This review presents a broad perspective of the state of knowledge in science education, a description of existing trends, the areas where research is needed, and tentative answers to persistent problems that have been identified from the research. The review is organized into nine categories. The first is an historical perspective dealing with some of the trends in science education. The next category presents research related to learning and development including Piagetian and Ausubelian studies, instruments and measurement, and student characteristics. This section is followed by research related to teaching and learning strategies. Cognitive styles, student characteristics, teacher training and behavior, evaluation in science education, concepts and content, and miscellaneous studies are presented in the remaining sections. Each section and subsection is followed by a summary, if enough research was available to permit generalizations to be made. At the conclusion of the paper, a general summary is given which relates the science education research of 1978 to that of the past and discusses the trends and implications for future research. (Author/SA)
A SUMMARY OF RESEARCH IN SCIENCE EDUCATION--1978

Dorothy L. Gabel
Martin H. Kagan
Robert D. Sherwood

Indiana University
Bloomington, Indiana 47405

National Association for Research in Science Teaching

Clearinghouse for Science, Mathematics, and Environmental Education
The Ohio State University
1200 Chambers Road
Columbus, Ohio 43212

December, 1979
This publication was prepared with funding from the National Institute of Education, U.S. Department of Health, Education and Welfare under Contract No. 400-78-0004. The opinions expressed in this report do not necessarily reflect the positions or policies of NIE or HEW.
INDEX

A

Advance organizers 20, 21, 36, 37, 41
Affective education 54, 60, 91-96
Africa 53, 81
Anxiety 38, 43
Aptitude by treatment interaction (ATI) 8, 17, 38, 42-43, 49, 53, 101
Attitude
toward ecology and environment 78, 80, 82, 84-86, 89, 103
toward environmental education 60, 85, 86
toward science (students') 1, 6, 8, 9-10, 12, 22, 23, 24, 25, 26, 27, 28, 30,
31, 38, 39, 40, 43, 46, 51, 53, 54-55, 56, 65, 66, 67, 77, 78, 79, 80, 81-82,
83, 88, 89, 91, 97, 101-102
toward science and science teaching
community's perceptions 77
inservice teachers' perceptions 60, 61, 65, 66, 68, 69, 75-76, 77, 86, 89, 102
preservice teachers' perceptions 49, 60, 66, 81, 89
Audiotapes 35, 57, 58
Audio-tutorial 21, 22-23, 28, 78, 93
Audio-visual materials 2, 38, 53, 73
Australia 88
Ausubel 20-21, 36, 99-100

B

Behavioral objectives 3, 21, 36-37, 41, 59, 90, 92, 100
Bilingual education 32, 34, 52
Biology 8, 17, 18, 19, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 33, 35, 39, 40,
43, 44, 47-48, 52, 53, 54, 62, 65, 67, 71, 74, 77, 78, 80, 81, 82, 84, 85, 86,
88, 90-91, 93, 94, 95, 96
Bloom's taxonomy 7, 91
Botany 27
Brazil 97
British Columbia 85
BSCS 22, 67, 80, 86

C

Careers 25, 82, 88, 97
Case studies 67, 73-74
Cerebral hemispheric dominance 19, 20
Chemistry 6, 8, 9, 11, 18, 19, 24, 25, 26, 30, 35, 36, 44-45, 47, 48, 53, 54-55, 58,
65, 67, 77, 78, 95, 96
Classification 10, 17, 19, 94, 95
CLOSE 33
Cognitive conflict 10
Cognitive structure 6, 12, 20, 45, 93, 95
Cognitive style 5, 17, 20, 42-50, 64, 101-102
College 5, 7, 8, 11, 12, 14, 15, 18, 22, 23, 24, 25, 26, 27, 31, 34, 35, 37, 38, 39,
44-45, 47-68, 49, 53, 54, 55, 57-60, 77, 78, 81, 83, 85, 89, 91, 92, 93-94, 96
Combinatorial analysis 15, 16, 17, 18
C (continued)

Competency, teacher education 60, 62-63
Computer assisted instruction (CAI) 23, 37, 96, 100
Concept mapping 93, 95, 96
Convergence/divergence 20, 43, 91
Conservation (process) 10, 19
Controlling variables 7, 12, 13, 14, 16, 27
Correlative reasoning 14, 18, 19
Creationist theory 84, 95
Creativty 19, 28, 30, 39
Critical thinking 11, 22, 25, 36, 42, 54, 101
Cross-cultural studies 19-20
Curiosity 28, 29, 79, 80

D

Delphina technique 62

E

Earth science 68, 71, 72, 77, 81, 82
Ecology 78, 80, 85
Elementary school 5, 10, 13, 14, 15, 18, 19, 20, 27, 28, 29, 30, 31, 32, 34, 36, 38, 39, 40, 42, 49, 53, 58, 59, 60, 61, 63-64, 66, 68, 69, 71, 73-74, 75, 76, 77, 78-79, 85, 88, 89, 90, 96, 102-103
Energy 86, 96
Engineering 2, 15, 55, 65, 75
England 82, 88, 95, 97
Environmental education 60, 80, 82, 83, 84, 85, 86, 88, 89
ES
ESS 75, 79
Essence I 79
Evaluation models 92
Europe 97

F

Feedback 28, 37-38, 41, 57, 61, 100
Field dependence 20, 30, 42, 43, 44, 46, 48, 49, 50, 101
Flander's interaction analysis 50, 67, 69
Frye readability 33, 85

G

Games 40, 86
General science 30, 37, 74
Genetics 21, 23, 36, 82
Grammatical construction 34
Graphics 8-9
Group Embedded Figures Test 20, 42, 46, 48, 49, 50
H

High school (See also Secondary school.) 6, 7, 8, 9, 13, 14, 16, 18, 21, 24, 28, 29, 31, 33, 38, 40, 44, 45, 52, 55, 58, 63, 65, 71, 73-74, 77, 80, 85, 90, 95

I

Inductive thinking 28, 30, 31, 43, 58, 66, 100
Inquiry 28-30, 31, 40, 44, 58, 59, 60, 66, 67, 74, 79, 80, 89, 100, 102
IPS 38-39
IS 59-60
ISCS 51, 80
ISIS 46
Israel 26, 31, 47, 65, 76, 80, 94, 95

J

Junior high school (See also Secondary school.) 7, 11, 12, 14, 16, 18, 27, 33, 37, 38, 39, 69, 71, 87

K

Kinetic structure 32, 35-36, 58, 101, 102

L

Laboratory 6, 11, 23, 24, 26, 28, 30-31, 35, 37-38, 80, 90-91, 95, 100
Language arts 32, 61, 100
Latent partition analysis 91, 93
Lecture 9, 23, 24-25, 30, 31, 36, 38-39, 47, 57, 80, 83
Locus of control 46, 48, 49, 50, 101

M

Mainstreaming 100
Marine science 81, 82
Mastery learning 24, 26
Metrics 65, 91
Microteaching 57, 58, 61, 68
Middle school 7, 60, 66, 67, 83
Minorities 55, 71-72, 75
Morality 19, 83, 84, 101
Museums 28, 38
National Assessment 71-72, 102-103
National surveys 71-74, 102-103
Nature of science 60, 62, 63, 65-66, 87-88, 89, 91
Nigeria 65, 81, 91
Nova Scotia 66
Nuclear science 83
Nursing 54

OBIS 85
Objectives (See Behavioral objectives.)
Operative comprehension 19, 45

Peer interaction 10
Philippines 90
Photographs 34, 40-41
Physical geography 24
Physical objects 7, 8, 9, 11, 12
Physical science 9, 12, 18, 19, 23, 38-39, 40, 53, 54, 55, 66, 71, 72, 82, 92, 93-94, 96, 97
Physics 11, 12, 15, 18, 19, 21, 24, 25, 27, 31, 34, 38, 45, 47, 48-49, 53, 55, 61, 67, 69, 81, 82, 92, 93, 97
Physiology 8, 10
Piaget 6, 7-20, 42, 44, 46, 48-49, 50, 87, 94, 99-100
Pictorial diagrams 33-34, 36, 38
PP 69, 81, 82
Prescribed activities 26
Probability 14, 16, 71
Problem solving 7-8, 9, 11, 12, 13, 78
Proportionality 7, 8, 9, 12-13, 14, 15, 16, 18, 24, 45, 99, 100
Propositional logic 14, 17, 19, 25
Processes of science 23, 32, 51, 54, 59, 68, 90
PSCS 81
PSI 11, 12, 23, 24-26, 28, 30, 48

Question format 13
Questioning 6, 34, 38-39, 41, 44, 57, 58, 59, 91-92, 100

Readability 33, 35, 36, 85
Reading 6, 8, 19, 25, 32-33, 36, 46, 74, 78, 100
Review 26
S

SAQA 32, 78, 90
Science attitude inventory 82
Science misconceptions 94
Scientific literacy 79, 87-88
SCIS 50, 57, 66, 75, 78-79, 90
Secondary education (See also High school or Junior high school.) 6, 12, 13, 14,
16, 18, 19, 30, 34, 41, 43, 44, 45, 46, 51, 52, 60, 63, 64, 66, 67, 71, 73-74,
77, 78, 80-81, 82, 83, 85, 86, 91, 102-103
Self-actualization 64-65
Self-concept 19, 33, 45, 52, 53, 79
Self-pacing 22, 24-26, 48
Semantic memory 6
SES 72, 89
Sex differences 17, 18, 20, 27, 30, 40, 43, 44, 46, 51, 52, 55, 66, 71-72, 85, 89
Sex education 76
Simulation games 40
Study habits 24, 39
Study questions 39, 41
Superstition 52
Supervision 62, 64, 73, 74
Surveys 71-74, 102-103
Sweden 13

T

Tape-tutor 23
Tasks 14-17
Teacher training 40, 49-50, 57-64, 86, 89, 102
Teaching assistants 61-62
Test characteristics 15-17
Textbook evaluation 7, 33, 91-92
Topological groupings 15
Transparencies 40
TV 40, 58

U

Unit-box approach 40, 59
University teaching (See college.)

V

Values education 50
Verbal interaction 50, 68-69
Videotapes 17, 58, 62, 63, 67-68, 69
Visual perception 38, 40, 47, 53

W

Wait-time 38-41, 57, 59, 102
Women 55
Workstyle 51, 101, 102
INTRODUCTION

Historical Perspective

The review of research in any area is best seen in relation to research of the past. By comparing present-day trends with those in the past, we can see where we have been and may be able to predict where we are going. For this reason we begin with a brief summary of science education research as synthesized from earlier reviews in this area. Mallinson, in the introduction to A Summary of Research in Science Education—1975, gives a detailed description of the history of these reviews.

The earliest years of science education research (pre-1937) reflect a trend that was common in most educational research. Many articles included in the reviews reported on new curricula and gave statements of objectives for science instruction. Most of the articles that were labeled "research" were ones that described two techniques of instruction and used correlation coefficients to indicate the strength of relationships. No studies of this period used inferential statistics, not even t-tests. Randomization of groups was not employed; matching was the accepted method of group equilibration. There was, however, some trend toward the end of the period to look at more general questions in instruction rather than simply comparing two methods in one school setting.

The second period of science education research can be described as extending from 1938 to approximately 1957. During this 20-year period changes in research perspective and techniques are evident. In earlier years researchers had shown some interest in student attitude, but this was usually very closely related to an ability measure, i.e., do student attitudes towards Y affect their ability in Y? This second period showed a trend of looking at science attitudes as a basis for developing programs in science.

The number of methods articles reported in the reviews showed a decline during this period. While these types of studies were
probably still being done, the reviewers of most summaries of the period
did not choose to include them in their works. This shows some develop-
ment in the area of methods in that there was an attempt to look at
larger generalizations rather than at specific methods. However, this
approach did raise a serious question. Had the theoretical basis for,
learning been developed sufficiently to provide a firm support for
science education? It appears that much of the work was based on applying
the "scientific method" to science instruction. Little research had
a basis in learning theory.

The large growth in methodological changes in this time span is
interesting to note. While the first ten years of the period were very
similar to the pre-1937 years with only correlational work being
reported, during the last ten years many advances were made. Random
sampling first appeared in the 1949 review. It is interesting that
R. A. Fisher first proposed randomization in 1925,\(^1\) a 24-year gap between
theoretical considerations and actual practice. The use of inferential
statistics such as t-tests also began to occur at this time as well as
the use of linear regression. Psychometric considerations of reliability
and validity through the use of the Spearman-Brown formula and of "panels
of experts" for content validity became more common in the literature.

The development of education technology brought about another trend.
Audio-visual materials became more available to schools in the fifties.
This availability introduced questions about the usefulness of these
types of materials and spawned several research studies. Problems of
controlling intervening variables that were of concern in earlier
"methods" research reappeared at this time.

In the sixties, a change in focus in educational research occurred.
Two major events, one external to education and the other internal, had
enormous effects on the science education community. The external event
was the Russian satellite program and the resulting push for the United
States to develop more scientists and engineers. The large increase in
federal money produced many curriculum development projects. With these
projects came the resulting evaluation studies and research on the bene-
fits of one curriculum versus another.

The second event was the Woods Hole Conference. Although the major
purpose of the 35 scientists, scholars, and educators was to discuss how
education in science might be improved in elementary and secondary
schools, the final result was curriculum projects with more psychologi-
cal bases. This stimulated research with a psychological thrust.

Methodological changes were, in some ways, as dramatic as the
philosophical changes of the period. Inferential statistics, includ-
ing analysis of variance and multiple linear regression, became more
widely used. The publication of Campbell and Stanley's Experimental
and Quasi-Experimental Designs for Research in 1963 represented a major
step in the attempt to provide educational research on a stronger set
of paradigms. The book, written in a very readable style, offered a

\(^1\) Fisher, R. A. Statistical Methods for Research Workers (1st ed.).
Edinburgh: Oliver and Boyd, 1925.
A summary of design characteristics that helped researchers develop more soundly controlled experiments.

In the seventies, a trend that continued from the sixties was the theoretical orientation of many research studies. Piagetian developmental theory was dominant. Early Piagetian work centered mainly on the pre-operational and concrete operational stages. Later studies focused on the formal operational stage.

Teaching methodologies, however, continued to be the basis for the majority of studies. In the early sixties, the popularity of Mager's book, Preparing Instructional Objectives, resulted in the proliferation of behavioral objective use in the schools. As a consequence, many studies in the early seventies compared the effectiveness of behavioral objectives with other teaching strategies.

In the early seventies individualized instruction was also at the forefront. This resulted in a number of studies comparing aspects of individualized instruction with conventional teaching. Both of these areas continue to be researched, although the number of studies concerned with behavioral objectives has decreased dramatically.

Curriculum evaluation reports continued to abound in this period. Many of the curriculum projects developed in the early and mid-sixties were revised at the end of the sixties and therefore new evaluations were reported.

Methodological changes were in the area of psychometrics. Inferential statistics were increasingly used as well as first-time applications of some non-parametric inferential statistics. There was increased emphasis on the careful documentation of procedures used to establish the reliability and validity of instruments.

Research studies in the past decade have become increasingly more multivariated in approach. (This was one of the factors that led Peterson and Carlson to organize A Summary of Research in Science Education--1977 according to the type of research study conducted rather than according to the content studied.) This has resulted in more frequent use of multiple linear regression and multivariate analyses of variance.

The trends in science education research over the past 50 years can be summarized as follows:

1. A movement from descriptive to inferential to multivariate studies.
2. An increase in consideration of psychological models as the basis for research.
3. A tendency to focus on a given area of research for a period of time.
4. An increase in the number of studies per year (an order of magnitude over the 50 years of active research).

5. Increased consideration of the psychometric problems of reliability and validity.

6. Increased consideration of the proper statistical models of analyses.

Purpose and Organization

The purposes of this review are virtually the same as those expressed by Rowe and DeTure: A Summary of Research in Science Education—1973. These are (1) to portray the state of knowledge in science education, (2) to describe any existing trends, (3) to identify areas that need to be researched, and (4) to provide tentative answers to persistent problems, if any seem to emerge from the research.

In order to adequately perform all of the above functions more time was needed by the reviewers than the four months devoted to the task. Even though previous research reviews in this series and others had been used extensively in a science education research course, and the reviewers were very familiar with them, the synthesis of studies with those of the past is a time consuming endeavor and is always tinged with the bias of the reviewer. For this reason, the organization of this review is similar to those of earlier reviews in this series and an index is included. This should enable science education researchers to compare studies in successive volumes and come to their own conclusions about research in their particular areas of interest. Another feature of this year's review that should aid in this is the inclusion of page numbers in the list of references. Perhaps now is the time for science educators to consider conducting meta-reviews as advocated by Glass. By reanalyzing data from a number of related studies in this manner, personal bias would be eliminated from the review and general conclusions reached without gathering additional data.

We hope that the format of this review will not discourage its use by teachers. We recognize that the highest four priorities in science education research, as determined by the membership of the National Association for Research in Science Teaching (Battis et al., 1978) had to do with research in the classroom. This research certainly should be of interest to teachers, and this review is one source of information. However, more effective alternative methods of translating research findings to teachers may be in graduate level science methods courses and through other publications that review specific aspects of science education research. Two sources for the latter are the short and succinct reviews that occur in publications of the National Science Teachers Association, Children and Science, The Science Teacher, The Journal of


Although the organization of this review is similar to those of past years, the reader will note a few changes. Some minor additions and deletions have been made according to the prevalence of the research during 1978. A section of reading related to science education has been added, as well as a section on cognitive style. In the methods section there was an attempt to relate studies according to the topic researched rather than strictly according to whether the subjects were elementary, secondary, or college students. We felt that this gave a better overview of the area. Within these topical areas, studies are organized according to the grade level of the subject, beginning with elementary and progressing to college.

Selection of Studies

The majority of studies in this review were identified by Dr. Stanley Helgeson and his staff at the Ohio State University who supplied us with abstracts of articles from major journals, dissertations, papers presented at meetings, and other ERIC holdings. Several additional reports were added as a result of the authors' scanning of journals not included in the ERIC search. These abstracts were categorized, checked for applicability to science education, and non-repeatability from the previous year's review.

The same research could hypothetically appear three times in a review of this nature. It might first appear as a dissertation, then as a paper at a professional meeting, and finally as an article in a journal. Where an article appeared more than once, the citation given is that of the journal. There are two exceptions to this. First, when the journal article is taken from only a part of the dissertation, both citations are used. Second, the research reported by Johnson and Howe (1978), that received the Journal of Research in Science Teaching Award for 1978, is described again even though a summary appears in the 1977 review by Peterson and Carlson.

The reviews of articles appearing in journals are based on the reviewers' reading of the articles in their entirety. The same is true of papers presented at meetings. Because research from dissertations frequently appears as a journal article at a later date, the decision was made to review only the dissertation abstract. Sorry to say, little improvement in abstract writing has occurred since Rowe and DeTure and Herron et al. listed the deficiencies of the abstracts in both 1973 and 1974 editions of Summary of Research in Science Education. In some cases, the level of the sample (elementary, high school, or college) was not identified; in others, results were not given.

Because we felt that it would be useful to the reader to see the scope of the research being conducted in a given area, an effort was made to include, at least in passing, a brief description of every study reported during the year. Longer reviews were made of those studies deemed more generally applicable to science education research and to
those studies that had a theoretical base. In some cases, an article is included in the review in more than one section. In general, if the research had a theoretical base, it is included in that section in greater detail.

Other Reviews and Summaries

The following reviews of science education research were published during 1978.

A book edited by Rowe (1978) entitled *What Research Says to the Science Teacher, Volume 1*, contained reviews of science education research with classroom applications. This collection, published by the National Science Teachers Association, contains the following:

"Science: A Basic for Language and Reading Development,"
Ruth T. Wellman

"Analyzing the Questioning Behaviors of Science Teachers,"
Glenn McGlathery

"How Teaching Strategies Affect Students: Implications for Teaching Science," James A. Shymansky

"Relating Student Feelings to Achievement in Science,"
Ronald D. Simpson


"Learning Science from Planned Experiences," Fletcher G. Watson

DeCarcer et al. (1978) reviewed science education research studies concerned with the development of formal thinking in adolescents according to Piagetian theory. The article reviewed a selection of studies concerned with the effects of schooling and the effects of training that were reported in 1975.

Preece (1978a) reviewed the research on the organization of scientific concepts in semantic memory. Topics included are ostensive concepts and conceptual structures, maps of cognitive structure, and organization of scientific concepts in semantic memory.

R. K. Fletcher's (1977) review of literature predicting success in college chemistry is an additional review that may be of interest to researchers in the area of chemistry instruction. Recommendations as well as summaries of articles are presented.
LEARNING AND DEVELOPMENT

Piagetian Studies

In a Summary of Research in Science Education—1975, Mallinson stated:

Education is replete with the ascendance and descendance of messiahs who capture the fancies of those who teach. The Twentieth Century has had its share; for example, Thorndike and Connectionism, Dewey and Learning-by-Doing, and Bloom and his Taxonomy of Educational Objectives. There have been, of course, many others but the current star is Piaget (p.10).

In 1978, Piagetian research continued to dominate studies with a theoretical thrust. In fact, the number of studies based on Piaget more than doubled in the time span from 1975 to 1978. The focus of the studies was on the formal operational stage of development rather than on the concrete stage. There was a shift in emphasis, however. Research during 1978 centered more on examining teaching strategies that would enhance concrete students' achievement, and studies of specific schema such as controlling variables and proportional reasoning. This is in contrast to the deluge of earlier studies primarily to determine the proportion of concrete students enrolled in high school or college science courses. The latter type of information, although still being gathered, is set in a larger context such as in the study by Han (1978) who, in analyzing the learnability of science content of a textbook for middle school students in Korea, determined the proportion of students in the concrete, transitional, and formal developmental stages. He found that 60 percent of the textbook objectives were at the formal level whereas only 22 percent of the students were at the formal or transitional-formal developmental stage.

General Piagetian Experimental Studies

Seventeen Piagetian studies reported during 1978 examined methods and procedures that would either have an effect on students' cognitive development or would increase their achievement in science courses. Some of these studies investigated specific interventions in science instruction; others reported modifications of entire courses.

Wollman and Lawson (1978) pretested students in two average seventh grade classrooms. The students were then randomly assigned to two training groups. The active group manipulated physical objects (Cuisenaire rods, geometric shapes, and cylinders of water) whereas the control group used standard textbook procedures among which was the proportionality algorithm of cross multiplication followed by division. Students were posttested immediately and one month later via individual interviews (disks task) and written tests (math quiz, Mr. Tall-Mr. Short, machine problems). The active group was found to score significantly higher than did the control group on two immediate posttests. Although scores changed during the intervening month, delayed posttest scores were still significantly higher for the active group.

In a study of slightly longer duration but of less rigid control, Boulanger (1978a) reported similar results. Using two different approaches
to teach problem solving (manipulative versus verbal) he found that the manipulative treatment produced higher achievement. No interactions between the treatment and proportional reasoning ability were found or between formal capability and proportional reasoning ability. In an attitude survey, however, he found that formal students preferred the verbal approach and concrete students the manipulative.

On the high school level, Goodstein and Howe (1978) examined the effect on chemistry achievement of making high school chemistry instruction more concrete. Four intact chemistry classes studied stoichiometry for a six-week period using various models of atoms. Results indicated that concrete students did not profit from the use of models whereas formal students did.

Another study of chemistry achievement by high school students was done by Cantu and Herron (1978) in which they examined the effect of using pseudoexamples for the acquisition of six concepts (MIB, ideal gas, isomer, metal, acid-base [operational], acid-base [Bronsted]). They postulated that (1) concrete students would perform as well as formal students in attainment of concrete, (2) formal students would perform better on abstract concepts than concrete students, and (3) concrete students would improve their attainment of abstract concepts more than formal students when pseudoexamples were used to reveal critical attributes of the concepts.

Twelve minutes of instruction using pseudoexamples were followed by a short quiz on each topic over a six-week period. Pseudoexamples consisted of illustrations, diagrams, and models that could be used "to focus attention on critical and variable attributes of a concept in the same way that real examples and nonexamples are used to teach the attributes of concrete concepts" (p.136).

Cantu and Herron found support for their second hypothesis, but not for the first or third. Formal students attained concrete concepts better than did concrete students and pseudoexamples were beneficial to both formal and concrete students. However, when a comparison was made between students who used the pseudoexamples and those who did not, there was a significant difference for the concrete students but not for the formal students, thus indicating that the instruction was more beneficial for the concrete students.

On the college level, Raven and Cole (1978) made the learning of nine units of college physiology more concrete by requiring students to form three types of models (graphic, formal, and inferred) after the instruction. Evaluation consisted of an achievement test for each of the modules containing recall, application, and synthesis questions, the Raven Content Comprehension Test (RCCT), the Gates-MacGinitie Reading Tests, Survey F (GMRT) and the Raven Test of Logical Operations (RTLO).

Using a multivariate analysis of variance, the researchers found that for the models group there was no difference on recall and application items, but that they were superior to the control group on synthesis items. They also found that the multiple correlation coefficient for the graphic model scores (the most concrete treatment) was the best predictor on the
achievement tests. Other findings were that RCCT was a better predictor of physiology achievement than were GMRT and RTLO.

Shyers and Cox (1978) designed a study to observe (1) the acquisition of the proportionality concept through training on the operations of group structured tasks, and (2) the transfer, both between isomorphic tasks, and to subgroups or extensions of groups exemplified by the tasks. Subjects were college freshmen enrolled in a remedial mathematics class for educationally and financially disadvantaged students. Concrete tasks used were the balance beam task, the rings and shadow task, the half-balance task, and a wheelbarrow task.

Shyers and Cox found that, using the last three tasks as demonstrations, students acquired the ability to solve proportion problems and this was transferable to the first task. Order of training did not prove to be an important factor in acquisition of the skill, and the skill remained after one month.

A third study on the high school level which made instruction more concrete was reported by Schneider (1978). He randomly assigned one high school ninth grade physical science class to concrete instruction, whereas another was taught with lectures, films, demonstrations, etc. The concrete section used materials to gather data about an unidentified concept they were to learn. From the description given in his abstract it appears that the general Piagetian-type learning cycle of exploration, invention, and discovery (application) was followed. Schneider found that the concrete instruction group showed significantly greater gains in intellectual development at the end of the instruction but that there were no differences three months later.

The above seven studies lead to the general conclusion that making instruction concrete tends to aid achievement and the acquisition of specific schema. Three of these studies (Wollman and Lawson, 1978; Shyers and Cox, 1978; and Boulanger, 1978a) were directly concerned with improving students' proportional reasoning ability. All three reported positive, immediate results and the former two, which tested for retention, found the improved ability lasted for at least one month. Schneider's study, on the other hand, showed no retention of general intellectual development after a three-month period.

