for example, it is difficult or impossible to use direct experience in teaching and learning this material. If the learners exposed to this material are sophisticated enough to deal with the abstraction implied by simulation, it seems reasonable to use simulations or other media in substitution for the direct experience. However, in the case of students who cannot deal with abstractions, there is a serious question whether this does any good. The question arises, "Should we limit the elementary curricula to concepts which can be represented as direct experience?" The answer would seem to be "Yes." Enough concepts seem to be available that can be presented at a concrete level to provide a reasonable science curriculum. The more abstract concepts which require less direct methods of approach, should be saved for those learners who can deal with them. Although approaches such as audio-tutorial systems of instruction can be justified on the basis of their cost, time, and manpower efficiency, it is still not clear whether these methods of instruction are more effective from attitudinal or cognitive standpoint than are more traditional approaches.
The Concepts, Processes, and Content of Science

The research summarized in this chapter will consist of studies of the concepts, processes, and content associated with science. Most of these studies are descriptive in nature. Many of these studies look at specific units of content, concepts and processes. Others concern themselves with syllabi of content in a particular area. A few of these studies are experimental. These experimental studies are concerned with the optimum conditions under which specific processes or concepts can be learned. Several of these studies looked at aptitude/treatment interaction effect. Student characteristics and instructional methods are the parameters most often studied.

Syllabi

Most of the studies summarized here are surveys, attempts on the part of the researcher to sample the content, processes, and concepts which are being used in elementary and high school science courses in various locations.

Fournier (100) tried to determine if cultural differences existed between Mexican- and Anglo-American students in how they perceive natural phenomena. In order to do this, he used the Science Concepts Instrument (SCI) to survey the science knowledge of 298 fifth grade students, some of whom were Mexican-American and others who were Anglo-American. The available fifth grade text and other curriculum materials were used to select the science concepts tested in the instrument. The following conclusions are based on interpretation of the survey data. Scores on the SCI correlated significantly with the father’s educational level. A significant correlation also existed between the scores on the SCI and sex within the Mexican-American group, but not within the Anglo-American group. Scores on the SCI correlated significantly with school science grades for the Mexican-American group only. Correlations between the Science Concept Instrument and cultural background, language background, age, and father’s occupational level revealed no significant differences between Mexican-American and Anglo-American fifth grade students.
Batten (20) studied ninth grade earth science students and their ability to use the processes of science. He tried to correlate this ability with the student achievement level, the students' science curriculum experience, and aspects of their teacher's educational instructional experience. The Test of Science Processes was administered to each student in 40 selected sections of earth science. The analysis of the data indicated the following. Results indicated that SCAT quantitative test scores were significant student achievement predictors for the use of science processes as measured by the Test of Science Processes. Previous science experience was significantly related to the student's ability to use science processes. Students who had completed an Introductory Physical Science (IPS) course showed a higher ability to use the processes of science than did their counterparts enrolled in eighth grade physical science of an Intermediate Science Curriculum Study (TSCS) course. Educational experience, age, sex, hours of mathematics, area of certification, years of NSF academic year institutes and number of mathematics and science center workshops attended were teacher variables which were identified as significant predictors of the use of science processes by students. Considering the instructional experience of the teachers in the sample, Batten concluded that the fewer the number of years of teaching experience, the higher the ability of the teacher's students in using science processes. This result may be accounted for by the fact that younger teachers took their college training at a time when process instruction was emphasized heavily.

Johnstone and Mughol (139) studied secondary and college level students to find out what fundamental physics concepts were the most difficult for them to understand. Concepts reported as most difficult included differences between mass and weight, magnification, indirection, the idea of fields, and potential difference. Abel (1) analyzed questions used in the introductory high school biology correspondence instruction syllabi issued by the National University Extension Association. The major findings of this study were that zoology was emphasized over other content areas. The methodology and history of science were least emphasized. Ecology received very little emphasis in the test. Test items emphasizing low level intellectual skills were over-represented. Abel concluded from these results that there was an overemphasis of biology content in the high school correspondence syllabi.
The literature of biology teaching suggests that biological instruction for the present and future should increase the emphasis on problems of conservation and human concerns, yet the findings of this study indicated that only 10 percent of the questions examined concerned these areas. While a review of the literature suggests that ecological levels of biology should be more emphasized, this area was not sufficiently covered in the syllabi examination.

Levin and Lindbeck (167) surveyed five secondary school biology textbooks to study their coverage of controversial issues in biosocial problems. The three BSCS versions were found to rank higher than two other widely used high school biology texts in the quantitative and qualitative use of these controversial issues.

Wofford (315) surveyed secondary school advanced biology teachers, college professors of biology, and college biology majors to develop a model set of cognitive behavioral objectives for a secondary school advanced biology curriculum. A list of behavioral objectives was constructed from the literature on advanced biology. These behavioral objectives were used to construct a questionnaire in which each behavioral objective could be rated as being highly suited, suited, or unsuited for use in a high school advanced biology curriculum. Each behavioral objective was assigned a grade by each of the three groups. From an analysis of these data, a model set of objectives was constructed. Of 114 behavioral objectives, advanced biology teachers considered 31 to be highest ranked, college biology professors considered 16 to be highest ranked. All groups combined considered 13 behavioral objectives to be commonly highest ranked. Significant differences were found between the three groups for only three behavioral objectives.

Specific Content, Concepts, and Processes

The process skill of "hypothesis formation" was studied by Quinn and Kessler (238). These researchers studied the relationship between skills involved in acquiring language and those in an activity based science program. Two sixth grade classes participated in the study, one of which served as control group. The other, the treatment group, had lessons related to inquiry development and instruction on how to formulate hypotheses which
consisted of twelve films and six discussions. It was found that the students involved in the treatment did indeed improve their ability to hypothesize. One of the conclusions of this study was that individuals who are successful in science are also usually proficient in language centered activities. The converse was also true. This is in contrast with the common assumption usually made about language and science skills.

Thiel and George (292) studied the process skill of "prediction," which they defined as the ability to use one or more rules from the same or different rule classes to determine the outcome of an event or series of events. They hypothesized that four factors may affect the use of prediction: (1) experience, (2) ability to infer or to use rules, (3) the types of rules available, and (4) the dimensionality of the task. The subjects in this study were given Piaget tests concerned with seriation, conservation of length, and the classification of objects. The results of these tests were used to judge if the students were concrete operational. In the third grade, 55 percent of the students; in the fourth grade, 75 percent of the students; and in the fifth grade, 77 percent of the students were judged to be concrete operational. Ninety of these concrete operational students from the third, fourth and fifth grades were chosen for the study. Schools in the study had no previous formal science program.