Studies more related to specific science content by Goodstein and Howe (1978) and Cantu and Herron (1978) report conflicting outcomes. Both of these studies were directly concerned with how concrete students could be aided in the acquisition of formal chemistry concepts. Goodstein and Howe found that the use of physical models was not helpful whereas Cantu and Herron found that pseudoexamples were. Different outcomes may be due to the fact that the formal concepts in the Cantu and Herron study were carefully delineated in a programmed instruction format and were learned by volunteers in a short interval of time. These conditions would lead one to assume that all students paid attention to the instruction. In the Goodstein and Howe study, the concepts tested were much more complex. Stoichiometry problems require students not only to understand single concepts but to relate them to one another. The regular classroom setting is also not as conducive in making students...
pay attention to the instructional treatment. In the latter study there was a much greater possibility that no significant differences would be noted.

An effective method of making students pay attention to more concrete modes of instruction is to make students create their own graphic representation of the concept to be learned. Raven and Cole (1978) found this improved achievement in the study of physiology.

Another study, although not concerned with making instruction more concrete, did require pupils to pay attention to the instruction. Johnson and Howe (1978), in their Journal of Research in Science Teaching award-winning article, investigated the effectiveness of two conflict-producing procedures on the development of logical thinking of fifth graders. Students who were classified as conservers or nonconservers using four Piagetian conservation of area tests, formed five 14-membered groups, three of nonconservers and two of conservers. Students in one nonconserving group received individual conflict training by an experienced teacher whereas students in another nonconserving group were paired with conservers for peer interaction conflict training. The remaining two groups acted as controls.

Johnson and Howe found that conservers were not persuaded by nonconservers to deviate substantially from their conserving viewpoints and that, although both treatments produced significantly higher scores on the immediate posttest, only the peer interaction conflict training had long-term results (one month). Similar results were found by Fagal (1978) who investigated the effectiveness of students working alone, with someone of the same Piagetian cognitive developmental level, or with someone of a higher level on science achievement. She found that early concrete students achieved higher when working with someone of a higher level than when working alone. Late concrete operational students achieved highest when working alone. For formal students it made no difference.

Other science educators have studied the effect of modifying a whole course according to Piagetian principles to determine if achievement is enhanced or if there is a change in developmental level. J. A. Henry (1978) and Calvey (1978) modified and examined elementary science programs.

The major purpose of Henry's study was to determine if classification and discrimination activities in a science program for six-year-old children could affect the transition from pre-operational to concrete-operational thinking. If so, would there be a differential gain associated with number, substance, and length versus weight and area?

Classification and discrimination hands-on activities were introduced into the regular science program for first graders for a 15-week period. The control group received regular instruction. Findings showed that the experimental group exhibited significantly greater transition than did the control group, and that the gain was greater for tasks involving number, substance, and length than for weight and area.
Calvey (1978) reported that eighth-grade students using *Science--A Process Approach* had higher scores on achievement tests of Piaget's operative comprehension than did students in a traditional program. No differences were found for sixth-grade students, however.

On the college level, several researchers modified courses to make science instruction more comprehensible or to modify students' developmental levels. M. H. Baker (1978) made chemistry instruction more concrete through the use of desktop kits. The kits contained chemicals, model building materials, and demonstration devices which each student individually manipulated during lectures. Two classes used the kits and two classes were used as controls. All students were classified as concrete or formal, using three Piagetian tasks, and were given a comprehensive pre- and posttest of achievement. Growth quotients (posttest-pretest/perfect score pretest) were calculated.

A significant difference between the growth quotients and achievement of the entire experimental and control groups as well as between the concrete members of the experimental and control groups was reported. Among the formal members of the experimental and control groups, there was significant difference in growth quotients but not in achievement. (Results were similar to those obtained by Cantu and Herron.)

Mele (1978) studied the effectiveness of a problem-solving program in biology on the transition of students from concrete to formal operational thought. Students in the experimental group who used Kaplan's Problem-Solving Biological Curriculum scored significantly higher than did the control group on Piagetian concrete-formal operational thought, abstract reasoning, objective-subjective differential reasoning and critical thinking.

Other methods of modifying instruction according to Piagetian principles were not as effective. T. C. Campbell (1978) used the "Piagetian learning cycle" of exploration, invention, and application with physics students in a Personalized System of Instruction (PSI) format. He found that all students made significant improvement in the use of more formalistic reasoning abilities as measured by Written Puzzles and Written/Experimental Tasks. However, no significant differences in the learning of physics content were present. The learning cycle did aid in the earlier development of social skills and a more positive attitude in the laboratory.

Two other PSI-Piagetian studies on the college level were reported. Bunck (1978) developed a complete set of laboratory activities for chemistry based on Piagetian learning theory. Although no comparisons were made between groups, he reported a high degree of student satisfaction with the approach. On the other hand, Hardin (1978) tested the effects of PSI versus a traditional lecture course in college physics on students' cognitive reasoning ability. Using gain score comparisons, he found that neither PSI nor lecture course was effective in developing the logical reasoning or critical thinking ability of students. Both were equally effective in fostering physics achievement.
Regular physics courses apparently do not have an effect on developing logical reasoning either. Barnes and Barnes (1978) reported that no developmental progress was found as a result of students taking one introductory physics course for one semester.

What can be concluded about modifying entire courses according to Piagetian principles to bring about increased achievement or changes in developmental levels? Results are mixed. Some success was reported on the elementary level by Henry and Calvey. On the college level, two specific programs (desktop kits by Baker and a problem-solving program by Mele) had beneficial results. More general approaches such as PSI (Hardin) and PSI combined with exploration, invention, and application (Campbell) were not effective. This is not to say that these latter programs had no value, because positive attitudes were generally reported. It appears that more specific programs have the greatest probability of producing positive results.

Particular Schema

In examining Piagetian studies published during 1978, it became evident that there were two particular schema about which many studies centered. For this reason, studies concerned with proportionality and controlling variables are described separately whereas others are grouped together.

Proportionality

Six studies were directly related to problem solving and students' proportional reasoning ability. McBride (1978) found a positive relationship existed between proportional reasoning and achievement of selected mathematics and science concepts. He tested 136 ninth-grade physical science students and found that students classified as formal with respect to proportional reasoning had a significantly greater knowledge of science concepts involving simple machines and the structure of matter than had concrete operational students.

Preece's (1978b) report of a study with fourth-form physics students indicated similar findings. Through the use of word association tests and proportionality tests, he concluded that "the correlation reported above might be interpreted as implying that the lack of formation of the schema of proportionality imposes a constraint on the development of mechanics cognitive structure" (p. 398).

Wollman and Lawson found that when seventh-grade students manipulated physical objects they were more successful than when standard textbook procedures were used. Differences persisted one month later. Boulanger's manipulative strategy was not more successful than a pictorial approach with eighth-grade students in promoting proportional reasoning ability. Shyers and Cox (1978) found that disadvantaged college students could acquire the ability to solve proportionality problems and that there was transfer to another task. Again, results are mixed and appear to be highly dependent on the type of concrete experiences offered to the students.
Two studies on proportional reasoning took a close look at the finer structure of the schema and students' acquisition thereof: the one by Shyers and Cox discussed previously, and one by Lybeck (1978). Lybeck, in a study of 16- and 17-year-old Swedish students, recorded procedures students used in solving problems. Among his conclusions were: (1) the conceptual difficulties faced by students were contextual rather than structural; that is, a correct answer on one task does not imply a correct one on another; (2) if students operate according to the proportionality mathematical formula, \( y = h x \), this hinders or at least does not lead to concept formation; and (3) students need to study the content qualitatively before equations are introduced.

**Controlling Variables**

Of the five studies in this area directly concerned with controlling variables, two were reported by Wollman (1977a, b). Both pertain to the finer structure of the schema and the acquisition thereof. A colliding spheres type of experiment was used by Wollman (1977a) with several hundred students in grades 4–12 to determine whether the concept of controlling variables was developed over a long span of time. Results indicated that the concept develops gradually and is not fully attained until around 14 years of age. The study also shows that students of lower socio-economic status are slower to acquire the concept.

In another article, Wollman (1977b) described the method he used to determine that there were five levels between concrete and formal operation in the development of the concept of controlling variables. "The question sequence used in determining the levels may be helpful for designing sequential learning experiences and assessment instruments consonant with the course of intellectual development as seen from a Piagetian viewpoint" (p. 390).

Linn and Levine (1978) examined the influence of the question format and the type of problems used on 11- to 17-year-olds' ability to control variables. The three problems studied were the Ramp Problem, the Circuit Problem, and the Seed Problem. Formats considered were free response questions, multiple-choice questions, and screen questions. Results of the study showed that students did not perform consistently on all three tasks using three different formats. Multiple-choice items were consistently easier than were free response and screen questions. The difference between the screen questions and free response was dependent on the task involved.

Two other studies involving controlling variables investigated whether students could learn to control variables in a free-choice environment. Bowyer et al. (1978) found that sixth-grade children can learn to control variables from doing activities that involve free experimentation using activity cards. Rice and Linn (1978) reported that learning to control variables via activity cards was enhanced by a training program given simultaneously. A more complete description of these two studies is included in the section on Individualized Instruction.
Other Schema

A study by Lawson et al. (1978) examined the acquisition of propositional logic and its relationship to the formal schemata during the secondary school year. Using a broad sample of students from elementary, junior high, senior high, and college they administered a seven-item pencil-and-paper test that required use of the formal operational schemata of proportions, probability, and correlations (5 items) and the rules of propositional logic (2 items). Results showed that while the percentages of students who answered the items requiring formal operations generally increased with age, there was no such consistent improvement with age for the propositional logic items. This suggests that there are at least two rather distinct developmental trends rather than one as postulated by Inhelder and Piaget. A principal component analysis also revealed two loading patterns with the formal schemata questions forming a cluster quite distinct from the propositional logic questions. Although these differences might have resulted from differences in difficulty level of the test items rather than from a difference in psychological requirements, the researchers probed further and found this unlikely.

Adi et al. (1978) examined how high school students approach tasks that require correlational reasoning for successful solution. Two tasks (one causal and one coincidental) were administered to 80 secondary students in the San Francisco area. Responses were categorized according to whether students saw a relationship, how much information was used, and whether there was a quantitative correlation. They found that twelfth graders performed at a higher level than did ninth graders on both tasks but that a significant number of students did not use correlational reasoning.

Several conclusions can be made about the research reports reviewed in this section on particular schemata. Many more research articles were devoted to a particular developmental schema. In the summary of Piagetian research relevant for high school science teaching by DeCarcier et al. (1978) that covered a period from 1971-1975, only two major studies were concerned with proportional reasoning, and four studies with controlling variables. Most studies included in the review were training studies; this review contains six proportional reasoning studies and five controlling variables studies.

Comparing studies conducted today with those that were appearing in the early seventies is like comparing a Model T with a Citation. Studies focus on the fine detail of the structure of the schema, rather than on the age at which the schema are acquired. More sophisticated statistical methods such as multivariate analyses of variance are becoming commonplace and greater care is being taken to test for retention and transfer.

Instruments and Measurement

Ten research reports were directly related to instrument production and methodological problems in assessing students' cognitive developmental
level. Several of these were concerned with Piagetian tasks; others describe the validation of written Piagetian tests.

Linn (1977) analyzed two tasks first used by Inhelder and Piaget, and two developed at Lawrence Hall of Science, to determine what task characteristics influenced performance, and to consider what criteria were used to assign scores to subjects. She found that in the Pendulum Task, the amount of interaction with the apparatus and the characteristics of the variables affected the results. In both the Ramp Problem and Pendulum Task, subjects set up controlled experiments, but these were sometimes for the wrong reason or they drew erroneous conclusions. Sometimes reducing information available to the subject at a given time maximized reasoning skill. Linn concluded, "educators need to be aware that concrete apparatus used for explaining scientific concepts can sometimes simplify, but could also confuse..." (p. 367).

Bady (1978) investigated whether one task is an adequate and reliable measure of a schema. Sixty-six subjects were administered five proportionality tests and five combinatorial analysis tests. In each of the categories two of the tests were paper and pencil tasks. Point biserial correlations were calculated between the score on each task and the total score for that schema. The Shadows Test had the highest correlation for the proportionality schema and travel routes for the combinatorial analysis schema. From the moderate intertask reliabilities, Bady concluded that success on a single task was not a reliable indicator that the student possessed the schema or would perform successfully on another task purported to measure the same schema. He suggested that more research be done to determine why tasks differ in difficulty and why this varies from person to person.

Going one step further, Cohen (1978) investigated whether a hierarchical relationship existed among six Piagetian tasks designed to examine associated topological groupings as suggested by Piaget. He administered seven Piagetian tasks to 100 preschool, kindergarten, and second-grade children. He found that although the tasks did scale, the grouping was different from that reported by Inhelder and Piaget. He also reported that the number of children passing each task at each grade level was low and suggested that this might be rectified if children had more opportunities to manipulate concrete objects.

Four studies were directly concerned with methodological issues associated with test content and format. Linn and Levine (1978) reported differences in scores according to whether the question format was free response, multiple-choice, or a screen question (see Controlling Variables). Barnes (1978) investigated whether or not the familiarity of content was a factor on how well students scored on Piagetian-type tasks, and whether there were differences between education majors and engineering physics majors. He found that there were differences between physics and education classes, between biological and physical science majors, and between science majors and nonscience majors. Differences according to familiarity of content were also noted in several instances.
Van Harlingen (1978) suggested that "instrument noise" could obliterate the determination of students' cognitive levels. He found that a change in the character or nature of a task variable may result in misclassification of a student's cognitive level, and that it is difficult to predict which changes make a difference. For example, in Mr. Tall-Mr. Short, a change of units or a change from linear to volume units made no difference in performance, but a change from integers to nonintegers did. He called for the determination of a more detailed hierarchical conceptualization of formal thought if accurate and appropriate instruments are to be developed.

Gates and Jay (1978) studied children's conception of "all" and "some," two terms frequently used in Piagetian interviews. A summary of their conclusions is given in the section on concepts and content.

Three studies were directly concerned with instrument development. Renner et al. (1978) reported on the development of a paper and pencil test to measure displacement volume to be used to identify the beginning of formal thought. Seventh- through twelfth-grade students (N = 586) were administered three Piagetian tasks and the written test. From examining the percentage of students who were successful on the displacement volume tasks they concluded that this task was appropriate for measuring the early formal stage. They also found that in 94 percent of the cases, the written test and the individual interview results agreed.

Lawson (1978) developed a more comprehensive written test to measure concrete and formal reasoning in junior and senior high school students that could be administered in a relatively short period of time (75-100 minutes). The test included items on conservation of weight, displaced volume, proportional reasoning (4 items), controlling variables (4 items), combinatorial analysis (2 items), and probability (3 items). Each item involved a demonstration by the teacher using some physical materials and/or apparatus which posed a question. Students responded in a test booklet by checking a box next to the answer and then explaining their choice.

The test was validated in part by administering it to 513 eighth-, ninth-, and tenth-grade students and correlating their scores with those obtained from two Piagetian tasks (bending rods and the balance beam). Scores on the test ranged from 0 to 15. The mean was 7.41; standard deviation, 4.27; and standard error of measurement, 2.0. The KR-20 estimate of reliability was .78. Validity was assessed by a panel of judges and Pearson product-components moment correlations were .75 and .65. A principal components analysis showed that three factors accounted for 66 percent of the total variance, and that these factors corresponded to similar factors obtained from interview scores.

Lawson's classroom test was designed to allow teachers and/or researchers to classify student performance according to developmental levels. If a student scores 0-5, he is classified as concrete; 6-11, as transitional; and 12-15, as formal.
Another written test of Piagetian tasks was developed by Stever and Gabel (1978). The purpose of developing this test was identical with that of Lawson. The test, which can be administered in a 55-minute period, differs from Lawson's in that instead of the teacher demonstrating Piagetian tasks, these are demonstrated via videotapes. In this test students do not write down the reason why they selected a particular answer but select a reason from a number of possible reasons that is given to them. The test contains items based on four Piagetian tasks (conservation of volume, separation and control of variables, combinatorial analysis, and propositional thinking). Validation of the test was made using 84 students in grades 10-12 to whom the test was administered and individual task interviews made. The alpha coefficient was found to be 0.85. No breakdown of scores according to particular Piagetian stages of concrete, transitional or formal was given.

In addition to the above nine studies, two others are included briefly here because they give some information about measurement although this was not their primary thrust. Braun (1978) studied three factors that might influence performance on Piagetian tasks. He found that training in science, cognitive ability, and cognitive style had little effect on performance. He also found that the discrimination learning task he used was an inappropriate measure of formal operational ability. Data supported a relationship between the discrimination learning task and the level of concrete operations.

Lowell (1978), in a study of hierarchical classification and concrete and formal thought, found that his results supported the assumption that students capable of formal operations would have little difficulty on the hierarchical classification test. The assumption that concrete students would experience difficulty was not supported.

Studies reviewed in this section point out the difficulty and complexity determining a student's developmental level. We can conclude that one task is probably insufficient to measure a student's level (Bady) although if the task is selected carefully, there is a high degree of probability that the student will be properly classified (Renner et al.). There are also many inherent problems in conducting interviews. Such factors as the familiarity of the task (Barnes), the format of the questions (Linn and Levine), and the value of the variables (Van Harlingen) make a difference in results on both interviews and paper and pencil tests.

Two written tests that can be used to determine the cognitive level of large groups of students have been supported by Lawson, and Stever and Gabel. These tests should be useful to researchers who are interested in studying the interaction of different methods of instruction and Piagetian level.

**Student Characteristics**

What is the relationship between student characteristics such as age, sex, cognitive style and Piagetian developmental level? Most of
the studies discussed in this section were correlational studies whose authors sought to determine if Piagetian level or success on certain tasks were correlated with other factors in the student's background.

Previous research has established that intellectual development increases gradually with age or year enrolled in school. In 1978, research reports confirmed this finding for young children (H. G. Cohen, 1978), for elementary through high school students (Wollman, 1977), and for junior and senior high students (Adi et al., 1978; Linn and Levine, 1978).

Researchers who looked at sex differences were consistent in their findings. No significant sex differences were reported for young children by H. G. Cohen (1978) in his study of topological Piagetian groupings, by Gross (1978) for twelfth-grade students on a series of tests and tasks (although males did outperform females on the more manipulative balance task), and by Barnes (1978) for college students on six Piagetian tasks.

Several researchers examined the relationship between the kind and number of courses students took and their developmental level. Cohen et al. (1978) found that success on three Piagetian tasks was weakly correlated with the level of the physics course in which the student enrolled whereas there was no significant correlation with the grade the student received. Lloyd (1978) found, however, that students' cognitive level was a good predictor of success (final grade) in a college chemistry course. Working with fifth-grade students, Brown et al. (1977) found that a 30-item Piagetian test could be used to predict their letter grade in science.

Other investigators examined test scores rather than course grades. Gross (1978) found that science students had a distinct advantage over non-science students on Piagetian-like tests (except for one non-science, math verbal analogy test). Unruh (1978) found correlations of Piagetian task scores and test scores only in physical science courses. He found no significant correlations between test scores and astronomy, biology, chemistry or physics achievement. He did find that there was a high correlation (r = 0.90) between task scores and the number of math courses the student took. Cobb (1978) found no significant correlation between Piagetian task scores and biology achievement.

On the other hand, McBride (1978) found that there was a significant relationship between proportional reasoning ability and the achievement of selected concepts. Basmajian (1978) found that there were statistical significant differences between formal and nonformal operational students as to their gain in knowledge of subject matter, critical thinking ability, and attainment of laboratory skills in a college biology course. Others who found formal students learned more than concrete were Fagal (1978), Cantu and Herron (1978), and Boulanger (1978a).

A more detailed look at the relationship between specific science content and Piagetian level was made by Raven and Murray (1978). They
examined the effect of high school chemistry experiences on Piaget's operative comprehension. They found that students who had taken chemistry had made higher gains on the Raven Content Comprehension Test (RCCT). Some care should be taken in interpreting these results because pretests were not given. Raven and Murray recommended that a structure analysis of the conceptual framework of our present science curricula be undertaken to provide information on ways that logical operations currently are being used with science content, and how these may be more extensively implemented.

Raven and Adrian (1978) investigated the relationship between scores obtained on the Self-Concept of Ability Scale, the Sequential Test of Educational Progress, the Standard Test of Academic Skills and Piaget's operative comprehension as measured by RCCT. Results showed that there was a positive relationship between reading achievement, science and mathematics achievement, operative comprehension of reading, general self-concept of ability and self-concept of ability in mathematics, English, social studies, and science.

In 1978, two studies investigated the relationship between Piagetian levels of cognitive development and cerebral hemispheric dominance. Unruh (1978), studying college students enrolled in astronomy, biology, chemistry, physics, and physical science, found no significant relationship between the levels of development and cerebral hemispheric dominance as measured by the Torque Test. Developmental levels were measured by five Piagetian tasks as well as by a logical reasoning test. The verbal analysis subscore of this test was the only one that correlated with cerebral hemispheric dominance. Subjects tending toward right hemispheric dominance scored significantly higher (p > .01). The most positive correlation reported in the study was between the number of mathematics courses taken and Piagetian level (r = 0.90).

The findings of Goodrum (1978) with eighth-, ninth-, and tenth-grade students in Australia concur with Unruh's results. Goodrum found no positive relationship between Piagetian development as measured by Burney's test and cerebral dominance as measured by the Torque Test. However, he did find that there was a statistical significant relationship between hemispheric thinking (favoring the left side) with Piagetian development. No relationship between creativity and cerebral dominance on Piagetian development was found.

Two cross-cultural studies examined student characteristics and relationship to cognitive development. Kishta and Mays (1978) compared the levels of cognitive and moral development of second- and fifth-grade students in Jordan and Iowa. Cognitive tasks measured conservation, class inclusion, multiplicative classification, and spatial perception. The moral judgment tasks presented situations related to reciprocity, values, and perspective of viewpoints. Kishta and May found that both cognitive and moral development of students differed according to country and grade level. Results also indicated some support for a degree of parallelism between moral and cognitive development as reported by Piaget and Kohlberg.
Brown et al. (1977) studied the relationship of the acquisition of Piagetian concepts with surname, language, science letter grade, and science test scores of Mexican-American and Anglo-American fifth-graders. They found that the acquisition of concepts were predictors of all of these. Mexican-American children scored significantly lower on both the Piagetian concept test and a science achievement test.

An area of increasing interest and importance during 1978 was that of cognitive style. Four studies that relate cognitive style and Piagetian theory were reported. All used the Group Embedded Figures Test to determine whether students were primarily field dependent or field independent. Because we have included a section of this year's review on cognitive style, the four studies (Lawson and Wollman, 1978; Braun, 1978; Lloyd, 1978; and Cantu, 1978) will be discussed in detail in that section.

In this section we have found that researchers' studies generally confirm the Piagetian theory that cognitive growth develops over time (Cohen, Wollman, Adi et al. and Linn and Levine) and that formal students achieve more than concrete students (Lloyd, Brown et al., McBride, Basmajian, Fagan, Cantu and Herron, and Boulander).

Researchers also found no differences in Piagetian levels according to sex although males may perform better on specific, highly manipulative tasks. Success in a science course as measured by either high grades or test achievement has not always correlated positively with a student's developmental level. If this question is important, more research must be done.

Two areas of importance that are relatively new on the scene are those of relationship between students' cognitive developmental level, their cerebral dominance and their cognitive style. Many more studies in these two areas need to be conducted in order to make generalizations with wide applicability.

**Ausubelian Studies**

While the number of Piagetian studies continues to increase, the number of studies based on the learning theory of Ausubel has been fairly consistent over the past five years. During 1978, five studies were reported. Two of these were concerned with the effect of prior knowledge on subsequent concept learning, one on the ordering of concepts within a course, and two on the effectiveness of advance organizers on achievement.

Pines (1978) used 126 first-grade pupils to determine whether their relevant existing cognitive structure was an important variable affecting learning. Relevant existing cognitive structures were ascertained before and after 24 taped science lessons. Pines found that the prior knowledge did have an effect on subsequent concept learning. The most compelling evidence came from the divergence of initially different cognitive structures subsequent to an identical instructional sequence where convergence might have been expected.
Wesney (1978) examined the influence of prior cognitive development in physics and in mathematical reasoning on concept attainment in the study of mechanics. The cognitive backgrounds of 300 students were measured using the Dunning-Abeles Physics Test, Form E and an author-made mathematics test. Results of the study showed that students with more fully developed backgrounds of physics content and mathematical reasoning skills demonstrated better overall concept attainment.

Moreira (1978) modified the content of a physics course according to Ausubelian theory. Using Ausbel's principle of progressive differentiation (i.e., the most general and inclusive ideas are presented first and are then progressively differentiated in the terms of detail and specificity), he arranged topics in electromagnetism accordingly. This course sequence was used in two instructional modes, self-paced and the conventional lecture. Achievement of students in each mode using the Ausubelian structure format was compared with that of students using a traditional topic format. No significant differences in achievement were found in terms of measures such as quizzes and exams. However, in terms of concept learning, there was evidence that the Ausubelian approach fostered concept differentiation, relatedness, and meaningful hierarchical organization to a greater extent than did the traditional approach. This was especially true for the self-paced program comparisons.

Studies comparing the effectiveness of advance organizers were done by Varano (1978) and Kahle (1978). Both studies involved high school biology students.

Varano examined the effectiveness of advance organizers versus behavioral objectives. One hundred three tenth-grade students participated in the experiment for a two-week period. Although students receiving the advance organizer treatment had slightly higher scores, there were no significant differences observed between groups on immediate and delayed (eight weeks) posttests.

Kahle's study was similar in nature to Varano's. Her experiment involved disadvantaged students who studied three audio-tutorial units on genetics over a six-week period. The experimental group studied mitosis using behavioral objectives, meiosis using an advance organizer, and chromosome abnormalities using both an advance organizer and behavioral objectives. The control group studied all three topics using an historical review. Results showed that there were significant differences in achievement for the experimental group on the two units that were studied using advance organizers. Although an immediate posttest over all three units showed no significant differences between the experimental and control groups, there was a significant difference favoring the experimental group on a retention test administered three weeks later.

The above group of studies shows general support for Ausubel's theory of meaningful verbal learning in concept learning. In particular, support is given for the significance of prior cognitive structure on learning and for the usefulness of advance organizers.
TEACHING-LEARNING STRATEGIES

A continuing interest by science educators in investigating methods that might improve science achievement is reflected by the large number of such studies (88) reported during 1978. Although the effectiveness of specific methods of individualizing instruction remains in the forefront, an area of growing prominence is that of science achievement and language arts. Nineteen studies were reported during the period on this area.

Individualized Instruction

In this section are grouped together a large number of studies which are characterized by one main feature—a planned part of the instruction is that all students in a given classroom are not doing exactly the same thing at the same time. The four most common variables in programs that are individualized are: (1) the person (teacher or student) who makes the decisions about learning, (2) the content to be studied, (3) the time spent in learning the content, and (4) the method used to learn the material. In some individualized programs only one factor, such as time, is varied in a self-paced program. In others, combinations of the above are varied. If a research study on individualizing instruction is to be replicable, the report should clearly state which factors vary, how they vary, and how others are controlled. This review will begin with the effectiveness of very structured programs in which the primary variable is time and will conclude with findings related to more open approaches.

Audio-Tutorial Methods

The audio-tutorial (AT) approach to instruction has been reported to be effective in the past. Six studies in the period examined its effectiveness for teaching science to college students.

Hayden (1978) studied the usefulness of AT instruction in a biology course for non-science majors. Half of the students studied selected BCS Minicourses using AT instruction for 17 weeks whereas the remainder studied under a traditional format. Results indicated that there were no significant differences in achievement as measured by the Nelson Biology Test or in critical thinking as measured by the Watson-Glaser Critical Thinking Appraisal. She did find that students using the AT instruction had more positive attitudes than did those in the control group.

Cobb's (1978) findings were similar. She compared the effectiveness of minicourse instruction in college biology for nonmajors using an AT approach with conventional instruction for two semesters. Dependent measures were achievement, cognitive development (Piagetian tasks), attitude towards biology, and personality characteristics. She found no significant effect on achievement, no consistent findings for intellectual development over both semesters, and no related personality characteristics. Attitudes for the AT group were significantly higher for the first semester only.
Vinson (1978) studied the effectiveness of AT instruction for a physical science course over a two-semester period. He compared a traditional lecture approach with an individualized approach incorporating features of both Postlethwait's AT approach and Keller's PSI. Although he found no significant differences in achievement, students in the AT groups performed significantly higher on transfer tests and on a problem-solving test. No differences in attitudes toward physical science were found.

In a shorter study on one specific topic (genetics), J. F. Ross (1978) found AT instruction effective. His study was undertaken to compare the effect of AT laboratory instruction with that of traditional instruction for a three-week period in a nonmajor biology course. He found that there was a significant difference in achievement favoring students in the AT group. He also found that AT instruction was particularly effective for students who had "low entering competency" and for male students. Although he did not compare attitudes in both groups, he reported that the AT group had positive attitudes toward the method of instruction.

A study by Hertzler (1978) is included in this section, even though it is a pilot study, because it calls our attention to the possibility of improving conventional AT lessons. He compared the effectiveness of TAPE TUTOR Assisted Instruction with that of the letters format for two one-hour periods of instruction. The TAPE TUTOR (TT) goes one step beyond the conventional AT in that it asks students questions and allows responses. (It appears to be a hybrid of AT and computer-assisted instruction, PLATO, but much less expensive than the latter.) Hertzler found that for one of the topics TT was superior to the lecture, in that scores on a posttest were significantly higher. For the other lecture, there was no significant difference in scores, however, the TT students had a lower mean grade point average to begin with. Future plans call for comparing the effectiveness of TT with AT instruction.

Romero (1977) also found that AT instruction was effective, but in a slightly different context. He compared the use of AT instruction with the lecture discussion methods for teaching science process skills to prescience elementary teachers of three different majors. He found that subjects experiencing the AT instruction achieved significantly higher than did those in the control group. He also found that science majors' performances surpassed those of social studies and humanities students. This latter outcome should have been anticipated because science students generally enter science courses with more skills than do other students.

Learning Activity Packages

A somewhat more flexible method of individualizing instruction is the learning activity package. Generally there is a higher degree of student choice of activities and a greater latitude in time. Three studies examined their effectiveness.
On the high school level, Simes (1978) studied whether learning packages improve the study habits and attitudes of biology students. The self-instructional learning units contained optional learning activities, allowed self-pacing, and provided an opportunity for mastery. No mention is made in the abstract of the duration of the treatment. The findings did not support his major hypothesis that individualized instruction has a more positive effect than a traditional approach on the study habits and attitudes of students. He found that students of average mental ability obtained significantly higher scores on study habits when learning in the traditional manner. The traditional approach also produced higher attitudes for students of average family status.

Two studies on the college level reported on the usefulness of learning activity packages. Oszogomonyan (1978) developed and tested the effectiveness of an individualized package to teach stoichiometry to underprepared college chemistry students. Design criteria of the package included teaching visualization of molecules, using proportional reasoning rather than the unit factor label method, flexibly deploying instruction, and establishing a completion deadline. Although the details of the sample and the treatment given the control group are not found in the abstract, findings indicated that the experimental students performed significantly better on a posttest and on a retention test.

McCall (1978) compared the effectiveness of learning activity packages with lectures on student achievement and attitude in a physical geography course over two three-week periods. He concluded that students learned significantly more cognitive content when taught by the learning activity packages than when taught by traditional classroom lectures. Students also had a higher level of satisfaction with the package approach for both phases.

Haugen (1978) investigated how students used their time during self-paced instruction in small (9 students) and large (24 students) classrooms. Work-related activities and some forms of misbehavior occurred with similar frequency in both classes. No major differences were reported.

Personalized System of Instruction

A method of individualizing an entire course through use of objectives, learning materials, mastery learning, self-pacing, tutors and frequent tests was devised by Keller and is known as the Personalized System of Instruction (PSI). During the period, nine reports of curriculum development and/or instruction using PSI appeared in the literature. All pertained to the teaching of college science courses. Six described using PSI for physics instruction.

Four studies compared the effectiveness of PSI versus the traditional lecture-recitation-laboratory approach. Lagendijk et al. (1978) found that there was a significant difference in achievement between students in PSI classes and those in conventional laboratory courses.
They also found that students in the PSI classes were able to achieve higher scores in a lesser amount of instructional time. However, the number of hours instructors spend on individual students was considerably higher for PSI students. Balfour (1978) also reported that in a physics course for pharmacists, self-paced students scored significantly higher than did those enrolled in the traditional course.

T. C. Campbell's (1978) findings differed from those reported by Lagendijk et al. and by Balfour. He found no difference in physics achievement between PSI and conventional course students. He did find that PSI students had a more positive attitude toward the course and that they were not as likely to withdraw from the class as were conventional students.

Hardin's (1978) results were similar. He found no differences in physics achievement, in the development of logical reasoning, or in critical thinking ability. The only difference found was that students had a higher perception of achievement than had lecture students.

Hedges (1978) investigated the long-term effects of PSI. He compared grades in three courses taken by students who had three introductory physics courses using PSI with those of students taking the three courses taught by conventional methods. In a tightly controlled study, he found that students who had the PSI instruction in the three basic courses had significantly higher grades in two chemistry courses and a biology course taken later. Because this increase cannot be attributed to learning physics more thoroughly through the PSI instruction, the differences in achievement were attributed to the development of better study habits by PSI students.

McFarland et al. (1978) provided a detailed description of a PSI biophysics course for life science students. Although they did not provide a comparison between PSI and conventional instruction, the study is worth noting because they did compare the amount of instructional time per student in conventional and in PSI classes. They found that there was little difference in the amount of instructional time spent and that PSI could actually reduce time if large numbers of students were enrolled in the course.

In the five PSI studies described above, the entire course was taught by this method. Newhall (1978) described a course in which PSI instruction was used for the last one-third of each semester in a two-semester course. Students who were free to select from among a number of course topics had a very positive response to the course.

Two reports were made of the use of PSI in biology instruction. R. G. Ross (1978) sought to determine predictor variables for the two instructional modes. The three most important predictors for both groups were reading ability, maturation, and past academic achievement. Grabe et al. (1978) compared the effectiveness of the Phase Achievement System (PAS) (similar to PSI in format and objectives) and traditional lecture instruction with respect to affective consequences. They found
that students in the PAS classes had more positive attitudes toward their course than did other students. They were particularly favorable to the freedom PAS offered and to the test and grading system.

Only one report of PSI in teaching chemistry was identified. Bunck (1978) reported on the development of a complete set of laboratory activities incorporating Piagetian learning theory in a PSI format. Student response was very favorable to the mode of instruction.

Prescribed Review Activities

One method of individualizing instruction is to prescribe review activities for students who need extra practice or help. Two such programs were reported.

Palma (1978) sought to determine whether the use of individually prescribed review activities, selected on the basis of pretest results, would result in significant gains in achievement on a posttest for community junior college astronomy students. High risk students as determined by the pretest were randomly assigned to a control or an experimental group. After a one-week review period, the experimental group scored significantly higher than did the control group.

In chemistry, Martin and Carlton (1978) devised a help program for students who were failing chemistry. They reported that five hours per week of remedial work during the final six weeks of the semester led to higher final marks than those for a group not receiving remedial aid.

Mastery Curriculum

Saber and Kaplan (1978) reported on a new biology curriculum for use in Israel based on mastery learning. The curriculum incorporated strategies designed to increase maturation, interest, confidence, and comprehension of science content. Biology concepts were taught in a five-lesson sequence that employed an introduction, story, concept list, laboratory activities, programmed lessons and enrichment activities.

Comparison of the achievement of students using the new curriculum with that of those in a conventional program indicated that more students reached the mastery level with the new curriculum. Attitudes toward biology were also significantly higher for students enrolled in the new program.

Free-Choice Activities

One of the components of many individualized programs is that students are encouraged to determine the time spent on an activity and/or the activity itself. Three studies examined student behavior in a free-choice environment.
Bowyer et al. (1978) examined how children operate in a free-choice environment and determined whether or not they could learn to control variables. Sixth-grade students had the opportunity of selecting science activities for two 45-minute periods per week over a 12-week span. The directions for activities were provided on activity sheets that suggested an interesting activity, described a solution, and provided the learner with related challenges to solve. Performance on controlling variables was compared with that by students who did not have the science activities. Results indicated that sixth-grade, middle class children can work constructively, efficiently, and with enthusiasm over an extended period of time in an environment which encourages student autonomy. Children can make decisions concerning choice of activity, organization of work time and degree of cognitive involvement with materials. In addition to being able to operate effectively in a free-choice environment, there is evidence that children can learn about controlling variables... (p. 106).

In a study utilizing the same type of activity cards, Rice and Linn (1978) examined the effect that specific training in experimentation might have on modifying the choice behavior of seventh-grade students. They also studied whether subjects who learned experimentation from a special training program would perform differently than from those who did not have the program or who entered the free-choice environment with an understanding of experimentation. Students in the "free-choice group" selected science activities for twelve 45-minute classes, twice a week, for a six-week period. For some students this was accompanied by the training program that took 15 minutes per week for six weeks. Results indicated that students who had the training program alone or the training program in conjunction with the free-choice activities were more successful in learning to control variables than were students receiving only the "free-choice" treatment. Although boys were more successful on the pretest, there were no significant sex differences on the posttest. Other conclusions drawn from the experiment were (1) once students learned how the program operated, they could work autonomously; (2) students had positive attitudes toward the activities; (3) subjects who were successful on the pretest engaged in more social activity; and (4) the only observed differences between boys and girls were that the boys engaged in disruptive behavior more frequently and that the girls were more teacher-dependent.

On the college level, Coombs (1978) compared the effect of allowing students in a botany class for nonmajors to choose activities rather than having them assigned by the instructor. He found that students with prescribed activities achieved more cognitive knowledge but that those with a choice of activities had increased long-term retention and had more positive attitudes toward learning. Newhall (1978) also found that students had a positive attitude toward instruction when permitted to select one-third of the content of an introductory physics course.
In examining the studies related to individualized instruction the following conclusions can be made:

1. When methods such as audio-tutorial instruction, PSI, learning activity packages are used for instruction, students' attitudes toward the subject and/or method of instruction are generally positive. It is difficult to know whether this effect is stable over time, or due to the novelty of using a new method. Would students still retain their positive attitude toward AT or PSI if all courses they took were given in this format?

2. Cognitive gains from individualizing instruction have been mixed. With AT instruction and with learning activity packages, cognitive gains are generally reported when the method is used for a small number of units or over a short time span. Cognitive gains for PSI have not been clearly established.

3. Elementary school pupils (grades six and seven) profit from learning in a free environment.

**Inquiry Approaches**

Included in this section will be studies that examine inquiry, inductive-deductive approaches, curiosity and creativity, and laboratory approaches. The latter is included because frequently the purpose of laboratory activities is to promote inquiry. One of the major difficulties in grouping studies in this section is the diverse use of the word 'inquiry'. This lack of precision in the definition makes it almost impossible to generalize.

**Inquiry Strategies**

Five studies were reported that examine the effect of inquiry teaching on various dependent measures. Another study examines students' reaction to inquiry in different learning atmospheres.

On the elementary level, Van Rennes (1978) determined the effectiveness of inquiry approaches for the study of museum exhibits by sixth-grade students. She compared four approaches: (1) teacher-led inquiry, (2) written inquiry with written feedback, (3) written inquiry without feedback, (4) existing museum label only. She found that teacher-led inquiry was the most effective and that written feedback had no influence.

Marek's (1978) study was concerned with investigating the effect of inquiry treatment on the intellectual development, content achievement, and intelligence quotient of high school biology students. The experimental treatment employed was the Inquiry Role Approach program developed by the Mid-Continent Regional Educational Laboratories. Marek found that
this approach was successful in promoting cognitive development, increasing content achievement, enhancing inquiry skills achievement and producing gains in mental ability scores.

Sanford (1978) and Tamir (1978) both examined the relationship between teachers' use of inquiry and other parameters. Sanford investigated academic aptitude levels as a factor affecting the use of inquiry strategies in high school biology classes. Data analyzed from 15 teachers and over 1100 students showed no significant relationships. Tamir studied the relationship between the inquiry level of instruction in biology and students' levels of inquiry performance and scientific curiosity. He also examined the relationship between students' level of inquiry and their level of curiosity. Results indicated no significant relationship between the level of inquiry used by the teacher and students' inquiry or curiosity performance. A high positive correlation was found between curiosity and inquiry levels using classroom means but this correlation was low when using individual scores.

Wright (1978) examined the effectiveness of instruction on the open exploration or inquiry skills of students. He studied the influence of intensive instruction on either observation of details, or hypothesis generation on ninth-grade students' behavior. Open exploration skills included noting (1) the number of details, (2) the number of hypotheses, (3) the quality of the hypothesis, (4) the number of questions, and (5) diversity of questions about a discrepant event. He found that students who had received training in either observation of details or hypothesis generation were significantly better than the control group which received no training. The type of training received made no significant difference.

Tjosvold et al. (1977) examined how fifth-grade students react to inquiry and didactic strategies under conditions of cooperation and competition. They found that students in the competitive inquiry treatment disapproved of the way the lesson was taught and also of the teacher, significantly more than did subjects in other treatments. They concluded that the traditional interpersonal competition of American schools is incompatible with inquiry teaching methods.

In a subsequent study, Tjosvold and Santamaria (1978) studied teachers' supportive behavior of students' competence to make decisions in both cooperative and competitive classroom environments. Four preservice teachers, trained to create four different classroom conditions (supportive-competitive, supportive-cooperative, nonsupportive-competitive, nonsupportive-cooperative) each taught groups of fourth- and fifth-grade pupils under each of the above conditions. Students' commitment, confidence, and expectations about decision making were measured through the use of a questionnaire (reliability = .86) administered after the treatment.

Results indicated that the support condition was more effective in promoting greater commitment and confidence in decision making than was the nonsupport condition. While the cooperative condition promoted
more positive expectations than the competitive condition; there were no differences in pupils' confidence or commitment to decision making.

These results indicate that conscious efforts on the part of the teacher can facilitate greater confidence in decision making, even in competitive classrooms. This should make inquiry teaching more compatible with the environment in many American classrooms.

Closely allied with teaching students to become inquiry oriented is teaching them to become creative thinkers. Simonis' (1978) study sought to determine if students would become more creative by introducing them to brainstorming techniques and group practice in developing analogies. She found that undergraduate general science students showed a significant increase in verbal creativity as measured by the Torrance Tests of Creative Thinking after using these techniques. No retention tests were administered.

Inquiry teaching is sometimes equated with inductive rather than with deductive methods of instruction. Three studies examined the effectiveness of these modes of instruction.

Davis (1978) investigated whether a lecture-discussion method, a verification method, or an inductive laboratory method would result in higher achievement, attitude, and understanding of science for fifth- and sixth-grade students. She found that the method of instruction made no difference on achievement but that attitudes toward science and the understanding of science were significantly influenced by the inductive laboratory approach.

In a study of inductive versus deductive approaches in teaching chemistry (a one-lesson program on chemical equilibrium) to ninth graders, Hermann and Hincksman (1978) found that the deductive group performed significantly better on the immediate retention test. On a delayed posttest administered 14 days later. However, no significant differences in scores were found. There were no differences reported according to student anxiety. Girls outperformed boys.

Douglass and Kahle (1977), in a study of field dependence and field independence and inductive and deductive methods of biology instruction, also found that the deductive method was superior. No long-term retention test was administered.

Laboratory Approaches

The laboratory serves many functions in the teaching of science. One of its major uses can be to provide students with the opportunity to engage in inquiry. Studies described below examine the function of the science laboratory and describe its usefulness in this regard over nonlaboratory approaches. The reader should refer to the section on Individualized Instruction, in particular, PSI, for other research examining the usefulness of the laboratory.
Tamir (1977) reported on the actual use of high school and college laboratories in Israel. He found that students in the ninth- and tenth-grade labs spent proportionally more time on prelab discussion whereas more time was used for postlab discussion in the eleventh and twelfth grades. In college science courses the postlab phase was eliminated and individual reports were substituted. In relation to inquiry, he found that a larger percentage of inquiry-oriented teachers conducted postlab sessions than did noninquiry teachers, and that there were more inquiry-oriented than didactic teachers. On the high school level approximately 66 percent of the time was devoted to actual lab work. In college courses this varied from 34 to 86 percent.

Three studies examined the value of laboratory experiences. Webber (1978) sought to determine the effects on the delayed retention and consequent transfer of selected rules of proportion commonly used in physics instruction with and without laboratory experiences. He found it made no difference for either retention or transfer whether or not the laboratory or an expository method of instruction was used.

Davis (1978) found similar results in a study in which she compared the effectiveness of the lecture-discussion method, verification laboratory method or the inductive laboratory method on achievement of fifth- and sixth-grade science students. There were no significant differences in achievement according to method. The inductive laboratory approach, however, was more effective in producing positive attitudes toward science and a better understanding of science.

Spear (1978) also examined the effectiveness of laboratory instruction. He compared the success of students who took both lecture and laboratory in a college geology course versus those who took only the lecture. Students who took both lecture and lab averaged 10 percent higher on the total semester scores over those who registered for the lecture only.

Lee (1978) examined the role of the laboratory instruction. She surveyed science educators, college biology instructors, and students about the role of the laboratory in teaching biology. The five major functions of the lab (as identified in the literature) were affirmed by all groups. There were: manipulative skills, processes of science, knowledge of subject matter, attitudes, interests and values.

These studies show that the laboratory is not particularly effective in increasing students' knowledge of the subject matter, but that it does enhance attitudes. The effectiveness of the laboratory to teach manipulative skills, processes of science, and interests and values are areas that warrant further investigation.

Few conclusions can be drawn about the usefulness of inquiry teaching. It appears that deductive teaching is more effective than inductive teaching in enhancing science achievement. As with laboratory instruction, this must be carefully defined as there may be many facets of science achievement that an inductive approach might enhance.
Reading and Language Skills

An area of growing interest in science education is the relationship between reading and science achievement. Because one of the most common instructional modes of acquiring science concepts and facts is through the textbook, science educators have considered the relationship between reading achievement and science achievement, readability of science texts, and the kinetic structure of the textual materials.

Reading Achievement

Two studies examined the effectiveness of a process approach of teaching science on reading achievement with young children. Both failed to confirm earlier studies that showed beneficial results of a process-oriented curriculum. Merricks and Crocker (1978) examined the effect of a science process curriculum combined with supplementary reading materials. They found that the combined treatment was more effective than either of the other treatments in promoting reading comprehension for first graders. In grade three, however, the supplementary reading program alone produced superior results. Because there were some methodological problems in the research design of this study, the results must be interpreted with caution.

Similar results were found by Quorn and Yore (1978). They compared the reading readiness skill acquisition of kindergarten children who participated in a formal reading readiness program, an informal reading readiness program, and a science process program (SAPA). The formal program used The First Talking Alphabet materials. No significant differences according to program used were found.

Reading and Science Achievement

Two studies investigated the relationship between success in reading and science achievement. E. W. Fuller (1978) studied the science achievement of third graders using visual, symbolic, and manipulative instructional treatments. She found no difference in achievement on a science test on organisms, according to the treatments, however, she did find that when students were classified according to their reading performance, there were significant differences in achievement among high, middle, and low readers.

Goslin (1978) examined the relationship between oral/aural language proficiency, reading ability, and science achievement of Puerto Rican students enrolled in a bilingual education program. He found that oral/aural language proficiency in English was the strongest indicator of science achievement. Reading ability when assessed in Spanish was the second strongest indicator.

Adrian (1978) found that there was a positive relationship between reading achievement, science and mathematics achievement, operative
reading comprehension, general self-concept of ability, and self-concept of ability in mathematics, English, social studies, and math. Subjects in the study were above average ninth-, tenth-, and eleventh-grade students.

Readability

Because of the positive relationship between reading comprehension and science achievement, several studies were concerned with textbooks and their readability. Studies that examined the readability of science textbooks include those by Narke (1978), Middleton (1978), and Cohen and Poppino (1978).

One method that is sometimes used to measure reading ability is the use of a CLOZE test. Cohen and Poppino sought to determine whether there was a relationship between the students' scores on CLOZE tests and biology grades, whether it could be used to predict grades, and the effectiveness of a study skills course. They found that the students' scores on the CLOZE test were related to their success in biology and could be used to predict their grade. The study skills course was effective in increasing students' confidence in their biology skills.

Clarke was interested in determining the minimum reading level needed by eighth-grade students to satisfactorily comprehend their social science textbooks. She found that an eleventh-grade level was necessary. She also reported that the Smith Readability Formula was the most accurate in judging the readability levels of both the social studies and the science textbooks used in the study. The abstract did not indicate which readability formulae were compared.

Two other readability studies concluded that Frye Readability Graph accurately predicted readability of biology textbooks. Middleton (1978) investigated the possibility of predicting the readability of biology textbooks, students' general reading ability, and students' general reading ability. He found that each of the above variables could be predicted from the other two. The Frye Readability Graph was used to determine grade level readability.

Western (1978), in a comprehensive study of readability formulae and biology textbooks, found that the Frye Graph most nearly matched performance scores for the four biology texts published in 1973 that were examined. Scores from seven formulae were compared for the four biology texts.

Pictorial Diagrams

Because the readability of a science text is an important factor in science achievement, several investigators have examined whether the readability of a text could be improved by adding pictorial illustrations and diagrams. Thomas (1978) studied the influence of pictorial
Illustrations coupled with the textual material on the reading comprehension of fourth-grade students. He compared the use of color photographs, simplified line drawings of the color photographs, and no pictures at all. He found no significant differences between scores on a test that included both literal and inferential comprehension. Students were also classified according to their reading achievement. Significant differences according to achievement, but no interaction effects, were reported.

Although Thomas concluded that the use of pictures did not aid comprehension, this may be an overgeneralization. The pictorial figures in his study were limited to those that could be made from color photographs. Many other types of pictorial diagrams that show abstract relationships exist, and these indeed may be helpful.

The placement of written cues in relation to pictorial diagrams may also be important in learning from diagrams. Kauchak et al. (1978) studied the effect of cue specificity on learning by fourth-, fifth-, and sixth-grade pupils from graphical materials in science. They compared the effectiveness of asking specific questions (cued questions), incidental learning (nonced questions) and generalizing questions. They found that textual cues could significantly increase the amount of information gained from graphical material, the use of specific cued questions aided learning most.

On the college level, Kaplan (1978) sought to determine whether improving the readability of the text material or the introduction of labeled line drawings with or without improving the readability of the text material would improve chemistry achievement. Working with students in four community colleges, he concluded that improving the readability of the text material was most important in increasing comprehension of descriptive topics for both poor and good readers. For more quantitative topics, the original text was better. Line drawings had no effect on comprehension.

Grammatical Construction

Another approach that might be used to make reading material in science more comprehensible is to make students aware of the grammatical constructions frequently used in this type of writing. Randleman (1978) studied the grammatical constructions commonly used in secondary science materials. Of the five types of grammatical construction studied, prepositional phrases were used most. Physics and chemistry sentences were found to be more complex syntactically than sentences in the other sciences, and many sentences conventionally classified as "simple" because they consisted of a single clause were actually very complex from the standpoint of readability.

Foreign students frequently encounter much difficulty in understanding science texts. To help aid Iranian students' learning, Pakseresht (1978) studied the major syntactical and textual features of scientific phrases in the field of life sciences. On the basis of both the syntactic and discoursal evidence, special suggestions were made for both material selection and development and for reading instruction.
Kinetic Structure

Closely allied to studies of the readability of science texts are studies that consider the kinetic structure theory of O. Roger Anderson. His theory postulates that a communication with an appropriate degree of structure will facilitate increased knowledge acquisition over one with a lower degree of structure. Four studies tested his theory.

C. G. Cohen (1978) examined the effects of high and low structure lessons on the topic of how to perform two laboratory experiments in chemistry and how to compute mathematical data gathered during the performance of these experiments. Achievement was measured through assessing students' success on a card-sorting task that simulated the laboratory experiment and on a problem-solving task.

No differential effects were found between students using high and low structure lessons for a card-sorting manipulation, however, high structure lessons were more successful in teaching mathematics computation. No interactions between the structure and student ability were present.

Kittrell (1978) compared the effect of high and low structure in oral and written communications on knowledge recall by college students of varying verbal ability. Oral communications were presented by audiotape whereas written communications were presented in booklet form. She found that for high verbal students, the written mode was significantly more effective than the oral mode, but that there were no differences on the retention test. The study also demonstrated that for all subjects, communications with high structure produced greater recall than did communications with low structure on both the immediate recall test and the retention test.

Robbins (1978) had as his purpose the determination of the effect of high, low, and potency controlled structure lessons on the performance of a manual laboratory task and related concept acquisition in biology. Potency controlled structure lessons were defined as those that "have a repetition of key words between adjacent discourse units in a sequential pattern that builds a theme idea to a climax and then digresses into the next topic" (p. 6042-A). He found no significant differences in knowledge acquisition among three groups of students who received the lessons on the components of blood. The only significant differences found were in the psychomotor tasks between low structure and potency controlled structure lessons and between low and high structure lessons. It was postulated that failure to find differences on the knowledge acquisition task was due to this task being too difficult for students to comprehend.

Lu (1978) investigated the effect on knowledge acquisition of three models of teaching college astronomy based on the principles of kinetic structure analysis, and the psychological concept of massed versus distributed practice. Three topics in astronomy were each prepared on audiotape using three formats. They were initial nondistributed
integrative, terminal nondistributed integrative, and distributed integrative. Results indicated that distributive integrative lessons were more effective than nondistributed integrative lessons (both initial and terminal). These findings extend the validity and applicability of Anderson's kinetic structure theory and raise questions about the usefulness of Ausubelian advance organizers which, to a great extent when coupled with a lesson, resemble the initial nondistributed integrative lessons examined here.

The following conclusions can be drawn about this section on reading and language skills from the studies reviewed.

1. Contrary to previous research, no positive relationship was found between science instruction utilizing the science process skills and reading achievement of young children.