Students were given six prediction tasks. Half of the groups were given an algorithm for solving the problem, which consisted of a verbal description of a rule type. The other half were given no algorithm. The results of the study indicated that, for concrete operational children in grades three through five, the skill of prediction develops independent of formal instruction in science. There was no difference in the attainment of the skill between grade levels. It was hypothesized that concrete operational children may not use rules given to them, unless these rules correspond to procedures that already exist in their own cognitive structures. Thiel and George had originally hypothesized that students given the algorithms would have less trouble making predictions than those who did not receive these algorithms. This hypothesis was not confirmed by the data in the study. A further result of the study indicated that children in grades three through five had difficulty coordinating multiple dimensions when they were
asked to make predictions using these dimensions. One-attribute seriation and classification tasks were less difficult than two-attribute tasks, which in turn were less difficult than three-attribute tasks. Based on this study, it was also found that classification rules were easier for students in grade three to five than were seriation rules, when more than one attribute was contained in the prediction task.

Erickson (86) studied the concepts associated with the general topic of heat. In the first part of the study, interview data were collected and analyzed, while the second part involved the construction of an instrument called the Conceptual Profile Instrument (CPI). This instrument consisted of statements about heat obtained from analyzing the interview data. These statements represented typical ideas of children concerning the kinetic and caloric theories of heat. Children were asked to respond to each statement indicating their belief and their familiarity with the concepts. It was found that fifth, seventh, and ninth graders possessed beliefs about heat and temperature which were based upon common sense and intuitions developed from everyday experiences. Temperature of an object was thought to be related to the amount of heat possessed by that object, and so many children concluded that the temperature of an object depended in part upon its size. Heat and cold were sometimes thought to be substances capable of penetrating objects. Heat was therefore considered to be an active external agent accounting for the expansion and contraction in melting and freezing behaviors exhibited by many substances.

The specific concept "tree" was studied by Klausmeier, Schilling and Feldman (151) in order to determine the effectiveness of special lessons in facilitating and attainment of this concept by children. Using a learning model called the Conceptual Learning and Development Model (CLD), these researchers looked at the levels of concept attainment of elementary school children of the concept "tree." According to this model, there are four possible levels: (1) the concrete level, which consists of the ability to discriminate a concept; (2) the identity level, which consists of the ability to generalize two or more forms of the concept and see that they are the same; (3) the classificatory level, which is the ability to generalize two or more instances of the concept and see that they are the same; and finally, (4) the formal level, which consists of the ability to infer the
concept. The vocabulary associated is connected with a number of researchers, including Piaget, but the terms concrete and formal are not used in the same way that Piaget uses them.

Two experiments are reported by this group. One experiment with 103 fifth grade students showed that there was no significant effect of the treatment on this group. This result was attributed to the fact that students had already had a high level of attainment of the concept before instruction began. The second study had 64 third grade students divided into control and experimental groups. The control group had placebo lessons. The experimental group was instructed using two lessons on the subject "tree." The first lesson was 33 pages of written material which had several aspects including an introduction and a presentation of each of the defining attributes of a tree. Questions were asked following the presentation of each attribute and immediate feedback given to the subjects. The second lesson had a rational set of examples and non-examples to be used to discriminate the concept, and students were taught strategies for evaluating instances to see whether they were or were not examples of the concept. The experimental design went through a number of different stages, including an analysis of the concept to be taught, determination of pre-instructional student characteristics, identification of desired level of attainment according to the CLS model, assessment of students' prior instructional level of concept attainment using tests with exercises designed to assess the various levels of the concept, the designing of pictorial and verbal lessons, the actual instruction, the assessment of the students' post-instructional level of the concept, and evaluation of the results. Results with these third grade students showed that the experimental group performed better than did the control group on the tests for the formal level of attainment of the basic concept presented. The researchers recommend that such lessons can be valuable in facilitating children's attainment of concepts such as this one. At the same time, they recommend the use of concrete experience with younger children and children who have had experiences with examples of a particular concept.

Novick and Menis (212) did a small scale study of high school chemistry students in Israel to check their understanding of the "mole" concept.
The mole concept was developed with a textbook approach, not obviously experimental, although some of the concepts were used to interpret experiments using verification laboratories. A structured interview consisting of 21 items assessing levels of concept attainment from defining the concept to using the concept to solve problems, was developed. Student answers to the 21 items were taperecorded and analyzed. Misunderstandings about the concepts were developed from a "wrong answer" analysis. A product moment correlation between IQ and the score on the interviews was 0.69. Several misconceptions about the concept of the mole seemed to prevail. The mole was thought to be connected to mass rather than number, and to be an exclusive property of gases. The conclusion of the study was that most 15-year-old pupils in Israel do not have a coherent understanding of the mole's significance in interpretation and its use in solving problems. The author suggests that perhaps these students do not function at the cognitive level necessary to understand this concept.

The apparent discrepancies between the ability of some students to learn concepts and the use of those concepts in existing curriculum materials may be the source of significant learning difficulties in the science classroom. For that reason, the study of specific processes and concepts and how students learn these processes and concepts, the age at which they are capable of learning them, and the levels at which the learning is possible all seem to be important questions to be studied by science educators. Most of the studies in this area seem to be descriptive studies. This seems to be a reasonable approach based on the present limit of our theoretical knowledge. However, one would hope that more experimental studies can be done in order to test some of the knowledge that we already possess. Evaluation tools for studying this area vary widely from informal interviews to standardized examinations. The clinical Piaget-type interview seems to be a widely used and valuable technique for studying the understanding of concepts by elementary and secondary students.
College Level Research

College level research was arbitrarily divided into two major parts. The first part includes research concerning comparisons of components of a course or instructional programs. Most of the studies involved a comparison of one teaching or program approach to another with student achievement and attitude toward the course as variables.

The laboratory portion of science courses appeared to gain the most attention. Several research studies modified the laboratory approach and compared the new approach with respect to student achievement and attitudes to what was most often labeled the "traditional" approach.

Another relatively large group of studies focused on the lecture portion of the science course or program. Several of these studies compared various forms of audio-visual, self-paced, individualized, and tutorial aspects to the lecture-recitation or "traditional" approach. In most cases, no significant differences in student achievement were found. Most studies did, however, show that student attitude toward science in general or toward a particular course was improved by whatever modification was used.

Some research studies in the first group focused on course objectives. Generally speaking, the studies did not investigate or question the value or appropriateness of objectives, but focused on if, or when, or how, the objectives should be revealed to students.