2. Reading and science achievement are positively related.

3. Readability formulas can be used to assess the reading level of science texts with good accuracy.

4. Pictorial diagrams may not enhance textual materials sufficiently to cause an increase in achievement.

5. Textual materials with a high kinetic structure are more effective than those with less structure.

Classroom Practices

Behavioral Objectives

Although research in this area is not as extensive as it was in the early seventies, four studies were reported during 1978. Two of these (Kahle, 1978 and Varano, 1978) were comparisons of the effectiveness of behavioral objectives with advance organizers. A discussion of their studies can be found in the section on Ausubelian research.

Netivinyoo (1978) determined the effectiveness of a lecture-demonstration film approach versus a lecture-behavioral objectives approach to the teaching of chemistry. Also examined was a comparison of the time spent in instruction (7 consecutive 55-minute class periods versus 11 consecutive 55-minute class periods). No significant differences were found for either the method or the time.

Anderson and Fowler (1978) studied whether different cognitive levels of behavioral objectives (high and low) and selected entering behaviors (prior knowledge and critical thinking) had any effect on the learning and retention performance of preservice elementary education majors. The unit studied was on population genetics and the medium was self-instructional booklets. Students (N = 122) were assigned to one of four treatments: low cognitive level behavioral objectives, high cognitive level behavioral objectives, high and low cognitive level
behavioral objectives, no behavioral objectives. All students were pre-tested for previous knowledge in genetics and critical thinking ability.

Although there was no main effect, Anderson and Fowler found that (1) the high prior knowledge group performed significantly higher on knowledge level questions when given both high and low level behavioral objectives, and (2) the high prior knowledge group performed significantly higher on the retention tests for both knowledge and comprehension items when given high level behavioral objectives.

Results of this study and the others reviewed indicate that behavioral objectives of themselves do not appear to be universally helpful in improving achievement. They may be a useful adjunct when coupled with advance organizers, with certain types of learners, and with certain types of instruction.

**Feedback**

Another method of directing students' attention to what needs to be learned is through the use of feedback. Feedback to students may be given less formally through the use of questions in the classroom or more formally through computer-assisted instruction.

Wooley (1978) studied the effectiveness of two types of feedback using computer-assisted instruction to teach mathematical skills needed for learning astronomy. One type of feedback provided students with knowledge of the results, knowledge of the correct response, and response contingent feedback. The other type provided knowledge of results only. The former type of feedback occurs in guided discovery lessons, the latter in discovery lessons. Wooley found no differences in achievement or in attitudes of college astronomy students toward the feedback modes.

Marble et al. (1978) examined the effect of a continuous feedback schedule versus a partial feedback schedule when both grades or grades and comments were made on eighth-grade students' laboratory reports. The study was conducted over an eight-week unit during which 12 laboratory reports were assigned. Both achievement and improvement on reports were examined. The authors found no significant differences for whether the grade alone or the grade with comments were used. The continuous schedule, however, was superior to the partial one (in which half the reports were graded) in producing higher achievement.

Davidson (1978) investigated whether laboratory reports or competency measures (quizzes) had a greater effect on achievement in a college general science laboratory course. He found that there were no long- or short-term differences in achievement according to the method used. Students who wrote the laboratory reports spent twice as long outside of class in completing them as students who took the quizzes, but no differences in attitude were found.

Bitner (1978) also compared feedback modes. He studied the effectiveness of both the amount of feedback (answer only or answer with
information) and the mode (visual or audiovisual) on achievement, retention, and attitudes of students of different anxiety in a college laboratory course. Results of his study indicated no differences in achievement according to the amount of feedback or to the instructional mode. There was no significant interactions between anxiety and the methods. The findings, however, did indicate a significant student preference for the feedback methods that provided correct answers with information and the audiovisual mode of presentation. Low anxious students performed better than did high anxious students on achievement tests.

Boulanger (1978b) assessed the effect of three different feedback modes on learning the balance-beam principle by undergraduate elementary education majors. He hypothesized that manipulative feedback (as compared to single pictorial feedback and multiple-pictorial feedback) would result in higher immediate application and transfer scores, a longer instructional time but no difference in instructional efficiency, and a mathematics x treatment interaction. An attitude test, as well as an application test, was administered. Its function was to control for possible increases in attending behaviors due to increased student interest for the manipulative treatment. Results of the investigation showed that manipulative feedback was significantly more effective on the application test than were the other modes. This type of feedback, however, required more instructional time. Because no differences in instructional efficiency among the three modes were found, an improved pictorial treatment might be equally effective in producing high application scores. Attitudes were generally more favorable toward the manipulative mode.

Two other studies on feedback were reported during 1978. Van Rennes (1978) found that, in a museum environment, inquiry lessons with feedback were no more effective than inquiry lessons without feedback for sixth-grade students learning physics. F. N. Finley (1978) compared the effectiveness of using cumulative feedback with isolated feedback in geology instruction. No definitive results were discussed in his abstract.

Questions

Three studies examined the effectiveness of different questioning techniques on student achievement. Two studies were concerned with high school students, one with eighth graders.

B. O. Anderson (1978) studied the effect of wait-time on high school physics students' response length, classroom attitudes, science attitudes, and achievement. He concluded that pupils, when given a long wait-time, made longer responses to pre-planned questions than did pupils given a short wait-time. He also reported that "long wait-time" pupils were more apathetic about the class and found it less difficult than they had anticipated. No differences in achievement related to wait-time were reported.

Carazo Santaliz (1978) compared the effectiveness of two lecture teaching strategies on concept achievement for high school Introductory
Physical Science students. The two strategies compared were the oral-explanatory (traditional lecture) and the answer-centered (no description given). The oral-explanatory method produced significantly higher achievement and retention than the other method.

Holliday et al. (1978) examined the effectiveness of including study questions as an adjunct to learning about fossils from written textual descriptions by eighth-grade students. Students were administered a verbal ability test prior to being placed in one of their treatment groups. These were: text plus 20 study questions, text plus no study questions, and placebo passage. It was found that students who were low verbal performers and who were provided with the text and no study questions scored significantly higher on the posttest than did those low verbal learners who were provided with the text and study questions.

Other Classroom Adjuncts

Included in the following section are studies of miscellaneous teaching strategies that might have positive effects on students. These range from improving students' creativity to general classroom procedures for aiding achievement.

Simonis (1978) sought to determine whether giving undergraduate students practice in transforming techniques and formulating analogies would stimulate their creativity. Results based on the Torrance Test of Creative Thinking showed a significant increase at the 0.001 level in verbal creativity for students having the practice.

Symington (1977) investigated the use of class discussion followed by a period of unstructured observation and experiments as a method of increasing fifth and sixth graders' ability to view scientific problems as useful starting points for investigation. He found that the discussion resulted in a decrease in data-gathering problems and an increase in data-processing problems.

Welch (1978) studied the effect of study skills instruction on science achievement and self-as-learner attitude of seventh-grade students. The study skill unit was taught as a part of the regular science lessons over a two-week period. A control group received instruction in library skills. No significant differences were reported in achievement of a science lesson, study habits or in students' attitudes.

Counceller (1978) compared the effectiveness of a note-test method to the traditional lecture method of teaching general introductory microbiology. No differences in achievement or in attitude of students toward the course were noted as a result of the different methods. Because the note-test method required less lecturing than the traditional lecturing method, it holds some promise as an effective alternative in some courses.
General Physical Adjuncts

The above studies focused on changes in classroom procedures that may affect learning. The following studies examined the usefulness of physical adjuncts to learning.

Trollinger (1978) compared the effectiveness of teaching biology through the use of simulation games with conventional techniques of instruction. No differences were found on a factual knowledge or on a retention test according to treatment. On two of the five games, students had more positive attitudes toward the activity than did the control group who watched films and completed worksheets.

E. W. Fuller (1978) studied the effectiveness of transparencies, booklets, and puzzle-flash cards according to reading level and sex on science achievement of third graders. No differences were found among the instructional treatments or between sexes.

Szabo and Welliver (1978) examined the effectiveness of combining TV instruction with an inquiry-oriented science program on teachers' and students' inquiry behaviors. Although the study is discussed more completely under inservice instruction, it must be noted here that this mode was effective in increasing the inquiry behavior of elementary school children.

Specific Physical Adjuncts

Studies included in this section are concerned with the effectiveness of strategies and learning aids specific to one subject area.

Batoff (1978) reported on the usefulness of the Unit Box Approach for teaching elementary science. The key conclusion from his study was that the Unit Box Approach was very useful in preparing student teachers and inservice teachers to teach science in the elementary school.

Dederick (1978) studied the use of photomicrographs as a means of improving high school biology achievement and attitude. Students in the experimental group were given the same treatment as those in the control group except that their instruction was supplemented with photomicrographs of plant anatomical features. Results indicated that the use of the photomicrographs significantly improved achievement in recognizing and properly identifying the anatomical features, and that students' attitudes were positive. No indication was made of how time was controlled in the abstract. It is difficult to tell if the result was due to the photomicrographs or to additional practice in recognition.

Tarcza (1978) determined whether the use of color photographs reduced students' perceived difficulty of physical science laboratory tasks. Two methods of using photographs were used. The first was similar to standard science textbook format in that they were included as aids to material described in the laboratory guide. In the second method, the photographs and written material were presented as a preview
to the laboratory guide. These methods were compared with using neither photos or preview, and with using the preview alone. Analysis of the data over 11 experiments indicated no differences in perception of difficulty by students with or without the use of color photographs. Beneficial results noted were that students with the photographs tended to seek less assistance from the instructor and formed fewer partnerships in doing the experiments.

J. K. Fletcher (1978) examined the effectiveness of the traditional type planetarium program with a participatory type for eighth- and ninth-grade students. A description of the participatory instruction was not included in the abstract. No differences in achievement were found between the two instructional modes on either immediate recall or retention tests.

The following conclusions can be drawn from the studies included in this section on Classroom Practices:

1. The use of behavioral objectives by themselves was not effective in improving achievement. When coupled with advanced organizers, achievement improved.

2. The type of feedback a student received makes little difference on science achievement. An exception was that a continuous feedback schedule was more effective than a partial one. Different types of students prefer different feedback modes.

3. In examining the effectiveness of question-asking procedures, it was found that (a) the wait-time did not improve achievement although the length of student responses increased and (b) the addition of study questions in textual material impeded achievement.
COGNITIVE STYLE

The definition presented by Cronbach and Snow (1977, p. 375), "cognitive style implies an habitual pattern or preferred strategy of information processing," will be used as a criterion for inclusion of studies in this section. This definition includes such terms as "cognitive preference" and, to some extent, "cognitive ability." The review will be divided into parts based upon the level of instruction and/or subject matter.

Elementary School

Using sixth-grade elementary school students (N = 54), Lawson and Wellman (1977) investigated the relationship of Piagetian developmental level, field independence-dependence, and the ability of the students to make "critical value judgments." Piagetian level was measured by a series of three Piagetian tasks (conservation of weight, bending rods, and the balance beam). Developmental results were similar to previously reported levels and consistent with Piagetian theory with only six subjects being fully formal.

Critical value judgment ability was measured by raters scoring oral responses to five situations in which judgments by the student were required. Group Embedded Figures Test (GEFT) was used to measure field independence and dependence.

Results showed high correlation (.89 corrected for attenuation) between the balance beam and bending rods which was hypothesized by the investigators. Correlations between the GEFT and the tasks were modest (.48 -.60) as was the correlation between the GEFT and the sum of the values task (.44). Correlations ranged from .57 to .61 between the Piagetian and values tasks.

The correlational results plus a factor analysis of the measures led the authors to conclude that there appeared to be a strong relationship between field independence-dependence, Piagetian developmental levels, and ability to critically judge value statements. They concluded that there may be an underlying psychological construct relating the three areas.

An aptitude by treatment interaction study by D. J. Harmon (1978) attempted to determine the effects of teaching structure on students' anxiety and dependency levels. Fifth- and sixth-grade students (N = 168) were instructed using worksheets of varying teacher-student control over five instructional periods. The topic of the instruction was force. Aptitudes were measured with the Children's Manifest Anxiety Scale and the Modified Dependence Proneness Scale. The Ankney-Joyce Reasoning Test was used as a covariate to reduce personality group differences in knowledge and process level concept understanding, even though students

were randomly assigned to treatments. The author found that high anxiety individuals gained more knowledge when teaching was highly structured, that low teaching structure enhanced low anxiety students' performance relative to that of the high anxiety students (a finding contradictory to the literature), and low dependency individuals performed better with low teaching structure and with complex tasks. Persons high in both anxiety and dependence did better on both knowledge and process levels of achievement, and overall, a high degree of teaching structure was better for difficult tasks.

Harjoribanks (1978) studied the relationship of convergent and divergent abilities, academic performance, and school-related affective characteristics among 12-year-old English students (N=429). Convergent ability was measured in terms of scores on verbal and non-verbal reasoning sections of conventional intelligence tests. Divergent ability was measured by tests of fluency, flexibility, and originality. A questionnaire was used to assess school-related affective characteristics, and standardized and teacher-made tests of English, mathematics, French, physical science and biological science were the criterion measures collected.

Using multiple linear regression, raw regression coefficients and regression surfaces were presented. The regression surfaces showed that, for physical science, girls who were high on convergent ability and low on divergent ability scored higher than girls with the opposite characteristics. The results for boys were not as clear-cut with convergers, divergers, and all-rounders approximately the same in physical science ability. The hypothesis that divergers would score better on an ability test in English was not upheld. Convergers were higher for both boys and girls. A hypothesis of no difference between convergers, divergers, and all-rounders in biology achievement was also rejected with the order of ability being all-rounders, convergers, and divergers. The authors also concluded that for girls, high scoring all-rounders and convergers tended to have more positive school-related attitudes, while for boys the relationship was curvilinear with a maximum at moderate convergent-divergent levels.

Secondary School

Biology

Douglass and Kahle (1977) investigated the relationship between one cognitive dimension, field independence, and type of instructional sequence, inductive versus deductive, with secondary biology students (51 males, 39 females). Students were classified as field independent or field dependent on the basis of their score on the Thrustone Gottschaldt Closure Flexibility Test. The top and bottom quartiles were used as the field independent and field dependent students whereas students in the middle section were the control group. After random assignment of the field independent and dependent students to the two instructional strategies, students studied self-paced mastery units on Mendelian genetics and probability. The control students had a similar type of unit on content from different biology topics.
A two-way ANOVA on the posttest scores with the two factors of field independence-dependence and instructional sequence yielded no significant F ratios. Analysis of covariance using IQ scores as the covariate did yield a significant main effect for the cognitive style factor.

The authors concluded that the sample of their study may have been too homogeneous for the hypothesized interaction effects to appear. Results of retention effects, calculated by comparing formative and summative test results showed some advantage for the deductive instructional sequence.

In order to determine if a relationship existed between several student characteristics and cognitive preference, Tamir and Lunetta (1978a) used two forms of the Biology Cognitive Preference Test (BCPT), with secondary students (N = 177) at a summer science workshop. The two forms differed in the method of presentation of the items. In one form, descriptive items were followed by normative items while in the second form the item types were reversed. Differences between the forms were not significant.

Several background variables were collected at the same time as the cognitive preference measures. Major findings included a preference by the group as a whole for questioning versus recall, a higher preference for principles versus recall for students in higher grades, a higher preference for questioning among students from inquiry-based biology curricula, and differences in preference across topics in biology. No differences were found in preference according to gender, region, or hobby.

Chemistry

Using both Piagetian developmental level and field dependence-independence as student aptitudes, Cantu (1978) studied the effect of concrete and formal concepts, and the use of pseudoexamples on chemistry achievement. The author concluded that the use of pseudoexamples was beneficial to concept attainment, that even though formal students achieved more on both types of concepts, concrete students improved more with instruction based on concept analysis, and that field independent students learned concepts involving spatial ability more easily than did field dependent students.

Hofstein et al. (1978) studied the cognitive preferences of four student groups: post graduates in chemistry (N = 48), secondary chemistry students in academic high schools (N = 193), secondary chemistry students in vocational schools (N = 116), and secondary students who were participants in science clubs (N = 95). Using a modified form of the Cognitive Preference in Chemistry, a 34-item questionnaire to measure the four modes of cognitive preference, the researchers found several significant differences in the responses to the items between groups. The academic high school students showed a markedly lower preference for the recall mode than did the other groups. Academic high school students were higher in their rating of the application
mode than were the postgraduate students and the science club members. The academic high school students were the lowest on the principles mode of the four groups, but only significantly different from the science club group. Additionally, the academic high school students were the highest in their rating of the critical questioning mode and were significantly different from the vocational students and postgraduates in this area.

**Physics**

The schema of proportionality and its relationship to the development of cognitive structure in the area of mechanics was studied by Preece (1978b). From previous work he predicted that the index of associate structure $\mu$ would be negatively correlated with measures of proportional reasoning among fourth-form physics students.

The measurement of $\mu$ was by a free association word test. Proportional reasoning was measured by written missing number problems that used the proportionality schema. Using a biseral correlation coefficient between dichotomized scores on the proportionality measure and the measure of $\mu$ resulted in a value of $-0.32$, significant at $p < .05$. This indicated some support for the relationship between the acquisition of the schema of proportionality and the development of cognitive structure in the area of mechanics.

In an attempt to see if the matching of students and teachers as to cognitive similarity was related to physics achievement, Frinks (1978) used cognitive abilities tests from the French kit and Siegel's Educational Set Scale with 336 students and nine teachers. Similarity in cognitive abilities profile accounted for 9 percent of the variance in physics achievement and no correlation was found between educational set similarity and physics achievement. Additionally, no interaction effects were found, although conceptual students scored higher than did factual students on the educational set test and on the physics achievement test.

**Science in General**

In a correlational study, Raven and Adrian (1978) studied the relationships among operative comprehension, self-concept, and science achievement among "average and above average students" in ninth- to eleventh-grade classes ($N = 249$). Instruments used were the Raven Content Comprehension Test (RCCT), the Self-Concept of Ability Scale (SCAS), and the Sequential Test of Educational Progress Series II (STEP). Reading and mathematics achievement scores on the Stanford Test of Academic Skills (TASK) test were obtained from school records.

The RCCT required the student to use the seven logical operations of Piaget to construct the concepts necessary for answering the test questions. The SCAS reports a general self-concept and self-concept ratings in mathematics, English, social studies, and science. The other instruments were of a standard achievement test format.
Results indicated significant positive correlations between the RCCT and reading, mathematics, and science achievement, as well as significant positive correlations with general self-concept and science ability self-concept. The authors concluded that the RCCT appeared to be measuring a somewhat different construct than standard achievement, in that the highest percentage of variance explained by an achievement measure was only 31 percent of the variance on the RCCT. No sex differences were found in the RCCT results.

Braun (1978) reported the results of Piagetian interviews, the Group Embedded Figures Test (GEFT), a four-dimensional discrimination learning task, and achievement tests with 30 ninth-grade and 30 twelfth-grade "science oriented" students. While little detail was reported, the results indicated that previous science training, cognitive ability and cognitive style had little effect on the performance of the Piagetian tasks. The author concluded that "science oriented" students develop formal level thought before they begin formal science study.

Using the GEFT and various measures of science achievement, science attitude, and IQ, Krajcovich (1978) attempted to relate field independence to various students' characteristics. Ninth-grade students of mixed (N = 204) and high ability (N = 47) were used as the subjects. Correlations indicated modest relationships between science attitude and science achievement, and IQ. No relationship could be determined between field independence and science attitude. Field independence was, however, positively related to IQ (r = .52).

In two dissertation experiments, Grandy (1978) and Burow (1978) investigated locus of control and its relationship to science achievement and other student attributes. Grandy found no changes in locus of control among the three experimental groups that were studying mini-courses from the ISIS series. The three groups were a student decision group, student-teacher decision group, and a teacher decision group. Internal students in the teacher and student decision groups did, however, score significantly higher in science achievement than did external students. External students scored higher in the student-teacher decision group. Burow's results indicated that internal students were higher achievers than external students and that the internal students had a better attitude toward science instruction and tended to view science instruction as more inquiry oriented.

Van den Berg et al. (1978) completed a lengthy validation study of cognitive preference by using science-oriented students at a summer science workshop (N = 144). The Science Cognitive Preference Inventory (SCPI) was the instrument under study. The instrument contains 30 items developed to measure cognitive preference for four different modes; recall (R), principles (P), questions (Q), and applications (A). Internal consistency and test-retest reliabilities were calculated with high results on the Q and R scales, but poor results on the P and A scales. Construct validity questions were addressed through the use of divergent and convergent validity techniques and factor analysis. The Q and R dimensions appeared to have stronger construct validity
than the P and A dimensions. An indication of change in cognitive preference over the course of the summer workshop was also noted, a finding somewhat inconsistent with earlier cognitive preference work. The authors concluded that more research is needed on the interpretation of cognitive preference modes, a conclusion definitely warranted by the results.

In another study of cognitive preference, Tamir and Kempa (1978) worked with a large sample (N = 599) of Israeli students to determine the reliability, validity and correlations of cognitive preference with student characteristics. Using three instruments, one each for biology, chemistry, and physics, developed by Kempa, they obtained modest to high interval consistencies and a factor structure consistent with the four-dimensional makeup of the tests. The general preference order for the students was principles, recall, application, and questioning although there was no difference between application and questioning for the biology instrument. Low but significant correlations were found between achievement and the principles, and the principles minus application, dimensions. Some differences, across dimensions; were also found between the three types of school in the study: occupational, agricultural, and city.

Rowe (1978b) presented an interesting summary of some of the personality type research using the Myers-Briggs Type Indicator that has been completed and possible future uses of the instrument. The large differences between groups of "average" students and "science-oriented" students in their personality type is quite revealing. "Science-oriented" students appeared much more in the intuitive area while the "average" students were dominant in the sensing area. Noting that instructional practices vary little between different students, even though their personality types may be very different, Rowe posed several questions that should be considered by teachers in developing instruction to fit students.

College Instruction

Biology

In a dissertation study, Frase (1978) investigated the cognitive styles of successful and unsuccessful introductory biology students in a community college setting. Students divided into lecture and discussion groups showed no significant differences in performance. The instrument used to determine cognitive style is not reported but characterizations of unsuccessful students as being more prone to learn from verbal communications and tending to possess high motor ability were noted.

Murray (1978) investigated the relationship of a visual recall test to achievement in an introductory college biology course (N = 225). The visual recall test was a slide representation in which 15 of the slides had been seen previously in a lecture and 15 were new. Achievement was measured by three course examinations, the first of which was more closely tied to the visual recall test content.
Results showed that there were significant differences in achievement on all the achievement tests across five groups made from quintile of the visual recall test. The results did decrease from achievement test one to achievement tests two and three as the author had hypothesized. Post-hoc analysis, however, was not applied to determine which groups were significantly different from each other. The use of ANOVA rather than multiple regression may have also resulted in some loss of information. The author concluded that visual recall tests may be a useful method to predict achievement in an unobtrusive manner and may be used in determining the relationship of visual and verbal cognitive structures.

Chemistry

Studies by Lloyd (1978) and Schiesler (1978) looked at the relationship of several students' attributes to achievement in freshman college chemistry.

Lloyd found that internal students as defined by the Group Embedded Figures Test, Piagetian transitional or formal operational students, and students who scored well on a test of mathematical concepts were much more likely to be successful in a science major's chemistry course. Schiesler's results indicated no relationship between locus of control in remedial chemistry students' achievement. Significant predictors of both mid-term and final achievement were college grade point average and SAT-mathematics scores. Race was also a significant predictor for mid-term achievement. The high drop-out rate of the course was a confounding factor that may warrant further study.

Drake (1978) attempted to look at relationships between matching of a student's preferred mode of instruction and several personality variables which have been seen in other studies. After determining the student's preferred mode of instruction by a questionnaire, students were randomly assigned to either lecture or self-paced environments. This resulted in some students being matched and others mismatched to their preferred mode. Other attributes such as field independence, Cattell's 16 PF test, Piagetian classification, and Procrastination Factor were also measured.

Results indicated no significant difference between the matched and unmatched students in achievement. Grade point average and prior knowledge of chemistry were the best predictors for overall group achievement. Personality factors and Piagetian classification were not significant factors while field independence showed some discrimination ability between high and low achievement students.

Physics

Only one article appears in this classification. Cohen et al. (1978) investigated the relationship of Piagetian developmental level, SAT math and verbal scores, and achievement in introductory physics.

Using a series of three or four Piagetian tasks, students were classified as concrete, post-concrete, and formal by two independent
raters. Relationships between course achievement and Piagetian level were somewhat different than expected. Using only the two most different courses, the easiest and the most difficult, achievement was significantly different between the Piagetian levels only for the elementary course (p = .046) not for the more difficult course (p = .435). SAT math scores did correlate significantly with course achievement in both situations. The authors pointed out that these results are inconsistent with some previous research and suggested that students of concrete developmental level may be successful in current physics courses without major modifications in the curriculum.

Teacher Training
Preservice Teachers

Preservice science teachers' attitude toward teaching (teacher-oriented or student-oriented) and cognitive style (factual, neutral or conceptual set) and their relationship to science process skill attainment was the area of research for Jantaraweragul (1978). While interaction effects of the two factors approached significance, neither main effect of attitude nor cognitive style was significant for the sample of students (N = 58).

Heller (1978), in a dissertation dealing with the ability of college students enrolled in education classes (N = 130) to solve written logical puzzles, found that students' abilities were independent of the students' field independence and conceptual tempo. Field independence was measured by the Embedded Figures Test and conceptual tempo by the Matching Familiar Figures Test. In addition to the main effects not significantly adding to the multiple regression equation, no interaction effects were found.

In an aptitude by treatment interaction experiment, Horak and Slobadzian (1978) reported an interesting disordinal interaction between locus of control and instructional treatment with elementary preservice teachers. The instructional treatments were "high" or "low" structured instruction on planning science field trips for elementary children. Locus of control was measured by the James Internal-External Locus of Control Scale. The two criterion measures, science content and science process achievement, were developed by the experimenters. A strong disordinal interaction was shown between the high and low structure treatments, locus of control, and the science content achievement, with low locus of control students (internals) achieving much better in the low structure situation, while high locus of control students (externals) achieved much better in the high structure group. While the lack of reliability data on the criterion measures is a drawback of the study, the appearance of the interaction effect indicates that further research in the area is warranted.
Inservice Teachers

The stability of cognitive style as measured by the Group Embedded Figures Test, and possible modifications of cognitive style by training with SCIS material was the area of interest for Pringle and Morgan (1978). Using intact summer session classes in which SCIS was a part of the curriculum in two classes but not in another two, the researchers pretested one of each kind of class with the GEFT. Posttests only were conducted in the remaining classes. Results indicated a significant increase from pretest to posttest but reasons for the increase could not be separated as to whether the effect was due to the SCIS curriculum or to the pretesting.

In another study where the GEFT was used, Harty (1978) found no significant relationship between field independence and various observable classroom behaviors such as praising or using student ideas. Flanders Interaction Analysis Categories was used to code teacher behavior for 100 secondary school teachers in the study. As well as the nonsignificant relationship of the GEFT, the author noted that lecture formats continue to be the standard classroom method of instruction.