Other relatively large categories included in this first part are Course Supplements and Focus or Orientation of the course of study. The categories in the first section are as follows: (1) Laboratory, (2) Lecture, (3) Objectives, (4) Tests, (5) Supplements, (6) Time, (7) Focus or Orientation, and (8) Multiple.

The second part includes research that deals with predicting success of students and evaluation of college instructors. The evaluation applied to college instructors is quite limited in scope and depth since it focuses
primarily on the evaluation of instructor performance by students in the class. The research studies involving students were aimed at predicting success in college, comparing the cognitive performances, and motivation. The categories in the second section are: (1) Students, (2) Instructors.

In some categories a brief introduction is given prior to the review of specific research reviews. In other cases an introduction was not considered to be necessary.

Component Comparisons

Almost one-half of the research reports reviewed in this chapter are classified with this group. The idea in most cases was to alter or modify some part of a course and then see how that affected achievement or attitude. Achievement was usually measured by performance on a standardized or institutionally constructed test. Attitude was most often measured on a semantic differential instrument.

Laboratory

Several investigators modified the traditional laboratory approach used in science. The first group of experiments fit into this category.

Townes (296) compared the effect of vicarious laboratory experiences with conventional laboratory. One group of students collected data without having direct contact with equipment, apparatus, and materials. This group was said to have a vicarious laboratory experience. Another group of students collected data by manipulating equipment, apparatus, and materials. This group was said to have a conventional laboratory experience. The vicarious laboratory group was labeled experimental. The conventional laboratory group was the control.

Townes concluded that the experimental group exceeded the control group on all instruments used to measure achievement. Also the experimental group showed greater competency in the use of science processes than did the conventional laboratory group.
Dickinson (78) also experimented with modifications of the traditional laboratory-lecture approach to determine the effects on student attitude and achievement. He used three different laboratory-lecture combinations. Forty-three students comprised the control group taught by the lecture method only. Another 43 students comprised one experimental group taught by the lecture-laboratory method. A second experimental group of 43 students was taught by the lecture-recitation method. Student achievement was measured on two subject matter achievement tests: one constructed by the instructor and the Nelson Biology Test. Student attitude was compared using the Scientific Attitude Inventory.

Dickinson reported that students taught by the lecture-laboratory method had higher achievement scores than students taught by lecture only. Students taught by lecture-laboratory did not achieve significantly (.05 level) better than those taught by lecture-recitation. Student attitudes toward science were changed more favorably by the lecture-laboratory and by the lecture-recitation methods than by lecture only.

Holloway (127) investigated the effects of 13 open-ended laboratory experiences in physical science on critical thinking ability and attitude of college freshmen. The experimental group consisted of 38 randomly assigned students who were taught by the discussion-lecture method and open-ended (TCCP) laboratory procedures. The control group consisted of 38 randomly assigned students taught by lecture only.

At the end of the 12 weeks instructional period, both groups were administered the Test on Understanding Science and the Watson-Glaser Critical Thinking Appraisal. Significant differences were found between the two groups in critical thinking and in attitude toward science which favored the experimental group.

Dorrance (80) compared two laboratory instructional treatments to a non-laboratory approach to determine their effects on manipulative and cognitive skills. In general biology, laboratory sections were assigned randomly to the three instructional treatments--four received lecture only (control), three received lecture with structured laboratory, three received lecture with a structured demonstration.
A 40-item test on cognitive skills based on the Bingman (BSCS-MCREL) analysis of processes of science was used to measure cognitive skills. The acquisition of manipulative skills was established by performance of a specified laboratory practice (serial dilution) and proper laboratory procedure.

The results showed the laboratory method of instruction superior to the demonstration method in the acquisition of behavior's characteristic of manipulative skills and cognitive skills found in the processes of science.

Cannon (46) compared the effects of student-directed versus traditional, highly structured laboratory on student interest and understanding of the process of science. Eighty students in a general education physical science course were randomly assigned to two laboratory groups. One group was encouraged to use the laboratory to develop and direct their own laboratory activities. The other group used traditional, structured laboratory activities.

Students were pre- and post-tested with the Welsh Process of Science Inventory and Interest Assessment Scales. Cannon found no significant differences between the laboratory groups with respect to interest and understanding of the process of science.

From Canada, Valeriote (299) experimented with a self-paced laboratory course in first-year chemistry. Two sections of 24 students each were assigned to a trial group that did not follow a strict laboratory schedule. Several other laboratory sections of 24 students each followed the weekly three-hour laboratory schedule.

Students in the two trial group sections were allowed to work during the regular laboratory period if they wanted to and at other periods as well. The laboratory was left open 18 hours per week for these trial groups. A certain number of set experiments was required but other experiments could also be done. A written laboratory examination on the set experiments was administered at the end of each term. Based on grades assigned to students for laboratory performance, the self-paced group scored
higher on laboratory examinations and on the final laboratory grade than did the regular group. The difference was statistically significant. The students' responses to a questionnaire about the course revealed that they liked the self-paced format.

Goodson (109) used objective based diagnostic tests and help sessions based on these tests in undergraduate physical science laboratory. The study attempted to assess student learning during various types of help sessions.

Ninety-seven students were randomly assigned to one of four treatment groups and a control group. Each group was a laboratory section. Ten laboratory exercises of two hours duration each were scheduled in conjunction with the physical science course. Three of the ten sections were selected for this study and were taught by Goodson and five teaching assistants.

There were four different treatments for the experimental groups utilizing the results of the objective-referenced diagnostic tests and one non-treatment or control group. The treatments were as follows: (1) A list of objectives for each exercise studied, a diagnostic test based on these objectives, and a help session designed for reteaching incorrect responses on the diagnostic test. (2) Treatment was the same as (1) except the help session simply encouraged students to ask questions concerning items missed on the diagnostic test. (3) Treatment was the same as (1) except students were to use various certain resources to find answers to incorrect responses. (4) Only diagnostic test and help session which advised students to use various sources to find answers to incorrect responses were used. (5) The control group received no treatment.

At the conclusion of the study a criterion test of physical science laboratory achievement was administered to all students. A similar test was administered six weeks later as a retention test. Goodson reached several conclusions: (1) Student achievement was greater in the help sessions in which students were encouraged to ask questions. (2) Student achievement was higher when the diagnostic tests with remediation were
used than when no remediation was used. (3) Retention was greater when students were advised to discover, on their own, correct responses to diagnostic tests.