The previous section of 28 studies has indicated that several cognitive styles have been the subject of research investigations but the results of these studies have been mixed. The construct of field dependence-independence was examined in eight studies but only two showed strong results. This area may not be fruitful for future research. The Piagetian studies in this section which were mainly concerned with relationships between Piagetian level and cognitive style may be an area of future work, although inconsistencies in results such as the study by Cohen et al. make generalization difficult. While only three articles discussed locus of control and the results were mixed, validation studies of the construct might be a future direction of interest in science education.

The use of multiple regression techniques and multivariate analysis are areas that science educators may wish to explore to maximize the information obtainable from their data. Some articles in the section sacrificed information by "blocking" on aptitudes to create specific categories for analysis of variance. Recent work by Cronbach and Snow (1977) and by Cohen and Cohen (1975) offers alternatives to this technique through the use of multiple regression. While multivariate techniques advocated by Tatsuoka (1971) and Peterson and Carlson (1979) are more difficult to conceptualize than univariate techniques, they offer the experimenters the opportunity to look at multiple dependent measures.


STUDENT CHARACTERISTICS

While the previous section on cognitive style dealt with research of a correlational and experimental nature, this section deals almost exclusively with correlational studies predicting student outcomes from various student characteristics. While the dividing line between those student characteristics that can be considered to be cognitive styles, and those that are of a background nature (age, sex, etc.) is not always clearly delineated, this section is more directed towards the latter than the former.

Secondary School Students

General Science

In a study dealing with students in ISCS Level I classes (N = 393), McDuffie and Beehler (1978) used several characteristics to predict workstyle and achievement. Workstyle was measured by teacher ratings and achievement by a teacher-mode test. Both criterion measures showed high reliability (W20 = .95 and .82, respectively). Subscales of scientific attitude, aptitude, skills, and knowledge tests plus sex were the predictor variables. While the variance accounted for by the predictors of workstyle was not especially large (27.5 percent), the large role of scientific attitudes did appear, as well as a surprisingly low percentage for reading skills. Achievement predictors were similar with a slightly less percentage of variance accounted for (23.6 percent). By dichotomizing the group into low and high achievers, and low and high workstyle groups, discriminate analysis was performed to give an indicator of measures that would give the maximum discrimination between groups. Attitudes again played a large role in the discrimination between groups, along with sex and some knowledge components.

In a follow up study by McDuffie (1978), ISCS students were studied across levels of the program. A sample of 372 level I, 290-level II, and 379 level III students were studied using the same instruments as the earlier study. Results within levels were similar to the first study, although basic concepts appear to be a stronger discriminator at level II than would have been predicted just from the level I results. Comparisons between levels showed reading achievement to be the only major discriminator for low achievers. For high achievers computational skills and mathematical ability were significant between levels I and II, while reading, workstyle, and intellectual aptitude were significant for the change from level II to III.

Stevens and Atwood (1978) also investigated interest as well as science process achievement among seventh- (345), eighth- (196), and ninth- (529) grade students in the Midwest. Using subtests of The Test of Science Processes and the Science Interest Inventory as pre-and post-instruments resulted in increases in science process skills but a decrease in science interest. While the research design could not determine what the effects of this increase in process skill and decrease in attitude were due to, it does indicate further research in this area may be warranted.
In a multi-disciplinary dissertation, Wanous (1977) studied the relationship of science achievement to the nine cognitive abilities test of Guilford's SI model, as they related to bilingual (English-Spanish) seventh- and eighth-grade students. Students (N = 194) were randomly assigned to take either the English or Spanish SI battery. Results from the testing were then factor analyzed and used to predict scores on the Metropolitan Achievement Test science subtest. The factor structures did not replicate Guilford's model and the factor scores predicted rather low (17 percent Spanish, 13 percent English), although statistically significant, percentages of variance in the science achievement scores.

**Life Science**

Doran and Sellers (1978) investigated the relationship of mental ability, gender, biology achievement, and process achievement to the self-concept of biology students. Using an investigator prepared but well-validated measure of self-concept and standardized instruments for the other characteristics, the researchers obtained data from 320 students.

Results indicated that the best predictor of self-concept was biology achievement which accounted for 8 percent of the variation in self-concept. Other measures such as process achievement, gender, and mental ability accounted for only an additional 1 percent of the variation. Controlling for mental ability resulted in even lower percentages of variance being accounted for by the predictor variables (approximately 2 percent).

In a study attempting to relate the perceptual decentering ability of high school biology students and their success in biology, M. L. James (1978) used the Perceptual Ambiguity Test (PAT), the Perceptual Integration Test (PIT), and the Biology Structure Test (BST) with 40 tenth-grade students. The PAT and PIT tests were not significant predictors of biology achievement, but mental ability scores were. The PAT and PIT were, however, significantly correlated with the BST indicating a relationship to biology tasks that have a visual perception component.

In two studies done in international settings, both Conkright (1978) and Namuddu (1978) investigated adolescents' interpretation of certain naturally occurring events in relation to the superstitions or beliefs held by their peers. Conkright worked with grades eight through twelve in Sierra Leone, and Namuddu with adolescents and adults in Uganda.

Results by Conkright showed that a great majority of students did not have any "scientific" explanation for the phenomena such as eclipses and rainbows. A much smaller sample of American "university town" students showed much more "scientific" knowledge about the same phenomena.

Namuddu's results showed approximately 37 percent of the adolescents believed in and were influenced by selected superstitions, while 24 percent of the adults had the same response. Both studies showed that perceptions of the world, in developing countries, may be much different than Western ideas and should be considered in curriculum development.
Physical Science

In a dissertation study involving students from Sierra Leone (N = 297), A. B. Johnson (1978) studied the relationship of chemistry and physics achievement to scholastic aptitude, general science achievement, and self-concept. Self-concept scores correlated poorly with both chemistry and physics achievement. Regression results indicated that only scholastic aptitude was a significant predictor of chemistry and physics achievement.

Lynch (1978), working with another set of students in Africa, studied the relationship of experimental work and student achievement. Using a large sample (N = 2352) of white South African students, the investigator compared achievement, as measured by matriculation exams, with frequency of experimental work and types of experiments undertaken. While there were large and significant variations in frequency and type of experiments, but that the results make little difference in matriculation exam grades.

College Level Students

Life Science

Holland (1978) attempted to relate some student characteristics to success on audio-visual modules for a course in biophysical analysis as well as to evaluate the instructional units themselves. A small sample size (28) makes conclusions tentative, but results showed previous course work in computer science and biology were useful predictors for some units. Attitude changes were not especially predictive, although students' comments were generally positive for the program.

Using 303 students in college natural science class, Dallam (1978) developed a questionnaire of student behaviors and related it to success in the classes. The instrument was developed through pilot studies and used factor analytic techniques to select the most useful items. Comparisons of responses were made between high and low class achievement, teaching method (lecture, recitation, and self-paced), and previous college GPA (high, medium, low). Several main effect and interaction effects were found within the various subscales used in the analysis. Regression analysis did, however, show that much of the variation in student achievement was due to previous GPA.

In a dissertation study on relating visualization skills to learning science concepts, Dorsey (1978) used undergraduate college students enrolled in a premedical curriculum. Using standardized tests for spatial visualization and obtaining GPA and ACT scores, the investigator calculated correlation coefficients and did an analysis of variance between school year levels. Results indicated high visualization skills for juniors and seniors as compared to freshman and sophomores and significant correlations between GPA and visualization scores in juniors and seniors. No correlation with ACT scores was found.

Simpson and Wasik (1978), in a study to relate preference towards science and biology to achievement in a biology course, selected a course especially designed for elementary student teachers. Twenty-seven students, all female, took the course and were administered pre-
and posttests on science processes and science preference. A checklist of affective behavior characteristics was also developed by the authors and used by course instructors to rate students. Results showed that science and biology preference items predicted student test performance at a significant level. Test performance was, however, not significantly correlated with end-of-course preference, as was hypothesized by the researchers. The checklist correlated highly with achievement test scores indicating a stronger relationship than was expected. The authors noted that because instructors were marking the checklist, some bias may have occurred.

Chemistry

Several articles, mainly dissertation studies, are critiqued in this section. Researchers in this area might first wish to read Fletcher’s recent review article (1977) on predicting success in college chemistry for additional references.

Two of the dissertations in the group were mainly directed toward the mathematical preparation of freshman students and the relationship of this preparation to chemistry achievement. Malone (1978) found significant correlations between ACT mathematics scores, scores on a specially constructed mathematics preparation exams, and freshman chemistry grades. She also found that some students who took a remedial mathematics course for chemistry students made significant gains in mathematics ability but not in chemistry achievement.

Perkins (1978) found very similar correlations with a researcher prepared mathematics test and a chemistry final examination. Correlations were comparable for students who had completed a previous course in chemistry and for those who had not completed such a course.

In a study comparing two sequences of chemistry courses and the relationship between admissions, chemistry placement, and chemistry course grades, Kling (1978) found low correlations of ACT scores and chemistry course grades. Correlations between ACT scores and Toledo Chemistry Placement exams were modest. Students who completed the shorter sequence of courses scored higher on final course grades than did students in the longer course sequence.

In a rather abbreviated article on nursing chemistry achievement, Mamantov and Wyatt (1978) used results from ACT scores, course grade points, and previous chemistry course to predict results on the Chemistry Achievement Examination of the National League for Nursing. In two different groups, course grade points were the most significant predictor, although no $R^2$ values were reported to check the percentage of variance accounted for by the predictors.

Boyd (1978) used several predictor variables to estimate the criterion variables of critical thinking ability, perception of science classroom activities, views toward the tentativeness of science, and attitudes toward science classes among freshmen and sophomores in chemistry and physical science courses. All of the criterion variables except attitude
were significantly related to the predictors of number of high school science course, first semester grade, natural science section of the ACT, and GPA.

**Physical and Applied Science**

Three dissertations, covering wide variations of subject matter, are included in this section. Because of the specificity of the areas, generalization beyond the specific setting of each study are tenuous.

Borja (1978) found several student characteristics such as age, gender, high school and college background related significantly with academic achievement in college physical science.

R. B. Henry (1978) found that having taken high school chemistry was a significant predictor of success in an Army materials testing program. Mental abilities and the students' perception of the importance of their high school science courses were not significant predictors.

The relationship of calculus preparation to success in engineering, chemistry, and computer science was the topic of research by Spresser (1978). Results indicated an advantage for students who had completed a course in calculus at the high school level, but when mathematics aptitude was considered, this advantage decreased.

**Women and Minorities**

In an article dealing with students at the University of Washington, Remick and Miller (1978) point out the dramatic decline in participation in mathematics and science sequences by minority students and women. While Asian-American students were similar to Caucasian students in participation rates both for math and science, the females' rates were significantly lower than those of the males in participation in advanced courses. The black and Hispanic trends were extremely precipitous, with both male and female students going from approximately 90 percent participation in biology to only 20 percent in physics. The authors noted that interviews with students had indicated that the reason that many minority students who had decided on a science career did so because of a science teacher's encouragement.

Ott (1978) surveyed female engineering students in a large number of colleges across the United States. As expected, women made up a small minority of engineering students (generally less than 20 percent) with black women especially underrepresented (only 1.5 percent). Specific recommendations were made for the recruitment and retention of female engineering students.

Graham's (1978) dissertation study centers around the characteristics of male and female students that leave science majors. It was found that students' grades had less effect on the decision to leave science for women than for men, with 35 percent of the "A" female students leaving science as compared to 4 percent of the "A" male students. A regression equation of several variables was significantly able to predict some of the characteristics of students leaving the program, although it explained only 9 percent of the variance.
Because of the wide variation in settings, generalizations concerning studies in this section are difficult to make. It does appear that in many of the studies where attitudes and achievement were used as predictors, attitudes fared poorly. Many of the instruments that were used may not be valid for the situations in which they were employed. Researchers might wish to consider using interview techniques for determining attitudes rather than strictly paper and pencil methods.

The reporting of regression equation results by including both standardized and unstandardized regression coefficients, incremental and full model multiple regression coefficients and F values would help readers more fully understand the results. The use of zero order correlation coefficients is a useful descriptive technique but assigning causality based on correlational data should be avoided.
TEACHER TRAINING AND BEHAVIOR

A useful conceptualization of the place of the teacher in science education research was presented by Okey et al. (1978). Research in their view can be divided into two phases (1) teaching skill acquisition, i.e., teacher training and teaching; and (2) skill effectiveness, i.e., performance of teachers in the field. Okey et al. claimed that most research has been conducted exclusively in one or the other of these phases. They suggested that some research questions (e.g., the type of teacher training which produces greatest classroom effectiveness) require studies that span the gap between the two phases. The studies in the year under review bear out the perception of Okey et al. Few studies were found that investigated both aspects of the research. Apparently such longitudinal studies are difficult to carry out and are therefore eschewed by researchers.

The first part of this review is devoted to the teacher training phases and the second part to the skill effectiveness phase of research on teaching.

Teacher Training

Teaching Skills

Seven studies dealt with the effectiveness of various strategies for promoting the use of different teaching skills. Four of these dealt with questioning skills, undoubtedly one of the critical teaching competencies. One of the important factors in questioning is wait-time. Longer wait-times have been shown to raise the cognitive level of both questions and responses and to improve the participation of weaker students. Rice (1977) investigated whether wait-time, the number of questions asked, and the cognitive level of those questions, would improve if preservice elementary teachers were given instruction dealing with various question-asking strategies. Ten elementary education majors were assigned at random to experimental and control groups. The experimental manipulation consisted of viewing films on questioning strategy, reading an article on the importance of wait-time, and analyzing one's own microtaught lessons. In spite of the small sample, significant results were obtained for all three hypotheses.

A study with similar objectives was performed by Esquivel et al. (1978). In this study, four treatments were employed. Students microtaught three SCIS lessons. Audiotapes were made and analyzed in one of four ways: (1) self-feedback using a questioning and wait-time critique form, (2) peer feedback, (3) supervisor feedback, and (4) self-feedback using a form unrelated to questioning. Analysis of subsequent lessons failed to reveal significant differences among the groups with regard to level of question and wait-time.

Pearson (1978) studied the effects of three different training procedures (lecture-demonstration, simulation-activity, and task-card presentation) on the questioning styles of preservice elementary education majors in teaching science processes. Only minor differences between groups were revealed in analyses of audiotaped lessons.
Chawprecha (1978) examined three strategies for improving question-asking skills among Thai high school chemistry teachers. The three strategies were (1) instructional pamphlets, (2) listening to the commenting on audiotape model lessons, and (3) listening to audiotape model lessons and classifying teachers' questions. Strategy one was found to be superior in increasing both wait-time and the proportion of higher-level questions. If this result is replicable, it may have great significance because the use of pamphlets represents a much less expensive instructional strategy than using audiotapes.

Several programs designed to influence teaching competencies have met with some success. Yeany et al. (1978) trained preservice secondary science teachers to complete a systematic self-analysis of their use of strategies for teaching data collection, processing, and interpretation. Use of the Data Processing Observation Guide to analyze videotapes of their classes as compared to a control group. In a similar study, Yeany (1978) investigated the influence of using the Teaching Strategies Observation Differential (TSOD) among preservice secondary science teachers. This instrument categorizes classroom interaction along a 10-point continuum from inductive to expository. Three experimental groups were employed. Students in one group used the TSOD individually to analyze videotapes of their microteaching, whereas in another group, both the student and a supervisor used the TSOD. A control group viewed their videotapes and analyzed them in an unstructured way. Results indicated a significant shift toward the use of inductive strategies by groups using the TSOD in subsequent microteaching sessions. This was especially true for the supervised group.

Finally, a study by O'Shea (1978) concentrated on improving the kinetic structure of lessons of high school science teachers. Subjects' sensitivity to kinetic structure was measured and then three training sessions aimed at improving structure were held. As compared with control subjects exposed to a placebo treatment, a significant improvement was recorded for the experimental group, especially for subjects whose prior sensitivity to kinetic structure was low.

Szabo and Welliver (1978) used television programs to try to improve the inquiry skills of inservice teachers. The investigators adapted Science for the Seventies (SFTS) programs for TV by including breaks for pupil teacher interaction. To investigate their impact on teacher and pupil behavior, a sample of K-3 teachers (N = 44) was divided into two groups. One group experienced the SFTS television lessons and the other was restricted to the SFTS lessons only. Teacher and pupil verbal behavior was rated on the Social Substantive Schedule scale. The authors found that 11 of the significant findings favored inquiry-oriented classroom behaviors in science teaching.

In addition to the seven studies described above, two studies reported on the effectiveness of overall teacher preparation activities on teaching skills and strategies. Minstrell (1978) described the evaluation of a science program for elementary teachers developed at the University of Washington. The evaluation was performed during summer institutes held in 1973 and 1974. Results indicated that in addition to
cognitive gains, participants gained skill in using inductive strategies, higher level questions, and increased wait-time. Cotton et al. (1978) described the use of written materials for teaching prospective teachers' process skills and implementing inquiry teaching. Students who used the materials acquired six of the skills as well as decreasing the number of closed questions asked.

The above studies indicate that self-analysis of teaching and instruction in methods can help teachers improve in the desired directions. However, the failure of a few studies to produce positive results while using essentially the same broad strategies as successful studies, indicates that more detailed analyses of the critical elements of such training techniques are required to expose the subtle factors that influence success.

E. J. Anderson (1978), in an effort to improve teacher learning outcomes, studied the effect of combined learning cues. She presented inservice teachers with a list of expected behavioral objectives and competencies and gave them a pretest. A control group was given an unrelated inventory and a list of assignments. Although significant gains were registered by teachers given the pretest and objectives, no differences between experimental and control groups were evident.

A feature of elementary science teacher preparation which may be effective in promoting the teaching of science in the schools, is the preparation of teaching materials and equipment during preservice training. Batoff (1978) investigated one manifestation of this idea. Students prepared a Unit Box, that is, a multimedia package of materials designed to implement a classroom-tested, published unit. The study, which employed classroom observations and teacher interviews, indicated that graduates of the program did in fact make good use of their Unit Boxes years after completing the program. Teachers regarded this experience as being far more worthwhile than traditional assignments involving only the writing of a unit.

How does the training of a teacher affect the achievement of his/her students? This issue was tackled in two studies. Nelson (1978) provided one group of preservice elementary majors with instruction in ways of presenting three lessons to fifth and sixth graders while a second (control) group simply received lesson plans. Significantly higher scores were achieved by students in the experimental groups. Science knowledge, as reflected by the elementary majors' science grade point average, had no effect on pupils' performance. This study indicates that specific teaching strategies are more important than a high level of proficiency in general science concepts for teaching science.

R. L. Campbell (1978) studied the effectiveness of teacher training on the achievement of first-grade pupils. Four elementary teachers received instruction on how to utilize the Individualized Science (IS) program and then used it in their classrooms. Scores of their pupils on a basic science achievement test were compared to those of four teachers using a standard first grade program. Children in IS classes with trained teachers achieved significantly higher on the test. From the research report, it is impossible to tell whether all children received IS
instruction. If this was not the case, superior pupil performance might be due to the IS program itself rather than to teacher training.

Finally, Atwood and Atwood (1978) investigated one of the central issues in competency-based teacher education, namely the retest reliability of competency tests. They administered equivalent forms of a process skills test to elementary education majors in a methods class. Procedures were as follows: (1) students were tested on three competency tasks, (2) those who failed to demonstrate competency on all tasks were retested a week later, and (3) an unannounced delayed retest was administered ten weeks later. The principal finding of the research was that most students who demonstrated a competency early in the course maintained the competency nine or ten weeks later. There was an exception to this on the metric measurement task where almost 50 percent did not. Because of the continuing interest in competency-based teacher education programs, more research in this area is essential.

**Affective Outcomes**

Positive affective outcomes are important goals of most teacher training programs. Several studies focused on this aspect of teacher training.

Barufaldi et al. (1977) studied the effect of a modern elementary science methods course containing numerous hands-on, inquiry-oriented experiences on preservice teachers' views of the tentativeness of science. A comparison group in a mathematics methods course served as a control. Results clearly indicated that the science methods students held stronger views on the tentativeness of science than did the control group.

Jaus (1978) studied the effectiveness of 30 hours of environmental education at the elementary and middle school levels. A group of inservice elementary and middle school teachers participated in a graduate-level methods class. For half the period all teachers received instruction on the environment while the other (control) group was instructed in science process skills. A 30-item Likert scale instrument (reliability = .88) was administered to both groups. Results indicated a positive attitude towards environmental education for both groups but the experimental group scored significantly higher than the control group (p<.001). Interestingly, control students expressed positive attitudes towards environmental education but were negative about their ability to engage in such instruction themselves.

Finally, Lazarowitz et al. (1978) attempted to determine whether standard science methods courses had an effect on secondary and elementary teachers' attitudes towards inquiry. The Inquiry Science Teaching Strategies, a 40-item Likert scale, was administered prior and subsequent to the methods courses to both groups. In both cases a significant (p<.001) change in favor of inquiry strategies was registered indicating that the methods' courses did influence preservice teachers' views on this matter. W. R. Brown (1977) used the results of an attitude survey given to 80 preservice elementary teachers in a science methods course as a guide in modifying the laboratory component of the course.
Inservice Training

Continuous professional renewal must be regarded as a high priority in efforts to maintain high teaching standards and ensure teacher responsiveness to changes in curriculum and objectives. A number of the studies above were concerned with the inservice training of practicing teachers. Other studies are reviewed in this section.

Two studies were related to the form and content of inservice activities. L. D. Baker (1978) contrasted two methods of inservice instruction for teachers in grades one through three. These were: (1) continuous monitoring of students' academic achievements and attitudes and (2) integrating science with language arts, mathematics and social studies. Treatment two was found to be most effective in improving teachers' attitudes and commitment toward science teaching.

Meadows (1978) studied the effect of varying three factors concerned with the structure of inservice workshops for elementary teachers. He varied the topics of instruction, the length of the workshop (full or half day), and the time at which the workshop was evaluated. Significant findings, as measured by a project-developed rating scale, indicated that the more practical curricula-text-overview content was superior to the theoretical program, that full-day programs were superior to half-day programs, and that immediate evaluation produced lower ratings than delayed evaluation.

University Teaching

Graduate student support depends, in many universities, on undergraduate enrollments. It is not surprising, therefore, that in order to help maintain enrollments, university science departments have tried to improve instruction through the adequate training of graduate teaching assistants (Armenti and Wheeler, 1978). Renfrew and Moeller (1978) surveyed Ph.D and M.S. granting institutions with regard to their teaching assistant training programs in chemistry. Fewer than half of the institutions which responded to the questionnaires had such programs. The content and time committed to training varied widely (15-35 hours). Armenti and Wheeler (1978) described a particular teaching assistant training program in physics. Their program was based on the use of videotaped microteaching sessions in which graduate students employed three common teaching modes: prelaboratory sessions, short lectures, and office-hour sessions. The course has been most effective in helping students gain basic teaching skills.

Should student evaluations play a role in merit and retention decisions concerning college instructors? Stevens (1978) argued strongly that, as currently administered, they should not. He showed that students rating the same instructor twice within a two-week period under different conditions not related to the course itself produced widely discrepant ratings. Stevens did not deny the value of student feedback for course improvement. He did conclude, however, that if merit decisions are to be made on this basis, that student-faculty evaluations should be made more precise and unambiguous, and that they should be administered under well-controlled conditions with full student understanding of their purpose.
Kosinski (1978) took another approach in judging teaching effectiveness. He tested the hypothesis that more effective teaching should reflect itself in better student achievement in general biology courses at a large state university. All students experienced the same lecture course delivered by means of videotapes, but they had different teaching assistants for the laboratory and recitation sessions. The final exam, which was a multiple-choice test based on course content of both lectures and teaching assistant sessions, was used to judge teaching effectiveness. Although it was readily apparent that the teaching assistants varied widely in teaching ability, analysis of final test scores failed to reveal these differences.

**Teacher Competencies and Needs**

A number of studies inquire about what training a prospective science teacher should receive and what competencies a teacher would exhibit as the result of such training.

Tamir et al. (1978) prepared a science teacher inventory, a comprehensive list of possible components in a program, which can be used to help design a new program or to assess an existing one. The list was divided into the following nine categories: (1) integration and sequence of academic program and educational experience; (2) curriculum beyond the separate science disciplines; (3) nature of science in historical, philosophical, and social perspective; (4) objectives, competencies, and instructional skills; (5) communication and interpersonal relations; (6) experiences in teaching; (7) evaluation and application of research; (8) continuous professional growth; and (9) assessment of skills and program evaluation.

In a series of three studies, Chiappetta and his colleagues (Chiappetta and Collette, 1978a, 1978b; Chiappetta et al., 1978) investigated the perceptions of three groups of educators on the skills necessary for secondary school science teachers. The groups were members of The National Association for Research in Science Teaching (NARST), science supervisors, and science teachers. All three studies used a modified Delphi technique and had about a 50 percent response rate. Each of the three studies produced a ranked list of 15 cognitive competencies. Although the 15 items were not the same in each study, a fair degree of overlap was present. Competencies identified were in the areas of science content, techniques in science instruction, planning and organization, interpersonal relations, and classroom environment and control.

In a similar vein, Bybee (1978) conducted a survey of science educators' perceptions of the ideal science teacher. Subjects (N = 172) were a sample of members of the Association for the Education of Teachers in Science (AETS). The survey was structured as a Q-sort in which 50 items belonging to five categories were distributed on a seven-point scale of importance according to a predetermined frequency distribution. The mean response for each category was calculated and the following ranking of importance of the five categories was obtained: (1) adequate personal relations with students, (2) enthusiasm in working with students and teaching, (3) adequate teaching methods and class
procedures, (4) knowledge of subject matter, and (5) adequate planning and organization. These perceptions had a high correspondence with those of teachers and of students determined in other studies by the same author, compared with a similar study conducted ten years previously, a notable shift in perceived importance from knowledge and planning to personal relations occurred.

Butzow and Qureshi (1978) set out to determine whether teachers regarded as being highly competent actually demonstrate such competencies in practice. Twenty-one high school science teachers were sampled from a group of teachers nominated by their peers as being highly competent. In the first stage of the study, interviews and a questionnaire were used to generate a list of 12 competencies in rank order of importance. Then 10 to 20 minutes of actual teaching of each subject was videotaped. A five-minute segment of tape was randomly selected for viewing by a panel of judges who were asked to rate the degree to which each competency was demonstrated. Five of the six most important competencies in the list were significantly demonstrated by the teachers. Three of the six least important were significantly not demonstrated. The study shows that teachers are capable of generating a valid list of competencies and that their input should be sought in generating any list of criteria for competency assessment of teachers.