Hill (125) evaluated a set of commercial slide-tape units designed to teach laboratory technique in chemistry. The program did prove to be an effective means of presenting instruction in basic laboratory technique. Hill's interest was specifically related to creativity. One purpose of her study was to determine if creativity could be enhanced in a specific discipline, such as chemistry, if students received laboratory instruction in which they were encouraged to practice processes considered to be creative and were rewarded for such behavior. Based on the idea that creativity involves divergent production which involves the ability to synthesize and recombine material to form new solutions, this study used teaching methods believed to encourage divergent thinking in the chemistry laboratory instruction.

Hill designed a pretest and posttest for creativity in chemistry using the Minnesota Tests of Creative Thinking by Guilford and Merrifield. The investigation of creativity involved 176 students in 4 laboratory sections in general chemistry. Three of the laboratory sections comprised the experimental group. The fourth laboratory section was the control.

Students in the experimental group had access to the slide-tape instruction of Young and Fiel. The control group did not use the slide-tape presentation. The Modular Laboratory Program in Chemistry by Neidig and Young was the source of the weekly laboratory experiments for all four sections. All four laboratory sections received laboratory instruction which emphasized the importance of creativity. Pretest and posttest scores indicated that all four sections made statistically significant improvement (.05 level of confidence) in both laboratory technique and creativity. The experimental group excelled over the control group in laboratory technique. Hill further concluded that a system of teaching and rewarding creativity through grades can effect an increase in creative abilities.
DeLuca and Renner (76) compared achievement and attitude in traditional laboratory and a structured inquiry laboratory approach in introductory geology. All students attended lecture sessions three hours per week. The lectures were conducted by the geology staff.

The traditional geology laboratory required that a large percent of the total laboratory time (three hours) be given to explanations and clarification of terminology. It was common practice for the instructor to devote the first hour to lecture and providing information using the chalkboard. In some laboratory sessions, most of the three-hour period was used this way.

The structured inquiry approach that was used in the experimental laboratory began with a 10-to 15-minute introduction. Then the students were involved in the activity. A variety of learning aids was used: models, movies, slides, modeling clay, stereographs, etc. When working with the materials, students were encouraged to make observations, collect data, draw conclusions if possible, and answer questions. In the study, 83 students were randomly assigned to two instructors, two classes to each instructor. Each instructor taught an experimental group and a control group.

To measure achievement, an objective test was developed, tested, rewritten and used. Student attitude toward their respective course and self-esteem as a geology student was measured, using a 10-scale semantic differential test. Analysis of the data yielded the following findings: (1) There were no significant differences in achievement between groups, instructors, or methods of instruction. (2) Students in the experimental group had a significantly more positive attitude toward their course than did those in the expository (traditional laboratory) group. (3) Students in the structured inquiry group indicated greater self-esteem as geology students than did those taught by the expository approach.

Generally speaking, the two approaches were equally effective in promoting achievement in geology content. The structured inquiry approach, however, was significantly more effective in promoting favorable student attitude and self-esteem as a geology student.
Lecture

The first investigation in this group is the only research report reviewed dealing specifically with the effectiveness of lecture and lecture style. Since lecture is, no doubt, the most frequently employed teaching technique, this research has broad application in college teaching not only in science but in all curriculum areas.

Johnstone and Percival (140) investigated student attention patterns during lectures mostly in first year chemistry. The investigators were interested in the existence and frequency of general non-attention. More specifically, the research centered on determining the pattern of attention breaks if they did exist. They were also interested in determining whether or not the attention breaks were related to lecture style and which factors, if any, improved the non-attention pattern in a lecture.

An attention break was defined as "a period of general lack of concentration during a lecture involving the majority of the class and not merely isolated individuals." These attention breaks were identified by increases in background noise, students involved in doodling, chatting, looking around, etc. A general feature was a mood of restlessness in the class.

The investigators observed over 90 lecture sessions and sat among classes of 275 to 300 students to make observations. Several of the lectures were attended by independent observers so that the findings could be compared. A "fingerprint" of the lecture was recorded which included aspects of lecture style, times and length of non-attention, and precise content ideas presented during non-attention. There was a very high correlation of these notes taken by the separate observers although they were not in contact during the lecture.

A large class (550 students) of first year chemistry had been divided into two sections, one in the morning and one in the afternoon. Each section received an identical lecture. A comparison was made of these two classes to determine the effect of the attention breaks on achievement. For example, if one class had an attention break during a particular presentation of
content and the other did not, this would be expected to show up on one of
the diagnostic tests which were given monthly.

Patterns of general non-attention were observed. These periods generally lasted from two to four minutes. The first period of non-attention was at the beginning of the lecture, due to the class settling down. The next lapse of attention generally occurred some 10-18 minutes later. As the lecture proceeded, the attention span became shorter and often fell to three or four minutes toward the end of the lecture. This general pattern of shortening of attention span was found in every case where a lecture without a break was given. The rate of decline of attention spans varied from one lecturer to another of the 12 observed. The variables which appeared to affect the rate were difficulty of subject material, delivery rate, legibility of chalkboard work, and lecturer personality.

To determine the effect of non-attention breaks on achievement, certain items from the diagnostic tests were selected for comparison of the morning and afternoon sections of general chemistry. Some items represented periods in which both sections were involved in non-attention. Some items were selected to represent periods when both sections were attentive. Some items were selected to represent periods when one section was attentive and the other was not.

In the cases in which both sections were attentive or both sections were involved in non-attention, while the ideas represented by the test items were covered by the lecturer, there were no significant differences in test scores. In the cases where one section was attentive and the other was not, the difference in scores was highly significant.

Lectures with deliberate variations interspersed usually commanded a better attention-span pattern and had the effect of postponing or eliminating the occurrence of the attention break. The variations included illustrative models, experiments, buzz sessions or problem solving sessions and other such planned breaks.
Lawrence (158) compared the achievement and attitude of students in medical biochemistry in a standard lecture-chalkboard presentation to a slide-tape method and a combination of these methods at the student's discretion. Achievement was measured with regular scheduled quarter exams. Attitude was measured by using a pretest and posttest Likert-type scale directed at the student's attitude toward biochemistry. Another Likert-type scale was used at the termination of the course to determine each student's attitude toward the instructional method used. Lawrence found that informational gain was highest in the group taught only by the slide-tape method. The combination group had the highest rank in attitude toward the subject and the method of instruction.

Combs (64) used the self-paced format in physical chemistry lecture and testing. Lectures were taped so students could listen at their convenience. Also, students could take tests early if homework problems were completed. Taped previews and aids on problem solving for each unit were available to the students. After a period of adjustment to the various avenues available to them, the students using the multiple option approach improved steadily. Class attendance was even better under this option than with the traditional approach.