Moore and Blankenship (1977, 1978) made surveys of elementary and secondary teachers' perceived needs. In both studies the Moore Assessment Profile (MAP), a list of 117 need statements, was sent to a random sample of teachers in the 21 school districts of a large southwestern county. Return rate was 55 percent. Responses were submitted to factor analyses and the most significant factors were retained. The survey of elementary teachers produced 13 significant factors, the four most important being: providing realistic science experiences, developing basic science skills, developing an understanding of the relationship between science and society, and training in science teaching methodology. These four needs describe fairly accurately the thrust of current elementary science teaching programs and the study emphasizes the necessity to continue such activities at the inservice level. The survey of secondary science teachers produced 11 interpretable factors. Although the factor structure in this study was somewhat different from that in the elementary teachers survey, three of the four most important perceived needs were identical, thus indicating that science teachers at all levels discern a number of common needs.

Other Teacher Training Studies

Three other studies related to teacher training should be mentioned.

Stano (1978) examined secondary science teacher training in Catholic institutions in the Great Lakes region and concluded that consolidation of small programs would help to build quality programs.

Rydinsky (1978) examined the consequences of a three-week elementary school practicum for preservice secondary science teachers. This case study showed that the prospective secondary teachers benefited from the
exposure to elementary children and elementary science curricula. In addition, cooperating school districts benefited from the interaction with both the prospective secondary teachers and university personnel.

Finally, Taylor (1978) produced a step-by-step guide for teachers, administrators, and architects for the designing and planning of facilities for teaching secondary school science.

Supervision

Three studies were reported which involved the role of the science supervisor. J. Harmon (1978) investigated the changing role of the public school science supervisor during this century. The study surveyed published research. The 107 identified articles were classified according to the aspect of the supervisory concept that they considered. The majority of articles were limited to opinions with little writing on conceptions of the role of the supervisor. The role of the supervisor was seen to adapt to changing conditions.

The other two studies concentrated on perceptions of the supervisory role both by supervisors themselves and by teachers. Perrine (1978) used the Supervisory Behavior Questionnaire (SBQ) which was sent to all the elementary science supervisors in New Jersey and to a random sample of their teachers. Respondents were asked to answer the SBQ twice—once with regard to their perceptions of ideal supervisory behavior. Teachers and supervisors were found to have similar perceptions of actual behavior but teachers had significantly higher expectations.

Wolfe (1978) studied supervisors' and teachers' ratings of supervisor effectiveness and responsibility in a number of tasks. Supervisors (N = 132) and teachers (N = 905) in the State of Massachusetts were surveyed. A positive correlation between responsibility and effectiveness was found for both teachers' and supervisors' ratings. Situational factors related to supervisors' effectiveness and responsibilities were also identified.

The studies on teacher training indicate that changes do occur in teacher behavior when teachers participate in training sessions, workshops, and longer programs. A well-planned program can produce both cognitive and affective outcomes. The duration of these outcomes, however, has not been determined. It is a rare study indeed that tests the stability of outcomes, and whether they last beyond a year is highly questionable.

Teacher Characteristics and Behaviors

In the section on Teacher Training we concentrated on the influence of training on various teacher attributes and behaviors. In this section we report on research concerned with the determination of teacher attributes, attitudes, skills, needs, and behaviors. Cognitive style studies are included in the section with that title.

Teacher Characteristics, Opinions, and Abilities

A personality trait assumed to influence success in teaching is self-actualization. Self-actualization should promote higher achievement
and more favorable attitudes in students. Heintschel (1970) investigated this idea with a group of 16 high school biology and chemistry teachers. The Personal Orientation Inventory was used to classify teachers as self-actualized, normal, or nonself-actualized. Students were given standard achievement tests and the Science Attitude Inventory of Moore and Sutman. The teachers' level of self-actualization appeared to be related to student achievement with students of normal teachers performing better than others. Teachers' attitudes were not consistently related to the self-actualization of their students. No significant relationships were found between students' perceptions of teacher concern for students and self-actualization, or between students' attitudes and achievement.

Rowsey and Henry (1978) assessed secondary science and mathematics preservice teacher knowledge of metric transformations and the ability to estimate physical quantities in metric units. The knowledge component was tested by a 20-item completion test. The estimation test was administered by using slides of real objects whose physical aspects the students were required to estimate. Kuder-Richardson Formula 20 reliabilities for these tests were found to be .94 and .74, respectively. Analysis of the scores of 14 science majors and 18 math majors revealed that all students performed better on the knowledge test than on the applications test. There was a significant difference between math and science students on the applications test in favor of the science students, although in both cases scores were low.

White et al. (1978) conducted a needs assessment study for the National Heart and Blood Vessel Research and Demonstration Center (NRDC). The educational staff members of this center were interested in the current status and future needs of biology teachers concerning cardiovascular disease (CVD) and associated risk factors (ARF). The 32-item questionnaire was sent to teachers throughout Texas and about half responded. The survey indicated that teachers spend about as much time as they wish on the cardiovascular system itself but spend little time on CVD and ARF. Teachers felt a need for more up-to-date knowledge on CVD and ARF and felt that the self-instructional format of materials proposed by the NRDC would be appropriate.

Sagiv (1977) investigated whether different professional groups in Israel (engineers, supervisors, and teachers) with an interest in a vocational education program had similar evaluations of the objectives of a course in the program. The evaluations of the objectives of an applied science course by the three groups were submitted to separate factor analyses. Inspection of the emerging factors indicated that the engineers and teachers both had unique objectives for the course whereas those emerging from the supervisors' responses were held in common with either of the other groups. The author concluded that the formulation of new science programs and courses should be guided by a needs assessment of future employees and teachers of students for whom the course is intended.

Ogunniyi (1978) assessed the agreement of secondary school science teachers in Nigeria with various conceptualizations of "scientific law," "scientific concept" and "scientific theory" as espoused by various philosophers of science. The teachers were found to prefer a hypo-
theoretical-deductive conception of science as opposed to inductive conceptualizations preferred by a sample of midwestern middle-school science teachers.

In contrast, a study by Stevenson (1978) found that elementary school teachers in Nova Scotia viewed science as being a factual discipline, rather than abstract or theoretical. Stevenson's survey of 504 teachers further pointed out the limited background in science of elementary teachers. This is clearly one of the reasons for the limited attention paid to science in elementary schools. Stevenson also revealed that the organizational climate of schools in Nova Scotia were not conducive to the introduction of student-centered science programs.

Further evidence for the impact of environmental factors on teachers is given in a study of K. C. James (1978). He interviewed randomly selected high school teachers in the state of New York on factors affecting their innovativeness. Teachers perceived a number of socio-economic, academic, administrative, and local factors which influenced their ability to innovate.

This conclusion was supported in a study by Shrigley (1977) who used an open-ended questionnaire to determine what barriers elementary teachers see to more effective science instruction. Shrigley divided the responses into the categories suggested by Katz for classifying attitudes. For Katz, attitudes have four functions: (1) utilitarian, (2) knowledge, (3) value-expressive, and (4) ego-defensive. Shrigley's analysis, corroborated by four other raters, found that barriers to better science teaching were overwhelmingly utilitarian, resting on environmental factors outside of the teachers themselves, i.e., with resources and administrative forces.

Price (1978) determined the characteristics of teachers and schools in the Syracuse School District that were related to high achievement and positive attitudes of children enrolled in the SCIS program. Children were most successful with teachers who were 35 years old and had 8 years of experience. Other factors related to success were middle socio-economic status, a cooperative interpersonal environment, and positive attitudes toward science by the principal.

In regard to attitudes towards science, Carl sen (1978) found that elementary teachers were more positive about life science rather than physical science. He also found that male and older teachers held more favorable views about science.

Finally, in the study of Lazarowitz et al. (1978), the investigators attempted to determine what background and biographical variables affect preservice teacher attitudes to inquiry. For secondary science students, the number of hours of science taken had a positive relationship with attitudes toward inquiry. For elementary majors, greater maturity and desire to teach were correlated with positive attitudes towards inquiry. It appears that secondary students develop positive attitudes towards inquiry through their science courses whereas elementary majors do so through their science education courses.
Jones and Harty (1978) conducted a study to investigate the instructional and classroom management preferences of secondary science teachers. Volunteer science teachers of varied background and experience (N = 44) responded to two instruments. The Pupil Control Ideology (PCI) was used to measure the degree of custodial or humanistic pupil control ideology of the teacher. Another instrument developed for the study measured students' preference for the inquiry or traditional mode of teaching. The investigators found that the PCI was moderately correlated (p < 0.05) with the "traditional" scale indicating a correspondence between traditional instructional preference and a custodial control ideology. A similar moderate correlation (p < 0.05) between the "traditional" and "inquiry" scales indicated that these two preferences were not in opposition to one another in the minds of the respondents. This result seems to imply that, while teachers have accepted the "inquiry" mode, the "traditional" instructional mode is nevertheless a viable alternative for them.

**Teacher Behavior**

Mullinex (1978) investigated the relationship between students' gain in achievement test scores and teacher behavior in 15 BSCS Green Version classrooms. Teacher behavior was measured using a checklist administered to the students. This Likert-type instrument was formulated on the basis of the classroom interaction work of Flanders. It consisted of two parts, one measuring the directness or indirectness of the teacher's verbal behavior, and the other measuring the extent of inquiry-type procedural behavior. Gain scores showed a moderate positive correlation with the verbal rating of the teacher and a low positive correlation with the procedural rating. The teacher behaviors as perceived by the students accounted for some of the variance in gain scores but, clearly, other strong factors were also involved.

Rainey (1978) studied the relationship between science teachers' verbal and nonverbal behaviors and students' attitudes toward science classes and teachers in middle school classrooms. The experiment involved 16 teachers and about 360 students. Student attitudes were measured twice with a questionnaire over a six-week period. The IDER (Indirect/Direct; Encouraging/Restricting) system was used to classify the teachers' verbal and nonverbal behavior. No significant relationships were found between teachers' behavior patterns and student attitudes.

The use of systematic observation and observer-teacher interaction to produce changes in university instructor behavior was the subject of a study by Aubrecht (1978). In the first part of her study, Aubrecht videotaped ten physics classes, each with a different instructor. Five teachers were assigned to the experimental group and five to the control group. The experimental group's teaching was observed and analyzed according to the Bellack system. Any change in teaching behavior suggested by the teacher was supported by the experimenter. At a later date the videotapes of the teaching of both experimental and control groups were analyzed. Results indicated that, although the proportion of various teaching interactions was the same for both groups, the experimental group displayed a greater frequency of interactions of all types. The second part of the study was a case study involving eight teachers. It
was an attempt to discover how and why changes in teaching behavior take place as a result of using the Bellack system and having conferences with the experimenter. Interviews revealed that the reason for the improvement was due to increased awareness on the part of teachers of discrepancies between their actual teaching practices and the way they would prefer to teach. This enabled teachers to adapt in the desired direction.

The difficulties that many elementary teachers have with teaching science are well documented. In some school systems certain teachers are appointed to teach science in a number of classrooms, thus removing this burden from reluctant teachers and simultaneously allowing the science teacher more time for preparation. Nelson (1978a) surveyed 142 teachers with the expectation of finding differences in the science teaching practices of teachers who teach their own classes only and teachers who are responsible for science in a number of classrooms. Nelson's results indicated that teachers who teach science in a number of classrooms tended to (1) teach science to each class for more time each week, (2) be more satisfied with teaching science, (3) use more student-performed activities regardless of the curriculum materials used, and (4) attend more science-inservice activities. If these findings can be shown to have a positive influence on achievement, this research may point to new directions in implementing elementary science curricula.

Many elementary science teacher programs are heavily slanted toward training prospective teachers in the basic and integrated process skills on the assumption that proficiency in these skills will increase science teaching behavior consistent with new elementary science curricula. Sunal (1978) attempted to verify this relationship. Process skills of a sample of 42 elementary science majors were measured using a shortened version of the AAAS Science Process Measures for Teachers. At the end of their methods course the Microteaching Skills in Science (MSS) checklist used to record observations of the planning and teaching behaviors of the students. Near the end of student teaching, the Survey of Classroom Activities for Science (SOCAS) questionnaire was administered to the subjects. This instrument lists activities accepted by a panel of judges to be consistent with the new elementary science curricula. Discriminant validity was demonstrated by administering this instrument to a group of inservice teachers whose behavior, consistent with new elementary curricula, was regarded as superior. The Pearson correlation between the MSS and SOCAS was found to be .71 indicating concurrent validity for these scales. However, the correlations between performance in process skills and both behavioral measures were not significant. The author concluded that affective characteristics might better explain the behaviors of elementary teachers than would cognitive skills.

Bartholomew and Podio (1978) investigated various models for effecting behavioral changes in experienced teachers. Subjects were 30 experienced science teachers. Treatments were intended to increase interactions characteristic of an investigative mode of teaching. Each teacher taught the same earth science lesson twice, received one of the experimental treatments, and then taught the lesson a third and fourth time. All four lessons were videotaped and analyzed according to a modified version
of the Verbal Interaction Category System of Amidon and Hunter. In all, there were five treatments involving the following elements: viewing a model teacher, viewing their own lesson, and studying a written model. Analysis of the verbal interaction ratings indicated the efficacy of using models to influence teaching behavior. Both the written and video models had positive effects on teaching behavior. Viewing one's own teaching alone did not appreciably alter performance. In addition to the presentation and interpretation of their study, Bartholomew and Podio presented a three-dimensional method of portraying the matrices of verbal interactions using computer graphics. As a complement to conventional statistical procedures, these three-dimensional plots constitute a powerful way of communicating complex observations in a visual way.

Classroom Climate Measures

Reports on three instruments that measured climate are reviewed in this section. Fraser (1978e) describes the modification of the Learning Environment Inventory (LEI) that was developed for measuring classroom climate for Harvard Project Physics. It was modified for junior high school individualized settings by lowering the reading level, removing a number of scales inapplicable to the individualized classroom, and adding a scale related to "individualization." The modifications were performed by a committee of experts to ensure content validity. The scales were then administered to seventh-grade science students in both individualized and traditional classrooms. Nondiscriminating items were removed to reduce the test of 55 items. Alpha reliabilities for the nine scales ranged from .50 to .80. Acceptable discriminant validity between scales was demonstrated. The efficacy of the instrument is illustrated by the finding that individualized classrooms were rated higher than traditional ones on the scales entitled environment, satisfaction, and individualization.

Kyle (1978) developed an instrument for recording and analyzing student and teacher behavior in the laboratory. The instrument has separate categories for student and teacher behavior and involves recording categories using single-letter codes. Codes indicating to whom the behavior is directed and the sex of the person to whom the behavior is directed are also included.

Fagan (1978) determined the interrelationships among four classroom observation systems—two affective (Flanders and Hough) and two cognitive (Parakh and Brown). A collection of 30 high school science teachers were videotaped for 50 minutes in their classrooms and the tapes were coded according to all four schemes. Significant correlations were found between the two affective and the two cognitive schemes. The study was a statistical, exploratory one that might be used as a model to refine the numerous observation schemes presently being used to observe teaching behavior.

In the section on Teacher Behavior, three more instruments relevant to classroom interactions are described. Sunal (1978) described a checklist and a questionnaire related to behaviors in teaching elementary science curricula. Mullinex (1978) described a student questionnaire based on Flander's interaction approach.

Most of the studies included in this section were surveys on teacher characteristics and behaviors. Many served very utilitarian purposes of
finding out particular characteristics of teachers in specific programs or specific locations. Other studies, that had broader and more universal goals, frequently used small and limited samples that makes generalization questionable.
EVALUATION IN SCIENCE EDUCATION

The majority of studies included in this summary could be interpreted to be evaluation studies. This section will include the same major divisions that Renner et al. used in the Summary of Research in Science Education-1976. These are evaluation surveys, curriculum evaluations, affective studies, instrument development, and evaluation methodology. There is some overlap between the categories of evaluation, particularly in the area of affective studies. This was unavoidable because many curriculum studies evaluated both cognitive and affective outcomes.

Evaluation Surveys

Three large-scale evaluation surveys were reported in 1978. Data were collected during the 1976-1977 school year. The largest of these was the assessment of 9, 13, and 17-year-old American students by National Assessment of Educational Progress (NAEP) (Crane, 1978). This was the third such survey conducted. The first one took place during 1969 and 1970 and the second one during the 1972-1973 school year. Overall, the different questions were asked in the cognitive assessment of science at the three age levels (N = 72,000). The questions were developed according to a two-dimensional matrix. One dimension, the classification of the cognitive abilities required to answer specific science questions, was divided into four categories: (1) knowledge, (2) comprehension, (3) application, and (4) analysis, synthesis, and evaluation. The other dimension divided the domain of science into three major areas: content, process, and science and society. The questions illustrate various topics of the biological sciences, earth sciences, physical sciences, and integrated topics such as equilibrium, evolution, and probability.

The following findings are taken directly from the report.

Age and Grade Level. Performance levels of most of the reporting groups were highly consistent across the three age levels assessed. That is to say, groups that tend to perform above or below the national average at one age tend also to do so at other ages as well (p. 19). One would expect that achievement would improve with the greater amount of schooling. The data support this expectation for all the groups that National Assessment describes. The amount of schooling is positively related to learning for all groups. White third graders performed at the same level as Hispanic fourth graders and above the level of black fourth graders. White seventh graders outperformed black and Hispanic eighth graders. White tenth graders performed as well as Hispanic twelfth graders and above the level of black twelfth graders (p. 29).

Racial/ethnic background, sex, and region. The overall percentages of white students and male students who answered science questions correctly are consistently above the national levels at each age (p. 20). Whites who are male, go to schools that are at least 60 percent white, are at or above the modal grade for students their age, reside in advantaged urban communities and have access to at least four types of reading matter in their homes all performed above the
national average at their age. However, whites who are at low grade levels, come from disadvantaged urban communities and have less than three types of reading matter in their homes all performed below the national level for their age (p. 20). Hispanic students at the modal grade level tended to perform a little below the national average for their age. Among black students, the achievement of those living in advantaged urban communities (high metro) is near or above the overall performance of their age mates (age 13). The availability of reading matter in the home appears to be related to achievement as well. The percentage of black students at each age answering science questions correctly consistently increased with an increase in the reported types of reading matter in the home. Not all female students performed below the national level for their age. The percentages of females in the Central and Northeastern states, those who are white and those at or above the modal grade for students at their age answering science questions correctly tended to be at the national average. Females whose parents are educated beyond high school performed above the national average at all three ages (p. 20). Differences in the performance levels of male and female students were particularly noteworthy for several of the areas described in the science assessment... Although males had an advantage in almost every area described in the assessment, that was not the case in the decisionmaking and science-and-self areas. In the areas of physical and earth sciences, the male advantage is particularly large (p. 20).

**Socioeconomic status.** The data support the findings of other studies showing that the various measures of socioeconomic status are all highly related to achievement. Students whose parents have a post high school education, who have at least four types of reading material in the home or who live in urban communities with a high proportion of professional or managerial residents (high metro) all achieved well above the national level for their age. Conversely, students whose parents have not graduated from high school, who have less than three types of reading matter in the home or who live in urban communities where a high percentage of residents are on welfare or not regularly employed (low metro), all achieved well below the national average for their age. These patterns are consistent at all three ages (pp. 27, 30).

Two surveys were conducted in 1977 with NSF support. The Report of the 1977 National Survey of Science, Mathematics, and Social Studies Education (Weiss, 1978) presents data and conclusions concerning the following questions:

1. What science courses are currently offered in schools?

2. What local and state guidelines exist for the specification of minimal science experiences for students?

3. What texts, laboratory manuals, curriculum kits, modules, etc., are being used in science classrooms?
4. What share of the market is held by specific textbooks at the various grade levels and subject areas?

5. What regional patterns of curriculum usage are evident? What patterns exist with respect to urban, suburban, rural, and other geographic variables?

6. What "hands-on" materials, such as laboratory or activity centered materials, are being used? What is the extent and frequency of their use by grade level and subject matter?

7. What audio-visual materials (films, filmstrips/loops, models) are used? What is the extent, frequency and nature of their use by grade level and subject area?

8. By grade level, how much time (in comparison with other subjects) is spent on teaching science?

9. What is the role of the science teacher in working with students? How has this role changed in the past 15 years? What commonalities exist in the teaching styles/strategies/practices of science teachers throughout the United States?

10. What are the roles of science supervisory specialists at the local, district, and state levels? How are they selected? What are their qualifications?

11. How have science teachers throughout the United States been influenced in their use of materials by Federally-supported inservice training efforts in science?

Because of the comprehensiveness of this document (approximately 400 pages) and the limitation of space in this summary, no attempt will be made to report the conclusions. It is highly recommended that science educators carefully study the findings.

The other NSF study of interest to science educators is Case Studies in Science Education, Volumes I and II (Stake and Easley, Jr., 1978). It is a collection of field observations of science teaching and learning in American public schools during the school year 1976-77. The study was undertaken to provide the National Science Foundation with a portrayal of current conditions in K-12 science classrooms to help make the foundation's programs of support for science education consistent with national needs.

Eleven high schools and their feeder schools were selected to provide a diverse and balanced group of sites: rural and urban; east, west, north and south; racially diverse; economically well-off and impoverished; constructing schools and closing schools; innovative and traditional. They were finally selected so that a researcher with ample relevant field experience could be placed at each. To confirm findings of the ethnographic case
studies and to add special information, a national stratified random sample of about 4000 teachers, principals, curriculum supervisors, superintendents, parents, and senior class students were surveyed. Survey questions were based on observations at the eleven case study sites...

In the principal findings it was noted that each place was different in important ways, that each teacher made unique contributions. Nationally we found that science education was being given low priority, yielding to increasing emphasis on basic skills (reading and computation). Still, the CSSE-high-school science faculties worked hard to protect courses for the college-bound, with many of these courses kept small by prerequisites and "tough" grading. Only occasional efforts were made to do more than "read about" science topics in most elementary schools. Although ninth grade biology and eighth-grade general science flourished, general education aims for science instruction were not felt vital at any level. Seldom was science taught as scientific inquiry—all three subjects were presented as what experts had found to be true. School people and parents were supportive of what was chosen to be taught, complaining occasionally that it was not taught well enough. The textbook usually was seen as the authority on knowledge and the guide to learning. The teacher was seen to be the authority on both social and academic decorum. He or she worked hard to prepare youngsters for tests, subsequent instruction, and the value orientations of adult life. Though relatively free to depart from district syllabus or community expectation, the teacher seldom exercised either freedom.

The authors noted that because of the nature of the study, their summary is an "oversimplification of the circumstances observed by the field people." We highly recommend careful study of this two-volume work.

Curriculum Evaluation

Curriculum development and evaluation is one of the cornerstones of the enterprise of science education. Nevertheless, the demise of the national curriculum projects has substantially reduced interest in this area. What work has been done still centers around the "new" projects which are already not so new. Much effort was devoted to local implementation or adaptation although some researchers have maintained an interest in overall issues of curriculum and implementation.

Historical Perspective

A number of studies were devoted to the historical development and philosophical aspects of curriculum. Two studies explored the interaction between science curriculum and socio-historical events. Tkach (1978) examined the impact of events on the grade VI science program in Alberta between 1912 and 1963. Kim (1978) studied the influence of Sputnik and the Vietnam War on graduate science programs at a midwestern
university. Surprisingly, he concluded that there was no selective influence of Sputnik on science as opposed to nonscience programs after Sputnik, while nonscience programs declined more than science programs during the Vietnam period.

In another historical study, Alexander (1978) traced the development of engineering education in the United States, including the impact of political and social factors. He discussed the emergence of both private and public schools of engineering and speculated on future developments.

J. W. Fuller (1978) performed a historical and descriptive study of the contribution of black Americans to science. He discovered that most teachers never teach the contributions of black American scientists in their classes and gave recommendations to remedy this lack.

McCullough (1978) explored ways in which the organismic philosophy could be employed in developing science curricula and integrating science with the overall curriculum. The organismic framework can also be used in teacher training.

**Dissemination**

Seven studies were reported relating to the dissemination of elementary science curricula. Findings in four of these were discouraging. Brown (1978) carried out a questionnaire survey among administrators, supervisors, and teachers in California. The study concluded that the emphasis given to elementary science is extremely low and that the new materials have not been successfully implemented.

Gullickson (1978a), commercial television and direct mail were used to inform people of the existence of a call-in information center to learn about specific curricula. It was found that principals and teachers responded to the campaign by calling in, whereas members of the public overwhelmingly did not. A survey indicated awareness of the information center but familiarity with the curricula did not increase markedly during the period of the study. A later survey (Gullickson, 1978b) confirmed the low level of familiarity with elementary science curricula, especially among the general public, and suggested that traditional modes of communication are inadequate to the task. More effective methods should be sought.

The effect of an NSF-sponsored workshop on implementation of the Elementary Science Study (ESS) in two school districts in Washington state was studied by Craven (1978). Participation in the workshop was found to improve implementation. However, evidence of the decay of this effect was observed.

Koller (1978) reported on the use of three techniques in the implementation of SCIS in a school district in Washington during 1975-1976. The three methods were (1) summer workshops for principals and teachers; (2) school-year in-service activities for teachers, and (3) consultant services for principals and teachers. The study indicated that schools whose teachers had attended workshops tended to order more materials from the Resource Center set up to aid in curriculum implementation. Teachers,
principals and consultants were generally satisfied with the effectiveness of the implementation effort.

Eisen (1978) discussed the strategy used in implementing Matal elementary science project in Israel. The strategy depends on the situating of trainer teachers in project schools. It was hoped that these highly trained teachers would influence other teachers to adopt the philosophy of Matal and create the classroom climate encouraged by Matal. Comparisons were made among Matal trainers, Matal teachers, non-Matal teachers in project schools and nonproject school teachers. Results were generally in the direction expected.

The facets influencing dissemination of three new science curricula in Wisconsin were investigated by Stolsmark (1978). It was found that early adopters of the new curricula were those districts with the highest ratio of central office administrators to students and with the most differentiation (and greater size). Other variables such as presence of a science specialist, per student expenditure, population density, and urbanization had substantially lower impact than the two variables mentioned above.

In a similar vein, Pollard (1978) charted the process of curriculum implementation as it occurred in a small rural community. The study demonstrated similarities of the problems encountered with those in other districts. Factors influencing implementation that were pointed out included lack of adequate facilities and materials, and adequate teacher preparation.

White (1978) reviewed the status of elementary science education in the state of Virginia. Among other aspects he described patterns of curriculum adoption in the state and the impact of adopting accreditation standards.

As is clear from some of the above mentioned studies, curriculum implementation and inservice training often go hand in hand. One commendable inservice-implementation program has been established in the state of Iowa. Sheldon and Yager (1978) described the workings of the Iowa Alternative Strategies and Services for Improving Science Teaching (ASSIST) program. The program, heavily supported by the NSF and the Iowa State Legislature, functions through the Science Education Center at the University of Iowa and sixteen regional centers each led by a coordinator. The Science Education Center is responsible for general administration of the program, for needs assessment and evaluation and operation of a resource center. The Iowa-ASSIST concept combines central coordination and leadership with regional involvement. The program has had an enormous impact as regards inservice assistance and adoption of new curricula in the state of Iowa.

Finally, three studies of significance for implementation are reported. The first study attempted to find factors related to the amount of time devoted to sex education in school districts in Illinois and Missouri (Ehren, 1978). The amount of sex education was found to be correlated with the educational level of the community.
P. A. Johnson (1978) attempted to set up parameters for the cost-effectiveness of various subjects including science for grades K-12 in a single school district. Cost-effectiveness varied for different grades. Interpretation of the C/E ratios can only be effective as a trend over time. The results of this study are base-line data only, and future work will elucidate their validity for making comparisons.

Parental support for curricular innovation might be an important aid to success in some communities. With this in mind, a study by Hall (1978) is relevant. He used a tape-slide presentation to inform parents about a junior high program. A questionnaire was used to determine parents' understanding and attitudes regarding the science curriculum. The group of parents who had attended the presentation was compared with parents who had not attended. Significant differences in favor of the attending group for knowledge, comprehension, and attitudes indicated that this method of informing parents might be emulated more widely.

Curricular Objectives

Shami and Hershkowitz (1978) surveyed students, parents, members of lay public, and educators to determine the importance they attributed to 27 goals of public education in the state of Maryland. "Knowledge of Science Concepts" was ranked low on the list. On the other hand "Ability to Apply Knowledge and Skills to the Solution of Real Life Problems" received a very high ranking. The authors suggested, therefore, that scientific concepts be taught in such a way as to be related to real-life problems.

Two studies were related to the change in science education objectives over time. Ogden and Roy (1978) investigated objectives of earth science in periodical literature during the period 1918-1972. This article follows prior articles by the first author related to chemistry and biology and reported similar findings. First, the authorship of articles shifted from secondary educators to college instructors over the period. Second, knowledge level objectives decreased over the period although they remained important. Process objectives remained steady while objectives of attitude and interest type gradually fell.

Following up Ogden's previous work, Fraser (1978a) used content analysis to explore changes in objectives cited in the literature for science education as a whole over the period 1932-1974. He found little change in emphasis given to the various categories prior and subsequent to 1960.

Development and Evaluation

The era of large federally-funded curriculum projects is over for the time being. It is not surprising that those development projects which were reported were of limited scale. Five dissertations involved the development and evaluation of such projects. Shapiro (1978) developed a multimedia, independent study, mathematical skills course for allied health students. Students using the materials proved superior to students taught by the regular method, and parts of the project have been adopted as part of the regular program at the university.
Pickler (1978) designed and tested an integrated winter ecology program. Evaluation of the program demonstrated gains in knowledge and in attitudes towards the environment among a pilot group of students.

Womack (1978) developed and tested an audio-tutorial laboratory block on spiders. It was well received by students.

A preparatory chemistry course for undergraduates was developed by Weerasooriya (1978) for aiding underprepared students prior to entry into freshman chemistry courses. Students who participated in the preparatory program proved more successful in freshman courses than students who entered these courses directly.

Stewart (1978) made recommendations for inclusion of specific topics in introductory biology courses for science and nonscience majors on the basis of their interests.

Program Evaluation

The impact of the large curriculum projects is still being felt and reflects itself in a number of studies designed to assess outcomes of these curricula with various populations of students under various conditions. Some studies compare the "new" program with traditional approaches while others do not. These outcome studies are surveyed according to school level.

Elementary Programs

A number of studies demonstrate the effectiveness of new elementary science programs. Six were concerned with Science--A Process Approach (SAPA). Calvey (1978) showed that SAPA with its process orientation was superior to a more traditional program in promoting operative comprehension among eighth graders. However, differences were found for sixth graders who were in a transitional stage between concrete and formal reasoning.

Krebs et al. (1978) demonstrated that competency in the first five exercises of K-level SAPA could be achieved by preschool children. In four 10-minute sessions per week for a period of seven weeks, 90 percent competency in 90 percent of the measures was achieved by the predominantly middle class children.

Shaw (1978) showed that SAPA II was more effective in teaching problem-solving skills in science than a traditional program with the same subject matter. Moreover, transfer of these skills to social studies type problems was found. Studies by Quorn and Yore (1978) and Merricks and Crocker (1978) relating reading skills and SAPA are discussed in the section on Reading and Language Skills.

The Science Curriculum Improvement Study (SCIS) was the subject of six studies. Bowyer and Linn (1978) reported on the long-term effects of studying SCIS as compared with textbook elementary science materials, as regards processes of science (variable, analyzing experiments, histograms, reasoning from data) and understanding of major science concepts.
(relative position, energy source/energy receiver, energy transfer, life requirements, solution and evaporation). A pencil and paper task was developed for each of the items listed above and the validity of the Scientific Literary Test (SLT) was established by matching the tasks to the objectives of the SCIS program. Items related mostly to the fourth-through sixth-grade objectives of SCIS. Test-retest reliabilities were established using a group of 24 children who had been taught using SCIS ranged from .85 to .97. The test was administered to 531 sixth graders from rural Michigan, some of whom were enrolled in the SCIS programs. A multivariate analysis of variance established a significant overall difference ($p < .0001$) in favor of girls in tasks requiring greater verbal ability. This was interpreted as being consistent with the psychological literature which shows superior verbal skills for girls over boys at the 12-13 age level.

Disadvantaged rural fifth graders were the subjects in a study by Hendricks (1978) using SCIS materials in comparison with a textbook approach. Results favored the SCIS approach in the affective domain. Only—SCIS pupils displayed a more positive attitude and a greater curiosity towards science.

Hofman (1977) compared the attitudes towards science and scientists of eight-year-olds exposed to SCIS and the textbook program, Concepts in Science. No significant differences were found between the two groups although both apparently had positive attitudes.

Krockover and Malcolm (1978) also studied the effect of SCIS on children's attitudes toward science. They found that the attitudes of children in grades three to six enrolled in the SCIS program were more positive than attitudes of children in a conventional program. However, attitudes toward science did not improve as much as anticipated.

Rachelson (1978) investigated the implementation of two aspects of scientific inquiry in curricula—hypothesis generation and hypothesis testing. He asserted that the testing component is easier to incorporate into classroom practice. Examining the Science Curriculum Improvement Study (SCIS) and Essense I curricula, he concluded that SCIS incorporated hypothesis generation but in an inconsistent way, whereas Essense I did so consistently.

In another investigation, Krockover and Malcolm (1977) studied the impact of SCIS on children's self-concept. Changes in certain areas of self-concept were registered in pre- and posttesting among students in grades three through six over a four and one-half month period.

Only one study evaluated the Elementary Science Study (ESS) materials. Davies and Ball (1978) investigated the effect of ESS materials on selected science skills including communicating skills, observing, and inferring with educable mentally retarded at four age levels. Results differed for different age levels except for verbal skills where a significant positive effect was in evidence for all levels. These results indicate the need for further research on the use of regular science materials with EMR students.
**Secondary Programs**

During the year in review, three dissertations comparing the Intermediate Science Curriculum Study (ISCS) program and more traditional approaches were reported. Two of these indicated a more positive attitude towards science for ISCS students (Guffy, 1978; Del Barto, 1978). In addition, Del Barto did not detect any differences in achievement between ISCS and traditional groups. On the other hand, Matherne (1978) found that students who had gone through a traditional ninth-grade program scored significantly higher on scientific information tests than did ISCS students.

Interest in the effects of the major curriculum projects at the senior high school level appears to be waning. Three studies examined the impact of the Biological Science Curriculum Study (BSCS). Hart (1978) investigated the impact of BSCS Green Version on environmental attitudes and understanding. Ecology comprehension and environmental attitudes were found to be higher for students enrolled in BSCS than for nonbiology students.

Hayden (1978) found that the attitudes of students exposed to BSCS minicourses were to be more positive than those of students who studied by the traditional lecture-discussion method. No difference between the two groups was found, however, on measures of achievement or critical thinking. Tamir and Lunetta (1978b) described the impact of the introduction of BSCS in Israeli high schools on students' inquiry performance and curiosity. Results for whole classes seem to indicate a high correlation between inquiry performance and level of curiosity but this was not reflected by the performance of individuals.

Comparisons among traditional and new biology programs were made in two studies. Fido (1978) analyzed biology materials according to levels of biological organization. He criticized attempts by Schwab in the United States and Crossland in the United Kingdom to assign relative weights to the coverage of different levels of organization in various texts. According to Fido these methods have led to unwarranted contrasts between "traditional" and "modern" curricula as compared to the results he obtained via a more justifiable approach.

Lowery and Leonard (1978a) compared frequencies and styles of questioning in the three BSCS texts and Modern Biology, another commonly used American high school biology text. Differences between the ratios of questions to sentences were found among all four texts. All four texts were found to make little use of "higher" level questions and all, except the BSCS Blue Version, concentrated on lower order science processes in their questions.

Tamir and Lunetta (1978a) developed a task analysis instrument for analyzing laboratory materials as well as an instrument to measure the degree of integration of laboratory work with the text. These were used to examine laboratory inquiries in BSCS Yellow Version. Laboratory exercises were found to precede the text and to be integrated with it, giving students the opportunity to explore and then consolidate their knowledge. Higher level inquiry tasks such as formulating problems and designing and carrying them out experiments were found to be underrepresented.
In a questionnaire survey, Schlenker and Qureshi (1978) found that very few teachers in the United States teach secondary marine science but that there had been a dramatic increase in this in seven states.

In the area of physics, Quattropani (1978) compared the two federally-funded physics programs, Harvard Project Physics (HPP) and the Physical Science Study Committee (PSSC), on students' understanding of the relationships among science, technology, and society. Students who studied the historically oriented HPP course improved their understanding over the experimental period whereas the PSSC students did not.

In a descriptive survey Rolf (1978) compared the content of adopted texts with guidelines in Texas. He also investigated other issues related to earth science teaching and teacher training and made recommendations to improve earth science teaching in that state.

In a different type of experiment, Asun (1978) attempted to examine whether graduates of the Nigerian secondary school science program could apply their knowledge to situations outside the classroom. Results indicated great difficulty in doing so. The reason postulated for this was content orientation of the West African science program.

College Programs

Hendrix and Suttner (1978) surveyed geology educators on summer field course offerings and found that little change has taken place in the last ten years except that costs and enrollments have risen sharply. There was general agreement on the philosophy and goals of such courses.

This section has included a wide variety of evaluation studies, many of which compare one program with another. Generally speaking, the studies have shown that there are few cognitive gains for students enrolled in "new" programs versus "old." The "new" programs do appear to produce more positive attitudes. Several questions must be answered before too much hope is derived from these results. Is the change in attitudes due to the instructor's enthusiasm about the program that is conveyed to the students? Few studies control for this. Do the positive attitudes persist beyond the course? Very few studies test students' attitudes several months after the course has been completed. Both of these questions must be answered before confidence can be placed in these findings.

Affective Studies

Attitudes towards science and science teachers and how these attitudes might be improved is of concern to science educators. One indication of the possible magnitude of the problem is reported in a study by Failbanks (1978) that indicated that non-science majors with negative attitudes toward science delay fulfilling their science requirements.

Studies which have a bearing on attitudes and the relationship between science and society are reviewed in this section. Three groups of studies
are surveyed: (1) studies which simply measure attitudes or that relate attitudes to other variables, (2) studies concerned with the question of science and society, and (3) attitudes relating to environmental issues.

**Attitudes and Their Relationship to Other Variables**

Only two studies reviewed reported attitude measurements without attempting to relate them to other measured variables. Palmer (1977) presented eighth-grade science students in Wisconsin with an inventory of items comparing 15 categories of science (5 in each of earth science, biology, and physical science). Students tended to rank earth sciences and biological sciences ahead of the physical sciences. Reasons offered for this were a possible bias in the statement of categories in the inventory itself, transfer of preferences from teachers to students, and the maturational stage of students in the study.

In a study carried out in England, Smith (1978) measured the attitudes of secondary school students toward science as a career choice and their perception of science as a professional activity. Students' attitudes were unfavorable, leading Smith to posit this as a contributing factor to the poor social status of scientists in Britain today.

Several studies attempted to relate attitudes to other variables: cognitive, affective, and social. The relationship between knowledge and attitudes is addressed in a number of studies. Fortner (1978) obtained a relatively high correlation (.43) between knowledge and attitudes to marine science among tenth-grade students surveyed in Virginia. J. H. Finley (1978) found a lack of knowledge concerning recombinant DNA and genetic engineering and apprehension about advances in genetic technology among college students. On the other hand, students were interested in knowing about their own genetic make-up.

Nagy (1978) sought to find a distinction between the affective dimensions of feelings and beliefs in Moore and Sutman's Scientific Attitude Inventory (SAI) where beliefs have a greater cognitive content than feelings. In this view, feelings are expected to display a stronger correlation with achievement, unless the beliefs are specifically part of the course content. Nagy administered the SAI to 97 ninth-grade students. Using a cluster analysis technique, he was able to divide the SAI into five subtests which differed from the subscales as devised by the test constructors. He showed that the clusters designated as measuring feelings correlated independently with achievement, whereas the correlations between the belief subtest and achievement were moderated almost entirely by student ability. These results were not obtained for the subtests as defined by Moore and Sutman, leading the author to conclude that the cluster analysis procedure is superior to an a priori method for deriving the subscales of an attitude test.

The use of attitudes as predictors of success was demonstrated in two studies. Visco (1978) showed that there existed a correspondence between values as measured by the Rokeach Value Survey of students successfully completing three different high school courses (one of which was Project Physics) and those of students who elected these
courses. Visco commented on the possible use of values as indicators in predicting academic success.

In contrast, Stevens and Atwood (1978) found that science interest scores were poor predictors of achievement of science process skills among junior high school students. Pretest scores on the process test proved to be a much stronger predictor. A troublesome finding by Stevens and Atwood was the drop in interest in science they found over the school year. They explained this drop by a "let-down" effect as the school year progresses. The idea bears investigation lest we become disheartened by an apparent lack of success in arousing student interest in science.

In the same vein, a study by Crater (1977) failed to detect any change in student attitudes for high ability secondary students who participated in a summer program related to nuclear and environmental science. It should be noted, however, that for this selected group of students, attitudes toward nuclear science were high on both pre- and posttests.

Kruglak (1978) found that he was able to produce a negative intended attitude change toward astrology in students enrolled in an astronomy course. It may be that negative attitudes are easier to create than are positive.

Two issues dealt with attitudes in relation to social factors. Fraser (1978) explored the relationships between students' perceptions of their science classroom learning environment and attitudes to four different sources of scientific information: experiments, books, experts, and teachers. More favorable learning environments were found to promote positive attitudes towards experiments as a source of information, while less favorable environment promoted more positive attitudes to the more authoritarian sources of information.

Eggen (1978) examined student nonverbal behavior in the classroom. He found that, in a lecture mode of instruction for middle school classrooms, six nonverbal behaviors could be used to predict a positive or negative attitude toward the science teacher. Sensitivity to nonverbal behaviors might help the teacher promote more positive attitudes.

A characteristic of most attitude-related research in science education is the lack of a theoretical model of attitude change from which to work. Shrigley (1978) attempted to remedy this by adapting the learning theory approach for use in science education. One component of this model was the idea of "persuasive communication" as an instrument in promoting change. In his study, Shrigley surveyed the literature to establish general guidelines for performing research on attitudes in science education according to the learning theory model.

Science, Society, and Morals

A number of researchers concerned themselves with the overall relationship between science and society and the interaction with moral
issues (Clements, 1978; Dispoto, 1977; Magrane, 1978; Offurum, 1978; Wilhelm, 1978). Clements and Offurum, in particular, addressed the problem of the exploration of models of science teaching which promote moral and social responsibility. Clements tested his model for the encouragement of social action by conducting a workshop for 20 teachers. His results indicated both cognitive and moral growth among participants. Clements proceeded to further refine his model on the basis of these results and introspection. Offurum, in a theoretical dissertation, proposed a model for science teaching that simultaneously dealt with the dimensions of knowledge and moral responsibility.

A values-education approach was used by Magrane (1978) to present bioethical issues to biology students. Students enrolled in the course scored higher than did a randomly selected control group on tests of moral reasoning and biological values, thus indicating that such courses can contribute to the personal development of students.

In some cases, scientific theories contradict the values of members of society who may oppose the teaching of such theories in schools. A case in point is the belief of creationists and their opposition to the inclusion of evolution in the curriculum. Wilhelm (1978) analyzed the nature and results of attempted regulatory action in the United States between the years 1928 and 1978. He concluded that the issue is not resolved and that future action on the part of creationists may be expected.

One of the most attractive theories for interpreting moral behavior is that of Kohlberg who proposed six stages of moral reasoning, to parallel Piaget's stages of cognitive development. Dispoto (1977) conducted a study to relate levels of moral reasoning to three aspects of environment-related behavior (environmental knowledge, emotionality, and activity). His subjects were 140 undergraduates at an eastern state university. Level of moral reasoning was estimated using Rest's Defining Issues Test. Each subject received two scores, one for conventional values choices and one for principled values choices. Dispoto's analysis showed that the morality scores displayed varying relationships to the three environmental variables and explained only 5 to 10 percent of the variability in them. Humanities majors were as concerned about the environment as were science majors but science majors knew and did more about environmental problems. The two groups did not differ in the level of moral reasoning. There were, however, differences between the two groups with regard to the correlations between morality and environmental variables. For science students there were no important correlations, whereas for humanities students there were a number, accounting for about 10 percent of the variance. The surprising, albeit tentative, implication of this study was that environmental knowledge rather than moral level was a good predictor of environmental activity.

Environment

The area of environment is one where affective outcomes are often considered more important than cognitive ones. Hence, the development
suitable instruments for measuring attitudes has occupied some researchers' time. Hopeful that better knowledge and understanding will promote desirable attitudes, other researchers have sought to determine relationships between knowledge and attitudes.

Two articles under review concern themselves in detail with the problems of environmental attitude measurement. Hart and McClaren (1978) investigated the attitudes of twelfth-grade students (N = 382) in British Columbia toward a number of environmental issues using an instrument developed by Steiner and Barnhart. The Inventory of Societal Issues includes 60 items covering seven factors. The seven factors involve attitudes toward: value of human life, relationship between science and nature, cooperation with nature, population, personal involvement with societal problems, ability of science to deal with environmental problems, and individual freedom. The attitudes of Canadian students were found to differ very little from the attitudes of Oregon students who took the test when the instrument was developed. This prompted the authors to conclude that using the inventory represented a valid way of assessing environmental attitudes. Attempts to relate differences in environmental attitudes to variables such as location (urban or rural), gender, academic achievement, and non-economic status yielded a small number of statistically significant results capable of explaining a small fraction of the variance in scores.

A second study described procedures for validating the Moyer Unobstrusive Survey of Environmental Attitudes (MUSEA) (Moyer, 1977). The purpose of the instrument is to assess the inner feelings of students that they may not otherwise divulge. It uses techniques designed to lead the respondents into projecting their true inner feelings toward three environmental themes: pollution, population, and ecological relationships. The MUSEA can be completed in 25 minutes and has a reading level of 7.2 (Frye Readability Formula). The coefficient of internal consistency for the instrument and its subscales range from .60 to .80. It appears that this should be a useful instrument for use by both teachers and researchers to determine the effective outcomes of environmental courses.

Four studies were concerned with the relationships between knowledge and attitudes in the environmental area. Brock (1978) developed an affective instrument and a cognitive instrument for assessing ecological attitudes and knowledge of college students. The instruments were pilot-tested with 280 students in introductory biology courses. The analysis revealed a trend of low to negative correlation between attitude and knowledge level questions but an increased positive correlation at comprehension and higher level questions but an increased positive correlation at comprehension and higher levels.

Andrews (1978) investigated interrelationships between cognitive and affective attributes and participation in activities in an outdoor environmental education program for sixth graders. Outdoor Biology Instructional Strategies (OBIS) materials were used. He found a positive correlation between attitude toward environmental concepts and knowledge and between attitude and involvement in activities related to the concepts.
In a study of the influence of knowledge on attitudes about wildlife, LaHart (1978) tested 1300 eighth graders in Florida. Results indicated that knowledge played a relatively minor role in predicting attitudes.

In the study of BSCS and non-BSCS biology students, Hart (1978) found moderate correlations between environmental attitude and environmental information (.21) and environmental attitude and environmental comprehension (.33). The greater correlation for comprehension than for information seems to confirm Brock's results cited above. The seeming lack of unanimity in studies on the correspondence between knowledge and attitudes may be due to psychometric problems in measuring attitudes adequately. The problem merits further study.

Affecting change in attitudes may be possible if innovative techniques are used to enable students to investigate environmental problems in an involving way. Evidence for this possibility was found by Dunlop (1978) in a study to investigate the energy-related attitudes of several different groups of science students and science teachers before and after working with the energy-environment simulator for approximately one hour. During the interaction with the simulator, relevant variables were manipulated and the results indicated to the students. Significant changes in attitude, due at least in part to the interaction with the simulator, were reported for the different groups mentioned.

Who should handle environmental education, science teachers or social studies teachers? According to Pettus et al. (1978), it should be teachers who themselves have more favorable environmental attitudes. In their study, the environmental attitudes of 75 science and 75 social studies teachers were examined. The instrument used was adapted from that of Steiner and Barnhart and validated by submission to 25 judges. The final set of 35 items was submitted to factor analysis yielding three factors labeled (1) need for responsible action, (2) individual privileges and need to prepare for the future, and (3) need to protect the environment. The factor scores of subjects were cluster analyzed to yield nine clusters of teachers with similar patterns of response. Inspection of the clusters indicated two clusters dominated by science teachers and two dominated by social studies teacher clusters, indicating that perhaps science teachers should be entrusted with dealing with environmental issues.

Trollinger (1978) studied the effectiveness of games as compared with other strategies in teaching environmental topics. Since affective objectives are of such great importance in environmental education, we might be prepared to use games in place of more formal techniques if they enhance affective outcomes, even at the expense of cognitive ones. Although students enjoyed participation in the games, unfortunately no comparison of affective outcomes was measured in the study. Because of the diverse nature of the studies included in this section, few conclusions can be drawn. It appears that achievement and attitude are correlated, although it is impossible to tell whether the attitudes cause the achievement or vice versa. This is corroborated by the fact that in some studies high attitudes are a predictor of course success whereas in others they are not.
The progress of research in science education is critically related to the ability to measure relevant variables reliably. As the focus of research shifts with time, it is to be expected that new instruments will be developed to meet changing requirements. This section on instruments describes most of the instrument developments efforts in this summary. Two groups of instruments are not included because they are described in their respective topic section:

1. Instruments involving Piagetian tasks which are covered in the appropriate section under Learning and Development.

2. Instruments involving teacher needs, characteristics, and behaviors that are reviewed in the section under Teacher Training and Behavior.

Excluding the above two categories, in the year under review, the largest categories were instruments related to the broad topic of understanding the nature of science and instruments related to attitudes.

The Nature of Science

Rubba and Andersen (1978) reported a well-conceived study to develop an instrument for assessing students' understanding of the nature of scientific knowledge. The development and field testing of the instrument took place in seven steps. First, a model for the nature of scientific knowledge was established, based on Showalter's definition of scientific literacy. This resulted in six categories: amoral, creative, developmental, parsimonious, testable, and unified. Next, an item pool was constructed using the definition above. It was modified to suit the reading level of junior high school students and refined as to form and content. The items were tested on a group of high-ability junior high school students who helped identify problems and were then submitted to a panel of experts who were asked to match items to the model. The final pool of items judged content valid by the panel were field tested on 676 high school science students. This led to further elimination of items on the basis of item-total correlations. The final 48-item Likert scale instrument comprised of eight items (four positive and four negative) per scale was tested for reliability and construct validity. Alpha reliabilities for various groups ranged between .65 and .88 and test-retest reliabilities were .59 to .87.

Fraser (1978a) reported the development and validation of a new instrument for measuring understanding of the nature of science among elementary or junior high students based on the Test of Understanding Science (TOUS). Three subscales of the test were developed: philosophical, historical, social and normality of scientists. The items were submitted to a panel of experts for judgment of content validity. The new measure was administered to a sample of seventh-grade students. Item-remainder correlations were used to eliminate unworthy items as well as to eliminate items which correlated more highly with subscales other than
their own. The revised scale of 30 items was cross-validated on a second sample of seventh graders. KR-20 reliabilities for the subscales were between .51 and .62, and was .77 for the whole test.

Lamb (1977) described the validation of the Views of Science (VOS) questionnaire developed by Hillis to measure elementary education students' views of the tentativeness of scientific knowledge. The test was administered to a number of groups including university biology faculty members, master's level biology students, teachers, undergraduate science and nonscience majors, and high school students. The faculty members scored higher than all other groups and the graduate science students higher than most other groups. The results thus lend some support to the discriminant construct validity of the test.

McDermott (1978) reported on the development of a test to measure attitudes needed for scientific literacy. The item pool was submitted to a panel of experts for validation and piloted on a sample of 126 students. The 60-item final version of the test was administered to 1200 students. Split-half reliability was found to be .57, and correlation with a science achievement was .38.

Finally, Ogunniyi (1978) developed an instrument to measure teachers' concepts of science. The nature of this instrument is described in the section on Teacher Characteristics and Behavior.

Attitudes

This section reports the development of instruments to assess affective constructs other than those involving the environment which are reviewed above in the section on Environmental Studies. The construction of attitude scales is fraught with pitfalls, the avoidance of which requires careful construction and testing.

Fraser (1977, 1978b, 1978d) described the two stages of the construction of a composite instrument the Test of Science-Related Attitudes (TOSRA). The instrument contains seven Likert scales: (1) social implications of science, (2) normality of scientists, (3) attitude to scientific inquiry, (4) adoption of scientific attitudes, (5) enjoyment of science lessons, (6) leisure interest in science, and (7) career interest in science. In the first stage of development (Fraser 1977, 1978d) five of the scales (1,3,4,5, and 6) were constructed from previous scales used in England and Australia, and were tested on 1158 Australian seventh graders. In the later study (Fraser, 1978b), the five scales were improved and two new ones were added. Panels of experts were employed to ensure face validity of the items. The final instrument containing 16 items per subscale was tested on 1337 students of all socioeconomic levels in grades 7-10 in Sydney, Australia. Item total correlations were used to eliminate discrepant items and reduce the number of items in each scale to ten. Reliability coefficients were calculated for each subscale and the whole test in each grade level. Alpha coefficients for subscales ranged from .64 to .93 and for the whole test from .64 to .80. Test-retest reliabilities for a subsample of students ranged from .69 to .85 for the various subscales. Intercorrelations among TOSRA scales were
calculated as an index of discriminant validity. These were low except for the intercorrelations between the three interest scales. This instrument could be used advantageously in science curriculum evaluation.

It should be noted that the earlier study (Fraser, 1977) produced some interesting differential responses to subscales between various groups of students. Boys scored higher than did girls on enjoyment of science lessons, leisure interest in science, and adoption of scientific attitudes; high mental ability students scored higher than did low mental ability students on attitudes to scientific inquiry and adoption of scientific attitudes; and higher SES students scored significantly higher than did lower SES students on social implications of science and adoption of scientific attitudes.