Ott (218) compared the achievement and attitude of students taught by two methods in a one semester freshman level physics course. One group of students was taught by lecture-recitation-laboratory. The other was taught by audio-tutorial instruction. Students in both groups used the same texts, had the same homework, and laboratory assignments, and were given identical quizzes and examinations.

The standard lecture-recitation-laboratory consisted of two hours lecture and two hours of recitation per week, and a two-hour laboratory every other week. In the AT method, there was a one-hour recitation period per week. All other instruction was at the student's convenience in a learning center staffed by tutors 47 hours per week. This learning center contained materials for self demonstration, audio-tape commentaries and slides. The tapes, commentaries and slides were coordinated with a study guide. The measure of student achievement in the course was the student's final grades.
Another aspect of this study had to do with the method of assignment of students to AT or standard methods of instruction. Four different types of assignment were used as shown in the following diagram which also shows the number of sections in each category.

<table>
<thead>
<tr>
<th></th>
<th>AT</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Preference</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

The results of the study can be briefly summarized as follows: There was no significant difference in achievement as measured by final grades between students in the audio-tutorial group and students in the standard group if students were assigned on the same basis. Students assigned by preference indicated greater satisfaction with the method of instruction they received than did randomly assigned students.

There is much more in the research by Ott than can be dealt with in this brief review. The interested reader should refer to the complete article.

Spevack (276) compared the achievement and final course attitude of non-science oriented chemistry students using the Keller Plan and the Lecture Recitation System (LRS). The Keller Plan or the Personalized System of Instruction (PSI) is a self-paced, mastery oriented, student-tutored instructional method based on the psychology of reinforcement theory.

The distribution of students in the PSI (experimental) and LRS (control) sections was nearly random since students chose their section without knowing if it would be experimental or control. The sections were designated experimental or control on the basis of a coin flip.

Students in the control group had three periods each week conducted by the traditional lecture-recitation method. The students' grades were based on their performance on four regularly scheduled exams and a non-cumulative final. Students in PSI had one period each week designed to
motivate them to learn the course content on their own. During two other scheduled periods each week, the lecture room was turned into a study hall where students could consult with tutors. Students in the experimental group were also informed that their course had been divided into 12 units, each consisting of a reading assignment, instructional objectives, and homework questions. Each PSI student studied the units sequentially at his/her own pace and took mastery exams given by a tutor when she/he believed they knew the material. Alternate exams could be taken without penalty.

All students were tested during the first week with the American Chemical Society-National Science Teacher's Association (ACS-NSTA) Cooperative High School Chemistry Examination, Form 1971. The course instructors and Spevack constructed objective, multiple choice midterm and non-cumulative final exams. The sum of the standardized raw scores on these two exams, with corrections for guessing, was considered a measure of achievement. An analysis of covariance was used to analyze the data. There was no significant difference in the achievement of the experimental and control groups.

A course evaluation questionnaire was used to measure the final course attitude of the students. The instrument, administered during the fifteenth week, allowed the investigator to conclude that PSI students had a more favorable final course attitude.

Calhoun (42) conducted a study using the Personalized System of Instruction (or Keller method) in an undergraduate personality course. The results indicated that progress through the course was related to grade point average. That is, strong students progress rapidly but weaker students can achieve mastery if adequate time is available. The flexibility to go beyond the limits of the semester or quarter is desirable.

Mintzes, Littlefield, Shaub, Rakitan, Richard Crockett and Ronald Crockett (195) reported on five studies involving 693 secondary and college students in individualized biology courses. They identified student characteristics relating to high achievement in these courses. Prior knowledge, intellectual ability, and motivation were related to high student performance in all five studies.
Wolfson (316) compared lecture and individual research as methods of teaching a required science course for achievement and retention. Students in three classes of 30 students each, all taught by the same person, were permitted to select either of two approaches. One approach, the "formal" approach, required students to be present at lectures given once each week for two and one-half hours. They also attended a laboratory section that met once each week for two-and one-half hours. The second, or "informal" approach, required the student to select a research topic from a list provided by the instructor or decided on with the approval of the instructor. A statement and brief outline of the topic was required at three weeks, a bibliography at eight weeks, a formal outline at twelve, and the paper at fifteen weeks. Each student gave a 10-minute talk on his/her paper. The "informal" group participated in laboratory with the formal group.

Objective tests were given to all students in both groups three times during the course. There was no significant difference in science knowledge at the beginning of the course. After 15 weeks, the formal group performed better on tests but not significantly better. After a period of seven weeks a retention test was administered. The informal group outperformed the formal group at a level of significance greater than .05.

Objectives

These research studies dealt primarily with student achievement and attitudes related to variation in type and use of course objectives.

Miles (194) studied the relative effectiveness of behavioral and non-behavioral objectives on achievement in an introductory geology course. The behavioral objectives specified what was to be learned and how the learning would be demonstrated. The non-behavioral objectives, or outline objectives, consisted of listings of terms and concepts in hierarchical groups. A second objective was to assess the attitude and preferences of the students with respect to the two types of objectives.

Two intact classes, totaling 32 students, were randomly assigned as experimental and two classes (30 students) were designated as comparison
groups. The experimental group used behavioral objectives and the comparison group used outline objectives for one quarter. Tests were given on two-week intervals and the students were also given a comprehensive final exam.

Analysis of the data (multivariate analysis) indicated that the overall achievement of the experimental group was significantly higher than that of the comparison group. During the next quarter, all students were exposed to both types of objectives. An attitude scale showed an almost unanimous support for the use of objectives. Also, when students were given a choice of behavioral objectives or outline objectives, the majority of the students chose the outline form and appeared to view the outline form as more useful to them.

Leonard (165) conducted a study in which (among other variables) the students' perceived usefulness of, and attitude toward, prior knowledge of instructional objectives in a physical science course was assessed. The student groups which were given the instructional objectives found them useful and expressed a desire to have such statements of instructional objectives made available to them in other courses.

Tests

Trochet (297) studied the influence of a computer-based repeatable testing program on student achievement and attitude in a general education physical science course. Students (108) in the experimental group followed a repeatable testing program in which they could take unit tests a second and third time without penalty. The repeated tests were an alternate form of the first test. The control group students (100) had only one opportunity to take each test. Three unit tests, an opinion survey about the course and the testing program, and a midterm achievement examination were given.

Students in the experimental group scored significantly higher on two of the three unit exams. There was, however, no significant difference between the experimental and control groups on the midterm examination. It appears that repeatable testing allowed short-term learning, but there was not a significant difference in long-term retention. Students preferred
the repeatable testing but this preference had little, if any, influence on their attitudes about other aspects of the course.