Markle (1978) reported on the development of a Subject Preference Inventory (SPI). The instrument, tested with a sample of 28 preservice elementary teachers, required students to choose between pairs of subjects they preferred to teach. All possible pairs of eight subjects were presented for comparison. Average internal consistency over two testings and test-retest reliability were high (.92 and .86, respectively). Predictive validity was estimated by correlating science preference with the number of science-related activities in which the student voluntarily engaged. This was found to be .62. The test takes ten minutes to administer and could easily be adapted to other attitude preference constructs.

Bratt and DeVito (1978) described the construction of an inventory of attitudes toward science teaching and humanistic science teaching. Six scales, each with equal numbers of positively and negatively-phrased Likert items, were constructed. The first three scales related to attitudes towards teaching science and were adapted from a previous instrument, Moore's Science Teaching Attitude Scale. The items for the last three scales were devised by the authors and were examined for content validity by three experts. The instrument underwent three stages of refinement using samples of elementary education majors. Factor analyses (on the negative and positive scales separately) were used to eliminate items which did not match with their proposed subscale. Item loadings on factors were high for three of the factors (those identified with the science teaching attitude scales) but less determinate for the other three. Test-retest reliability for the whole instrument was .87 and for the subscales above .80 in most cases. One scale had an inexplicable near-zero reliability.

As part of a broader study reported elsewhere in this review, Krajkovich (1978) reports the development of an instrument to ascertain the Image of Science and Scientists Scale (ISSS). Alpha reliability and construct validity were assessed but the coefficients were not reported.

Two attitudinal instruments developed in conjunction with teacher training studies but having possible general application are described under Teacher Training and Behavior. They are an environmental attitudes questionnaire by Hart (1978) and an attitudes toward inquiry inventory by Lazarowitz et al. (1978).
**Processes of Science**

A number of new tests to measure science process skills among elementary students have been developed according to the processes delineated by the AAAS SAPA program. These have been used to evaluate both SAPA and other elementary science curricula. One test containing four process tasks was discussed in reporting a SCIS outcome study (Bowyer and Linn, 1978). In addition, a process instrument development study was described by Perez (1978). Perez outlined the development of a process skills test for sixth-grade students in the Philippines. The Perez Test of Science Processes (PTSP) is a 60-minute pencil and paper, multiple-choice test that tests eight processes. The processes and items were matched to the elementary science curriculum used in the Philippines. The PTSP was submitted to a panel of five Filipino science educators for content validation and then administered to 1705 students. The KR-20 reliability was found to be .87 and the mean item difficulty was .39. Construct validation was not attempted. The test will be used in curriculum evaluation in the Philippines.

**The Laboratory**

The laboratory has been a focus for many of the new curricula. Tamir and Lunetta (1978a) described two checklist instruments for assessing the laboratory content of a text. The first is a task analysis instrument for analyzing laboratory experiments. It has 16 items. The second is a four-item list for gauging the degree of organization of laboratory work with other components of a course.

Ben-Zvi et al. (1977) reported on the construction of an instrument to assess the attainment of laboratory objectives in high school chemistry. The instrument listed 23 objectives of laboratory work. It was validated by ten experienced high school teachers whose agreement that the objectives were those of the "new" curricula was used a criterion for inclusion. The questionnaire was circulated among twelfth-grade chemistry students, high school chemistry teachers and academic chemists who were asked to rate the importance of the objectives on a five-point Likert scale. In addition, the students were asked to rate the degree of attainment of each objective. Factor analysis of student responses was used to attain construct validity of the instrument. Student and adult scores were compared. Students tended to rate objectives related to manipulative skills and scientific method to be more important whereas teachers and academic chemists preferred objectives in the affective domain.

Lee (1978) described the construction of two other instruments related to the objectives of laboratory work in biology. The Laboratory Function Form listed 35 functions of the laboratory in a Likert format. Respondents rated the functions on a scale of importance. The statements were selected on the basis of an extensive literature review and submission of a pool of 120 items to a panel of experts. The Ranking Form consists of five functions, the subscales of the Laboratory Function Form. Here, respondents are required to rank order the five main functions of the laboratory. The forms were administered to faculty, teaching
assistants and freshman biology students. Respondents were found to accept the five major functions of the laboratory, namely: manipulative skills; processes of science; knowledge of subject matter, nature of science; and attitudes, interests and values.

Subject Matter Content

The evaluation of higher level objectives in Bloom's Taxonomy (analysis, synthesis, evaluation) cannot readily be achieved using objective formats, and scorer unreliability for open-ended test is notoriously low. Hambleton and Sheaehan (1977) described the use of a free-sort categorization technique for the evaluation of higher-order science objectives. Concepts in a given unit are sorted, by the student, into categories of like concepts. The number of categories is not specified ahead of time. A Latent Partition Analysis is used to analyze the responses of the whole group and loadings of concepts on categories are produced. The authors illustrated the procedure with 284 ninth-grade science students for concepts in "The Structure of Matter." The technique is useful in gauging changes in student perception prior to and after instructing or in comparing student perceptions of experts. It is thus similar to other concept-mapping approaches. Its drawback is that it does not permit the analysis of the conceptual map of individuals. Only the overall perceptions of a group can be mapped.

Two dissertations reported on the development of tests for specific purposes. In one study (Wolfe, 1978), the goal was to develop a Test of Higher Cognitive Learning in Chemistry for freshman college chemistry. The 40-item multiple-choice test contains four subtests—one for each of the three lowest levels of Bloom's Taxonomy and one for higher level objectives. The test was piloted among college chemistry students. A moderate degree of construct validity was established for the test.

In the second study, Oyebanji (1978) developed a test to evaluate performance in the eighth-grade Nigerian Integrated Science course. After refinement on a pilot group, norms were established using a sample of 1000 Nigerian eighth graders who had completed the course.

Finally, in the section on Teacher Characteristics, Opinions, and Abilities, two tests developed by Rowsey and Henry (1978) to test knowledge of metric transformations and application of metric measurement were described.

Textbook Evaluation

Lowery and Leonard (1978b) described the construction and use of their Textbook Questioning Strategies Assessment Instrument (TQUSA1). The instrument categorizes questions along two dimensions: (1) question types (five types from convergent to divergent) and (2) science/learning processes elicited by the questions. The distribution of question/process categories is calculated from a random sampling of pages from the text. The ranking of the categories in terms of increasing intellectual complexity was performed by a panel of experts. In addition to the categorization of questions, the position of questions relative to the textual information
is noted. The instrument provides a useful tool for comparing test
and analyzing the writing employed.

This section on Instrument Development has described in some
detail the procedures used to validate new instruments. In many cases
an entire dissertation or research project was devoted to this task.
This is certainly an encouraging sign as it is a very consuming process
to obtain instruments with high levels of objectivity, reliability, and
validity. The only discouraging thing about the proliferation of new
instruments is that this tends to make research in science education
more diffuse and research studies less comparable. What is needed are
several highly reliable and valid instruments around which large numbers
of science educators could consolidate their work.

Evaluation Models

The large-scale program development efforts of the 1960s spurred the
development of a myriad of evaluation models for assessing their worth
and as an aid in decision making. Any serious curriculum development
effort should be accompanied by systematic evaluation matched to the
program by a selection of elements from available models or the intro-
duction of new elements. In recent years, evaluation projects have been
concerned with rather more limited goals—the evaluation of single
courses or modules. Most of these have employed a Tylerian approach of
testing success by measuring performance against behavioral objectives.
It is not surprising that discourse on evaluation models has been cur-
tailed and that few studies in science education have sought to investi-
gate models of evaluation or to employ nonstandard ones. Some exceptions
to this will be reviewed in this section.

The Provus' Discrepancy model was used by Trotter (1978) to evaluate
two graduate science education programs at Temple University. This model
is cybernetic in nature, focusing on discrepancies between aims and
performance, and on recommendations to correct such discrepancies. In
the study cited, an opinionnaire was administered to faculty, graduate
students and post-graduates of the science education department to elicit
the desired aims and perceived performance of the graduate science educa-
tion programs. All groups used common criteria in their judgment of aims
and performance. The discrepancies between aims and performance were used
as a basis for suggesting improvements in the program. The Provus model
was adjudged adaptable to the evaluation of graduate programs.

Angotti et al. (1978) reported on a study of a new module for teach-
ing special relativity with prime emphasis on attitudinal objectives
which employed an "illuminative" model. This model is anthropological
in nature, relying on participant observation by the teacher for evidence.
The model proved helpful in elucidating some of the subtle conceptual and
philosophical difficulties experienced by the students in studying rela-
tivity.

Ogborn et al. (1978) described a novel model for conducting an eval-
uation. The transcripts of interviews with 155 physics students from ten
different British universities were analyzed. Each student was encouraged
to tell about one good learning experience and one bad learning experience.
Characteristics of good and bad stories were discussed. The interview
model proved to be an effective tool for analyzing teaching programs.
CONCEPTS AND CONTENT

The focus of studies reviewed in this section is the learning of particular concepts in the various science disciplines. Emphasis is on the nature of the concept itself rather than on teaching methods or strategies that can be used for concept attainment. Examined are such things as the prerequisite skills needed for understanding the concept, subconcepts, and misconceptions of the concept.

Experimental methods used in these studies frequently vary from more standard empirical techniques. In many cases students are interviewed to probe misconceptions and problems in learning specific concepts.

The outline followed for this section was to begin with descriptions of studies that have applicability to concept learning in general. This was followed by reviews of studies related to specific concepts in physics, chemistry and biology. For a review of research on scientific concepts, the reader is referred to Preece (1978a).

General Concepts

In order for science educators to learn what students know about particular concepts, a method of evaluation is necessary. One such method is the use of concept mapping. Concept maps are two-dimensional representations of concepts, examples, and concept interrelations within a given science area. Two studies involving concept mapping were reported during this time period.

Rowell (1978) used concept mapping to study its usefulness in evaluating AT instruction, the extent to which a child's scientific cognitive structure was uniform over several science concepts, and how science concepts developed over time. From examining data from interviews of 114 elementary children on four separate occasions over a two and one-half year period, he concluded that the AT program was successful, that science concept learning is, to a great degree, content specific, and that major concepts, rather than being learned in toto, develop progressively over time.

Hambleton and Sheehan (1977) described two methods that could be used in formulating concept maps. These are free-sort categorization and latent partition analysis (LPA). A description of the LPA technique as applied to an instructional unit entitled "The Structure of Matter" is described in their report. By comparing latent categories with groups before and after instruction, between high and low ability groups, or between students receiving different methods of instruction, it may be possible to evaluate the attainment of higher order science objectives. A further use might be in the construction of a concept map that would aid the teacher in determining a suitable order for the presentation of concepts.

Hunziker (1978) examined a strategy for the study of concept formation in the physical sciences through interviewing 18 undergraduate...
students in a general physical science course. Specific recommendations were not included in his abstract.

Treagust and Lunetta (1978) tested extensions of a mathematical model of concept learning proposed by Bower and Trabasso to a task in science education. The model was extended from a single binary classification system to a four-category system involving two binary dimensions used in the classification of leaves. Because some of what is taught in the science classroom is concept formation involving concepts with binary dimensions, the results have wider applicability than leaf classification. The investigators found that certain extensions of the concept learning models may apply to the complex identification task of identifying leaves "however, only stationarity of pre-revolution responses was fully consistent with the Bower and Trabasso model" (p. 352).

Goldsmith (1978) studied changes in the importance of certain popular science misconceptions from 1940 to 1973. In 1940, the 100 most prevalent science misconceptions were identified and ranked by science educators according to their potential for influencing the behavior of persons who held the misconceptions and acted accordingly. In 1973, this same set was ranked according to the same norms by other science educators. One general conclusion from the comparison of the rankings is that all of the misconceptions have become less important in the last three decades.

A misconception that many science students apparently have is in the area of causation. Bartov (1978) tested high school students in Israel to determine whether they distinguished between teleological and causal relationships. He then included instruction in the biology curriculum (five lessons) that would help students distinguish causal, teleological, and anthropomorphic explanations. This special treatment was effective in helping students make the ability persisted on a retention test given three months later.

Another area of misconception of terms commonly used in science and especially in Piagetian interviews is in the understanding of the terms "all" and "some" by elementary school children. Gates and Jay (1978) replicated an Inhelder-Piaget color and shape experiment in which they paid close attention to the relative difficulty of the various questions, looked at age-level trends using continuous substances, and examined more closely how children were misled by the fullness or non-fullness of containers. They concluded that their experiments supported Inhelder and Piaget's claim that six to nine is a transitional stage in children's learning of the concepts "all" and "some," but that even some 11-year-old children fail to use the terms correctly. The use of discrete substances and the use of counting techniques by children led to fewer correct responses. Children frequently confused the fullness of the container with the meaning of "all."
Specific Concepts

Biology

Deadman and Kelly (1978) reported on what secondary school boys know about evolution and heredity before they are introduced to the topic. Fifty-two boys from four English schools were interviewed. Seven broad categories of response were detected that collectively portrayed the pupils' conception of evolution and heredity. The authors suggested that these foci could be used on the basis of concept maps for the teaching of the concepts. They also reported that the concept of adaptation was particularly well established whereas that of chance was the least established. Naturalistic and Lamarckian interpretations of evolution were predominant.

Bliss (1978) compared the concept development of biology students studying the origin of life from a two-model approach (evolution and creation) with those using a one-model approach (evolution only). Analysis of the data obtained using a pretest-posttest, control group design showed that the students who received the two-model approach achieved significantly higher on an achievement test. In addition, the overall data suggested that the two-model group was more willing to make critical judgments in regard to origins, and to change their views when new data comes forth, than were those studying the evolution-only approach.

Ardoin (1978) studied whether high school students taught by the systematic approach would score significantly higher on a standardized measure of biology achievement than students taught by a traditional approach (taxonomic). No significant differences in achievement were found.

Hawk (1978) reported on the development of a short life cycle (45 days) higher plant, *Brassica campestris* L., to use as an alternative for *Drosophila melanogaster* for use in the laboratory instruction in genetics.

Chemistry

Cain (1978) developed a treatment of molecular orbital bonding theory that was suitable for high school students. Students at two of the four schools using the materials achieved at a 70 percent mastery level.

Novick and Nussbaum (1978) studied Israeli junior high school pupils' understanding of the particulate nature of matter. One hundred fifty-four students of high, average, and low ability gave explanations during interviews of eight physical science phenomena involving gases using either continuous or discontinuous descriptions. All students were in the eighth grade and had studied atoms and molecules as composing matter during the seventh grade. Results indicated that only 60 percent of the sample consistently used a particle model to explain the phenomena. Many pupils could not perceive "empty space" in ordinary matter, including gases.
Cody and Treagust (1977) studied the effectiveness of learning inorganic chemistry through the study of biochemistry. Twenty-six students were pre- and posttested on the American Chemical Society's Inorganic-Organic-Biological-Chemistry Test. In the intervening six-week period students studied biochemistry for approximately six hours a day. This resulted in significant differences in achievement on all three sections of the test.

**Physics Concepts**

Albert (1978) investigated children's concepts of heat by interviewing 40 children aged four to nine over an eight-month period. Eleven thought patterns that could be classified into six categories were identified. These were described chronologically thus providing a developmental approach to the question of how children develop the concept progressively.

Wong (1978) explored the meaning of energy concepts of an eleven-year-old girl engaged in a guided investigation of energy. Investigations centered on energy transfer, derivation of a definition of energy, energy requirements for growth, energy capture in photosynthesis, and energy release in respiration. Analysis of interviews showed that the subject's understanding of concepts related to energy were different from those held by the scientific community.

Bilbo and Milkent (1978) compared two different approaches to teaching the concept of volume to college students enrolled in a physical science course for nonscience majors. One approach presented length, area, and volume in the order given, and computation of volume was stressed. The other approach did not include teaching area, hence giving more time to developing an understanding of volume through direct measuring experience. Results indicated that the latter approach was more effective in teaching the understanding of volume and in helping students make estimations and measurements involving volume.

Shawhan (1978) studied the effectiveness of teaching Gauss' Law and Ampere's Law through the use of algorithms by means of computer-assisted instruction. Evaluation of the data indicated that students using algorithms achieved significantly higher than those in the control group. Results also indicated that students who used the algorithms in learning one of the above laws transferred the technique to learning the other.

Because of the wide diversity in subject matter, the results of studies in this section cannot be synthesized. An observation that may have some significance is that a large number of studies explored students' understanding of particular concepts. Researchers are becoming increasingly aware of the lack of evidence that particular teaching strategies are effective for teaching all the concepts in a given discipline and that further research may even be futile. Therefore, more research on specific concepts may be forthcoming. This is an area that appears to hold much promise for research by science educators.
MISCELLANEOUS STUDIES

The following four studies did not logically fit into any other section and, as a consequence, are included here. The first pertains to career awareness, the other three to research outside of the United States.

Careers

Korotkin et al. (1978) determined that the Science Career Awareness Training (SCAT) program was an effective means of increasing knowledge about science careers and interest in science as a career for students in grades four to six. Participants, however, became more indecisive as far as career choice was concerned. The SCAT program uses a highly-interactive, computer-based system to provide information about science career areas. The information is presented in the form of 30-minute dialogues that include both printed information and photographs as well as problem-solving experiences.

Research Outside the U.S.

Lybeck (1976) reported on science education research undertaken in Sweden in 1976. A summary is given of work done by government commissions and projects of national boards, a brief review of projects reported in Swedish journals is made, and a report on recent studies is included.

Frey et al. (1977) reported on research in science education in Europe from two vantage points: national reports, and perspectives and structures. Area reports include those from Austria, Belgium, Cyprus, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, and the United Kingdom.

Lorenz (1978) reviewed 143 Brazilian science education research studies. Findings included that most research was done on the university level with physics, in particular, and that few comparative studies were reported. Most of the studies were course descriptions and reports on teaching innovations.
SUMMARY, CONCLUSION, AND IMPLICATIONS

In this Summary of Research in Science Education—1978, we have included summaries at the end of sections and subsections where enough research was undertaken during the year to permit generalization. In this section, we attempt to relate this year’s research to that of the past, to identify trends, and to give implications for further research.

Learning and Development

During 1978, as in the past five years, studies based on Piagetian developmental theory dominated the literature. The focus of the studies, however, appears to be changing. As late as 1977, the major emphasis was on relating students' cognitive developmental level to other demographic data (Peterson and Carlson, 1979). Many of the earlier studies were training studies in which investigators attempted to change students' developmental level (DeCarcer, 1978). In the 1977 review, Peterson and Carlson advocated the areas that needed development in Piagetian research: (1) enlarging the theoretical horizons, (2) applying Piagetian principles to the teaching of science, and (3) developing a phylogenetic model. In 1978 there was a move toward implementing the first two of these recommendations. By focusing on the proportionality schema, the finer structure of this aspect of Piagetian theory may soon be elucidated.

Studies relating theory to teaching included those that examined the effectiveness of concrete experiences on achievement and those that determined the usefulness of the learning cycle of exploration, invention, and discovery.

The fact that Piagetian research during 1978 was more refined than in the past, that it focused in more detail on particular schema, that it examined practical implications for teaching, that instruments were more carefully validated, and that it attempted to link learning and physiological development were encouraging signs to us. If it continues along this vein in the next few years, we should see the full implications of the theory.

Unfortunately, so few studies were reported in the literature using a different theoretical base, that it is impossible to compare them. Perhaps a comparison is not even in order, as different theories pertain to different aspects of learning. Herron (1978) has succinctly described this in pointing out the differences in the function of the Ausubelian and Piagetian theories. He stated:

Ausubel's notions help us understand conditions that affect the acquisition of new information and incorporate it in our store of previously learned information so that interrelationships among concepts are clear and so that information can be recalled and applied to novel situations. Piaget's theory describes the

---

generalized logical operations (seriation, correspondence, reversibility, combinatorial reasoning, proportional reasoning, etc.) that are likely to be available to individuals at various ages and that are necessary to make sense out of information that is deliberately taught or spontaneously learned (p. 593).

During 1978, studies based on Ausubelian theory show general support of the theory. There was, as in other years, mixed results on the effectiveness of advance organizers. As Renner et al.9 concluded in the 1976 review, a more systematic approach in this area is needed.

In conclusion, we are encouraged by the good news that studies with a Piagetian base are more refined and that their number has increased. Simultaneously, we are discouraged by the bad news that many studies lack any theoretical base, and that there are too few studies exploring theories other than that of Piaget. The latter is a reiteration of writers of these reviews for the past five years.

**Teaching-Learning Strategies**

Studies on teaching and learning strategies continue to constitute the largest area of research included in the summary. During the past five years there has been a shift in the particular strategies that were investigated. The number of studies concerned with aspects of individualizing instruction has declined. In 1978, only 26 studies were reported as compared to 49 in 1975. Investigations on the success of specific methods that appeared to hold much promise for individualizing instruction; such as CAI, have not occurred. Only two or three studies are reported each year. Perhaps with the availability of low cost microcomputers, the research in this area will burgeon. Another area of waning interest is that of behavioral objectives. In 1973 this constituted an entire section in the review and dominated the teaching-learning strategy section.

Areas in which there have been a rather even but small number of studies reported each year are inquiry, inductive and deductive approaches, the laboratory, feedback, questioning, and teaching methods for the handicapped. The latter is an area in which we might expect to see an increase in the number of studies due to mainstreaming.

The largest increase in the number of studies in a given area appear to be those related to reading and the language arts. It may be that because these studies in this review have been grouped together, that this is illusory. Whether it is or not, the fact remains that this is a most important area that deserves attention by science educators. As students' SAT verbal scores continue to fall, and as students are expected to read more than ever in many individualized modes of instruction, it is imperative that science educators explore ways in which science facts, concepts, and principles become more available to students through the written word.

---

Results of the research on teaching-learning strategies have been summarized throughout that section of the review and the reader will find details there. In general, results of the research during 1978 confirm results of those of past reviewers. For almost every teaching strategy, reports of both success and failure (no differences) abound. Frequently there are positive affective rather than cognitive outcome whether these are stable over time goes untested. Strategies that appear to be effective in producing certain types of cognitive gains are kinetic structure, nonlaboratory approaches and deductive lessons. In some cases, these cognitive gains may be at the expense of other cognitive and effective outcomes.

Cognitive Style and Student Characteristics

A separate listing for the area of cognitive style has not been used in previous reviews. However, authors as early as Rowe and DeTure (1974) have advocated the use of aptitude by treatment interaction (ATI) studies as a possible field for future research. While few of the studies included in this section were actual ATI experiments, some of the cognitive styles studied may offer researchers possible appropriate aptitude measures. The number of studies has greatly increased from the 1977 review of Petersen and Carlson with less than a dozen studies covering this area versus the 28 of the current year.

The constructs that appear to have become widely used in the recent literature are field independence-dependence, cognitive preference, and locus of control. While the results have been somewhat mixed, especially for the field independence-dependence research, the area still holds possibilities for future work.

Researchers should carefully examine the literature on the various constructs to determine their suitability for study in their research situation. Even broad constructs, such as field independence-dependence, may not be important aptitudes for all situations.

Of the 26 studies on student characteristics, 21 had as one of the outcome variables a measure of achievement, while only three were interested in attitudes as outcomes. The other two studies were looking at variables such as "workstyle" and "critical thinking ability" as the dependent measures. This indicates some trend away from studying attitudes as outcome measures, although several studies used them as predictors. Their use as predictors, however, was not especially strong with several instances of attitude measures accounting for less than 10 percent of the variation in the outcome measures.


This trend of attitude measures as being rather mixed as predictors has been mentioned by previous reviewers. As Peterson and Carlson (1979) noted, "...a much stronger argument can now be made for saying that achievement creates positive attitudes and probably not the reverse, ..." (p. 71).

Future refinement of methods to measure outcomes, both attitudinal and achievement, may be an important step in future work. Instruments measuring very broad outcomes may not have a tendency to detect differences between the groups under study.

Teacher Training and Behavior

Teacher education continues to be an area of active research. One of the major reasons for this is probably because science educators find it both convenient and profitable to evaluate courses and workshops in their own geographic region. While this practice certainly has its merits in improving local instruction, when research is confined to one instructor and to materials the instructor has produced for use in his/her own classroom, generalizability of the research is questionable. Then again, if enough of these studies are carefully conducted, reported, and coordinated, results may be synthesized to yield something of significance.

During the year under review, several studies have shown that there are effective teaching practices that can be taught to teachers. Some of these are questioning skills (including wait-time), inquiry skills, and kinetic structure. A few investigators even attempted to determine if these practices had an effect on student achievement, thereby crossing the gap between skill acquisition and skill effectiveness.

Although effective training procedures can produce teachers with certain skills and attitudes, the duration of their acquisition is rarely tested. What is needed is a systematic attempt to coordinate preservice and inservice programs so that skill acquisition and behavioral characteristics of teachers can be monitored over a long time period. This would allow educators to determine what supportive measures teachers need to maintain and improve their teaching skills.

Evaluation in Science Education

Studies that were reviewed in the Evaluation section were very diffuse in nature and are, therefore, difficult to summarize. Particular mention must be made of the three large evaluation surveys that were reported during the period. The National Assessment of Educational Progress, the third such survey of 9-, 13-, and 17-year-old students in the last ten years contains a wealth of data that can be used to compare outcomes of science instruction under a variety of circumstances over a long period of time. The report included in the review is preliminary to this. It discusses only the results of the 1976-1977 survey. Research reports that compare outcomes of the three surveys should be forthcoming.
The two NSF reports also contain large quantities of data that should see future use. Because of the comprehensiveness of these reports it was not possible to adequately summarize them in such a way as to be useful to the researcher. Science educators are urged to study both reports carefully. The information contained therein should not only aid science educators in planning teacher education programs but may serve as a data base for future research.

Also worth noting is the expansion of the section on affective studies to include those concerned with attitudes toward the environment, and science, society, and morals. Interest in both of these areas appears to be increasing, and as suitable instruments are developed, studies are bound to multiply. These areas may hold high priority in future science education research.
REFERENCES

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

Numbers in boldface following each citation refer to the page number(s) in the Summary where the article is discussed.


Dallam, Jerald W. "Student Behaviors Related to Academic Success of Undergraduates at Kansas State University in the Natural Sciences." Mimeographed paper, 1978. ED 166 040


Dunlop, David L. "The Effects of an Energy-Environment Simulator upon Selected Energy-Related Attitudes of Science Students and In-Service Teachers." Paper presented at the annual meeting of the National Association for Research in Science Teaching, Toronto, Canada, March 31-April 2, 1978. ED 164 267


Fletcher, Richard K. "A Summary of Research Literature Predicting Success in College Chemistry." Paper presented at the annual meeting of the Tennessee Academy of Science, Tennessee Technological University, Cookeville, Tennessee, November, 1977. ED 155 054


Grabe, Mark and others. "Evaluation of the Multidimensional Affective Consequences of Traditional and Mastery Instruction." 1978. ED 152 521


Lorenz, Karl M. "A Review of Science Education Research in Brazil." 1978. ED 159 059


ED 164 268


ED 162 856


ED 155 051