Rosati (252) reported on Announced Repetitive Tests (ART), a teaching method similar to PSI. The method involves defining a minimum acceptable course content and expressing it in the form of a certain number of questions. The questions are grouped into question sheets and quizzes and given to the students together with study guides. The student studies the quiz until he knows the material. Then the student takes the quiz and has it checked by an instructor. The student may repeat incorrect questions or variations of them, until they are answered correctly. Students and instructors liked the ART method of teaching. Rosati feels, however, that the method would not be successful with classes of over 30 students.

Supplements

Research studies placed in this group are those which generally introduce some innovation in the form of supplementary materials or activities into an instructional program to enrich or enhance it or to solve a particular problem in the course of study.

Reiss (242) introduced the use of a personal journal kept by the student as a means of measuring students' learning and evaluating the students' knowledge of physics. He determined that the journal was an effective instrument by which the student could document and demonstrate learning.

Kromhout (153) investigated the effectiveness of computer review lessons as a supplement to an introductory physics course. Fifty-five computer lessons were made available to the 170 students involved in the study. Generally favorable results were achieved by users of the supplementary lessons.

Tamminen (290) supplemented a general chemistry course for non-science majors with a workbook of programmed chemistry problems. Four lecture sections (120 students) of college general chemistry were used in the study.
All sections met twice a week for a one-hour lecture and a two-hour laboratory session. Another one-hour discussion period was optional. Students in two sections used a basic chemistry text with the Programmed Supplement of General Chemistry Problems written by the investigator. The other two sections served as controls. A two-way analysis of covariance was employed to evaluate the data. Tamminen found that there was no significant difference \((p > .05)\) between the treatment and control groups.

Wooley (318) developed and evaluated a supplemental Computer Assisted Instruction (CAI) program to improve the ability of students to cope with the mathematics in introductory astronomy. Three types of CAI modules were developed for use: (1) guided discovery modules which provided feedback guiding students to correct responses; (2) discovery modules which supplied only knowledge of results feedback; and (3) placebo modules.

Wooley taught the two sections (94 students) involved in the study. A mathematics ability pre- and posttest was used. Course content, retention, CAI attitude, and course evaluation instruments were administered at appropriate times during the semester. Wooley concluded that neither of the experimental approaches resulted in an increase in students' mathematics ability or in transfer of learning to mathematical portions of the course.

Meade (192) studied the use of a computer as a problem solving tool in college physics. Three areas were explored: (1) achievement on course examinations and quizzes, (2) attitudes toward problem solving, and (3) feasibility of the use of computer problem solving on a regular basis. An experimental group of 46 students was required to write and execute computer programs based on problem assignments given to the control group of 89 students. Both groups had the same instructor and were given the same quizzes and examinations. Students in both groups were given surveys to measure their attitude toward physics, problem solving, and the particular course of study in which they were involved.

Meade found that students who were taught problem solving by the computer approach did not achieve higher scores on examination and quizzes and did not have a more favorable attitude at the end of the course toward
A problem solving or physics. Also, students in many cases reported that the computer programming took too much time.

**Time**

Blind (31) assessed the effects of compression of audio and video components on learning during an instructional television presentation. The results indicated that in this presentation learning, as demonstrated by a pencil-and-paper test on multiple choice items, occurred as readily with compressed material as with uncompressed material.

Studdard (283) demonstrated that the same goals could be achieved during an interim term in a college level physical science course as in a regular semester course.

**Focus or Orientation**

The following four research reports describe investigations in which a focus or orientation was utilized in the course that differed from the traditional course in that content area. One of the research studies, the last one, was included with this group because the evaluation was based in part on the election to take the second course of a two-semester sequence. The second course was not required.

Graham (110), attempting to make physics more attractive and familiar to students, developed a new course emphasizing the qualitative aspects of physics. He tested the hypothesis that there would be no significant difference in achievement by students who studied physics by the qualitative approach and those who used the quantitative approach.

A control group of 32 students enrolled in a physical science course which used the quantitative approach was used as the comparison group. Twenty-six students, the experimental group, were taught physics using the qualitative approach. Both courses were taught by the same instructor. The instruments used included the Dunning-Abels Physics Test (Form E pre-test, Form P posttest) and a science and mathematics interest inventory pre- and posttest.
Analysis of the data showed that there was no significant difference in physics achievement using the two approaches. There was also no significant difference in the two approaches with respect to interest.

Blomme (32) compared the attitude toward science of students in a traditional chemistry course and an environmentally oriented general chemistry course for non-science majors. The Science and Scientists Attitude Inventory (SASAI) developed by Laméne Motz was administered as a pretest and posttest. The pretest indicated that no significant difference in attitude toward science existed between the sample populations of the two groups. There was, however, a significant difference in the attitude toward scientists at the .01 level. The traditional chemistry group had a more positive attitude.

After completion of the two courses, there was no significant change in the attitude toward science by the students in the environmentally oriented course. There was a positive change at the .001 level of significance toward science by the traditional chemistry group. The environmentally oriented course produced a positive change in the attitude toward scientists in the sample population of the traditional course.

Williams (309) measured the effect of the course Physical Science for Non-Science Students (PSNS) upon the problem solving skills of non-science college students. Ten classes using PSNS were compared to ten classes using another science program.

The College Science STEP Test was administered as a pretest and as a posttest. The pretest score of the two groups were subjected to analysis of variance. The results indicated no significant difference between mean performance of the two groups.

The pretest and posttest class mean scores of the PSNS and non-PSNS groups were subjected to an analysis of covariance. The results indicated that there appeared to be no significant difference in the improvement of problem solving skill in the PSNS and non-PSNS groups.
Mauldin (180) studied the effectiveness of a physical science inquiry course in changing the attitude of college students toward scientific methods. In this study, which involved 300 students, use was made of the fact that the course was available as a two-semester sequence. The second half of the sequence was optional:

A pretest and posttest with a twelve-scale semantic differential applied to each of the phrases "physical science," "doing experiments," and "making inferences from observations" was used to assess attitude. Science interest was measured with two questions on the frequency of outside reading and science activities.

Analysis of data from the instrument used in the pretest and posttest indicated a significant change in attitude in the unfavorable direction. Also, no significant relation was found between student attitude level or change in the first course and student decision to take the second inquiry course.

Multiple Components

The first research report in this section is different from most of the others in two ways. First, it involves the use by students of many different components or variations from a traditional course. Most of the research studies focused on one or two components. Second, since the course of study and objectives are different from traditional courses, comparisons are probably not as meaningful.

The third research report indicates that chemistry courses offered by colleges and universities for non-science majors have no detectable patterns of content.

Naegele and Novak (204), in a paper presented at NARST, 1975, reported on a two-semester introductory physics sequence developed over a five-year period at Cornell University. The course involves between 20 and 30 staff members and from 500 to 700 students. The course is aimed primarily at students pursuing careers in life science areas.
The physical facilities occupy a 10,000 square foot area with some 90 carrels containing audio cassette recorders, film loop projectors, demonstration and laboratory equipment, etc. A large portion of the space is occupied by a testing center and post-exam tutoring rooms.

The function of the learning center is to provide a wide variety of instructional alternatives to be used by the student as often as she/he wishes. Students work at their own pace, with help from an instructor as needed, on nine modules each term. A standard textbook is also used with the course and students are provided a study guide containing a list of learning objectives, a list of recommended activities, laboratory instructions, audio-tape supplements, supplementary problems, programmed materials, and sample examinations. Examinations are self-paced mastery type and are non-scheduled and repeatable.

Student attitude was assessed using an end-of-course questionnaire. The information from this questionnaire, when compared to a background questionnaire administered at the beginning of the course, indicated substantial gains in student attitudes. The attitude of the staff, primarily graduate physics teaching assistants, was also reported to be extremely positive. The staff expressed a marked preference for the self-paced format over the more traditional format.

Because the content, objectives, and evaluation procedures have significantly changed since the course was taught traditionally, Naegele and Novak propose that it is virtually impossible to definitely compare overall achievement under the two formats. Castaldi (49), in a report on research at Cornell using the audio-tutorial format, self-pacing, and mastery testing, states that achievement in this format allows as good or better achievement than in a traditional lecture-laboratory-recitation format.

Blatt (30) used questionnaires to obtain information about introductory chemistry courses specifically designed for non-science majors. Questionnaires sent to four-year colleges accredited by the American Chemical Society established that most of these schools do offer a course for non-science majors. The courses differ from traditional chemistry courses in many ways and differed from each other such that there was no detectable pattern among the information obtained.
Instructors and Students

The college instructor, as a part of the educational program, is touched upon by two reports. In both cases the evaluation of the instructors was based on student opinion surveys. The research reports which dealt with students were richer in variety.

Counts (69) investigated the effects of certain student characteristics on the students' ratings of instruction in biology at the college level. The basic population for the study was 374 students enrolled in Freshman Biology. Among the findings were reactions of students to class size, workload in class, and course difficulty related to their biology instructors' performance. Students who thought the class size was too large gave instructors the highest ratings. Students who said that class size did not matter gave the lowest ratings. Students who perceived the workload as much heavier than other courses gave much lower ratings to instructors than those who believed the workload was "about right." Students who perceived the workload as lighter than other courses gave the highest ratings. Students who thought the course was either very difficult or very easy gave significantly lower ratings than those who thought the course difficulty was "about right."

McLaren (191) posed the questions in his study, "Do students rate instructors with training in professional education higher than those instructors who lack this training?" and "Do students rate instructors who have completed the Ph.D. degree higher than those instructors who have not completed their doctorate?"

Sixteen member colleges of the Council for the Advancement of Small Colleges in Ohio, Indiana, and Michigan were used in the study. The senior biology and chemistry students in those colleges evaluated science instructors using a simple numerical ranking scale, and the Student Instructional Report. Information was obtained from instructors by interview. There were no significant differences in the ratings of instructors on either of the two questions.
Students

Yekeson (326) tested a set of criteria to be used by academic advisors for the prediction of success and placement in general college chemistry. The independent variables tested were ACT scores, high school grades, scores on the 1958 ACS-NSTA cooperative examination for high school chemistry, and scores on the Western Michigan University Chemistry Placement Examination (WMUCPE). The sample consisted of 339 students who had completed chemistry 101 and 102. Yekeson found that no single predictor of success tested had a high relationship (coefficient of correlation ranging from .70 to 1.00) with the chemistry grade.

Using stepwise regression analysis, six variables were identified as effective predictors of the general chemistry grade. These variables were: (1) ACT English Scores, (2) ACT Mathematics score, (3) high school mathematics grade, (4) high school social studies grade, (5) high school natural science grade, and (6) the ACS or WMUCPE. Yekeson developed a prediction equation using the ACT mathematics score and the WMUCPE score.

Spencer (275) conducted a study to find the relationship of certain variables to the grade point average in general biology. The independent variables tested were high school grade point average, SCAT verbal and quantitative scores, and CGP mathematics, verbal, motivation and biology interest scores.

A random sample of 140 students was drawn from a list of 582 students who had already completed biology courses. The SPSS programs were used to conduct a Pearson product-moment correlation and a stepwise multiple regression analysis of the sample. Spencer concluded that the best single predictor of success in general biology was the high school grade point average.

Lipton (169) asked the question, "In introductory science classes, is there a differential relationship between academic ability and achievement, and between academic ability and attitude that depends on the personality characteristic of internal-external control?"
The predictor variables consisted of the verbal and mathematical scores of the Scholastic Aptitude Test (SAT), the scores of the Rotter Internal-External scale, and the linear cross product terms of internal-external control and ability. The criterion variables were course achievement as measured by pooled standard scores of mid-term and final scores for each class, and the intellectual and emotional scales of the Scientific Attitude Inventory. Three multiple regression analyses were performed to analyze the data. The results showed the SAT to be a valid predictor of academic achievement.

Hedges and Majer (120) determined the relationship between grades received in prerequisite biology, chemistry, mathematics, and physics courses and subsequent grades received in science major areas.

The subjects were 195 students who graduated with majors in applied physics, information science, applied mechanics and engineering sciences, biology, chemistry, mathematics or physics. Multiple regression analysis was used.

Hedges and Majer concluded that grades in major prerequisite courses were accurate predictors of upper division grades in the major area. Lower division physics grades were the best predictor for four of the seven areas and lower division mathematics grades were the best predictor in two of the three remaining areas. An interesting, and no doubt unexpected, finding of the study was that lower division biology, chemistry, mathematics and physics grades were not the best predictors of upper division grades in those areas.

Yett (327) conducted a study to determine the effectiveness of high school laboratory experience in preparing students with the necessary skills and knowledge requisite for satisfactory performance in college chemistry programs. A questionnaire consisting of 52 items of laboratory skills and knowledge selected from chemistry textbooks and manuals was sent to chemistry teachers in 110 high schools throughout the country. The teachers rated the importance which they assigned in teaching of the 52 items.
College instructors of general college inorganic chemistry from 298 two- and four-year colleges and universities rated students regarding their skills at the beginning of their general inorganic chemistry course. Analysis of the data indicated that although high school chemistry teachers emphasized certain laboratory skills, first year college inorganic chemistry laboratory students failed to exhibit knowledge of these skills. College chemistry professors expressed doubt about the value of high school chemistry laboratory courses in preparing students for college chemistry programs.

Motivation

Winsberg and Ste-Marie (314) studied the relationship of motivation and academic achievement in physics. For this study they used three types of motivation: (1) motivation to satisfy unfulfilled needs for security and belonging; (2) motivation to satisfy esteem needs such as the attainment of status within society through access to higher education, higher social status, higher income, etc.; and (3) motivation to satisfy growth needs such as the individual's needs to self-actualize, to create and divert his energies into intellectual, cultural, or humanitarian outlets.

The question posed for the study was: How does the individual student's academic achievement in physics relate to his measured strength in each type of motivation?

The sample for the study was the entire group of students registered in Physics 302. There were 78 students in the sample who completed the tests utilized. The sample was made up of approximately equal numbers of male and female students.

Each student's motivation to satisfy security needs, esteem needs, and growth needs was measured using the Merritt College Motivation Inventory (MCMI) developed by Coughren in 1972. Academic achievement was measured by the final or cumulative grade which the student received in the course. All of the subjects were evaluated by the same teacher.
The data collected and evaluated indicated a significant but negative relationship between the need for security and academic achievement in physics. That is, those who were not self-confident were not successful in physics. The correlations between motivation to satisfy esteem needs or growth needs and academic achievement in physics were not significant. Winsberg and Ste-Marie interpret their findings to suggest that a lack of motivation to study physics may be due to a need for security rather than a distaste for physics.

Cognitive Preferences

Wright (320) studied the different cognitive preferences of college undergraduate students majoring in science, mathematics, and engineering. A sample of 241 undergraduate students majoring in science, mathematics and engineering was used in the study.

The students were administered a revised CPE-II test which measures the subject's preference for memory, application, and questioning. Other instruments developed by Wright were also used. The students' cumulative grade-point averages and ACT were used.

The study revealed that there were no significant differences between the cognitive preferences of science majors and mathematics majors. There were significant differences between engineering majors and science and mathematics majors. Science majors displayed a significantly stronger preference for questioning than did engineering majors. Engineering majors displayed a significantly stronger preference for memory than did mathematics majors. Students in all three areas showed less preference for memory than for application and questioning.
References

Citations containing ED information indicate documents available from ERIC Document Reproduction Service, P.O. Box 190, Arlington, Virginia 22210.


35. Bowyer, Jane Ann Baker. "Science Curriculum Improvement Study and


on Attitude and Achievement in a High School Chemistry Program." Dis-

Grade Students in Learning Advanced Biologic Concepts." Dis-

39. Brush, Edwin Graham, Jr. "Science Education in a Philippine Village." Dis-

40. Buckley, John Thomas. "A Comparative Study of the Relationship of
the Presence of the Elementary School Science Specialist Toward
the Science Attitudes of Teachers and Students and Student

41. Bybee, Rodger Wayne. "The Implications of Abraham H. Maslow's Philo-
sophy and Psychology for Science Education in the United States." Dis-

42. Calhoun, James F. "Self-Pacing and Performance in a Personalized
System of Instruction." Journal of College Science Teaching,

43. Camp, David Lawrence. "An Investigation of Six Major Logical Groupings
of Concrete Operational Thought." Dissertation Abstracts, 36

44. Campbell, Richard L. and Luis A. Martinez-Perez. "A Study of Relation-
ships of Science Attitudes, Achievement and Self-Concept of Pre-
Service Teachers." Paper presented at the annual meeting of the
National Association for Research in Science Teaching. San
Francisco, California, April 23-25, 1976. ED 125 898

45. Canary, Pat, Carol Hudacheck and Robert D. Allen. "Student Response

Investigating Interest and the Understanding of the Process of
Science in a General Education Physical Science Course." Dis-


ED 123 093
ED 124
ED 126 093
ED 128 161
ED 130 161
ED 130 839
ED 131 617
ED 130 839
ED 121 617
ED 129 561
ED 130 839
ED 121 617
ED 129 561


155. Lantz, Alma; Anita S. West; and Linda Elliot. "An Impact Analysis of Sponsored Projects to Increase the Participation of Women in Careers in Science and Technology." Research Institute, Denver University, Denver, Colorado, June, 1976. ED 130 840


178. Markle, Glenn and William Capie. "The Effect of the Position of
Inserted Questions on Learning From an Activity-Centered
Science Module." Journal of Research in Science Teaching,

179. Martin, John F. "Analysis of John M. Mayfield's 'Factors Affecting
Rationality in the Discussion of a Problem by Small Groups of
Secondary School Students'." Science Education, 60(2):
185-192, April-June, 1976.

Science Inquiry Course in Changing the Attitudes of College
Students Toward Scientific Methods." Dissertation Abstracts,

181. Maupin, Pauline Hicks. "An Assessment of the Effectiveness of
Selected Aerospace Education Workshops in Tennessee." Disserta-

Achievement as Measured by Performance on the New York State
Regents Examination." Dissertation Abstracts, 36(8): 5051-A,
February, 1976.

183. May, David H. and Frank E. Crawley. "Assessment of the Effects of
the Competency Based Teacher Education Experience on the Acqui-
sition of a Teaching Model." Paper presented at the annual
meeting of the National Association for Research in Science
ED 123 107

184. Mayer, Victor J. "Requirements in Earth Science Teacher Preparation
Programs: 1965 to 1974." Science Education, 60(2): 223-234,
April-June, 1976.

185. Mayfield, John M. "Factors Affecting Rationality in the Discussion
of a Problem by Small Groups of Secondary School Students."

571-574, March, 1976.

187. McDonald, Leo A. "Accidents in the Chemistry Laboratory-Causes and
Prevention." Dissertation Abstracts, 36(9): 5781-A, March,
1976.

188. McGuire, Christine. "Simulation Technique in the Teaching and Test-
ing of Problem-Solving-Skills." Journal of Research in Science

189. McIntyre, Patrick J. and Jack A. Reed. "The Effect of Visual
Devices Based on Bruner's Modes of Representation on Teaching
Concepts of Electrostatics to Elementary School Children."


   ED 127 202
   36 p.

   ED 127 201
   35 p.

   ED 129 592
   27 p.


   ED 127 143
   53 p.


