DOCUMENT RESUME

ED 209 075 SE 035 716

AUTHOR	Robinson, James T., Ed.
TITLE	Research in Science Education: New Questions, New Directions.
INSTITUTION	ERIC Clearinghouse for Science, Mathematics, and Environmental Education, Columbus, Ohio.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.; National Science Foundation, Washington, D.C.
PUB DATE	Jun 81
CONTRACT	400-78-0004
GRANT	NSF-SED-80-17531
NOTE	198p.
AVAILABLE FROM	Information Reference Center (ERIC/IRC), The Ohio State Univ., 1200 Chambers Rd., 3rd Floor, Columbus, OH 43212 ($5.75).
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	*Classroom Research; College Science; Elementary School Science; Elementary Secondary Education; Higher Education; *Learning; Mathematics Education; Problem Solving; *Science Curriculum; *Science Education; *Science Instruction; Science Materials; Secondary School Science; Textbook Preparation; *Textbook Research
IDENTIFIERS	Mathematics Education Research; *Science Education Research

ABSTRACT

Presented is a compilation of papers focusing on new questions and new directions for research in science education. Papers are grouped into one of three related fields of research: (1) analysis of curriculum materials, especially science textbooks; (2) investigations of science understandings of students and experts; and (3) investigations of the contextual factors of science classrooms. A general introduction precedes each of these three sections and summarizes and compares papers. Individual papers describe a particular problem related to one of the general areas, the theoretical base underlying the research and how the problem is being investigated, and selected research findings. Topics include: (1) textbook assessment; (2) analysis of science textbooks; (3) structure, strategies, and comprehension in learning; (4) concept of change in scientific reasoning; (5) understanding and problem-solving in physics; (6) implications of classroom research for science and math instruction; (7) mathematics classroom inquiry; and (8) the case for the participant/observer in mathematics classrooms. (DS)

RESEARCH IN SCIENCE EDUCATION: NEW QUESTIONS, NEW DIRECTIONS

James T. Robinson, Editor

June 1981

CERE
A Continuing Program of the BSCS
CENTER FOR EDUCATIONAL RESEARCH AND EVALUATION
833 W. SOUTH BOULDER ROAD LOUISVILLE CO 80027 PHONE (303) 666-6558

ERIC®
CLEARINGHOUSE FOR SCIENCE
MATHEMATICS AND ENVIRONMENTAL EDUCATION

Supported by a Grant from the National Science Foundation

Published by the ERIC Clearinghouse for Science, Mathematics, and Environmental Education and by the National Science Foundation

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

Support for the conference and the preparation of this report was provided by the Research in Science Education (RISE) program, National Science Foundation, under Grant No SED 80-17531. The opinions expressed here do not necessarily reflect the positions or policies of the National Science Foundation

TABLE OF CONTENTS

INTRODUCTION TO THE CONFERENCE

Research in science education, and in education in general, has reached a stage of criticism, questioning, and ferment. New questions are being asked, and new research methodologies are being generated. At the same time, there is increased recognition of the "low estate into which science education in this country has fallen" (Carey, 1981). The report to the President of the United States prepared by the National Science Foundation and the Department of Education (October 1980) suggests that, over the past 15 years or so, our national commitment to excellence has been shrinking, as has our international primacy in science, mathematics, and technology. The time seems particularly right to look for new questions and new directions for research in science education.

Such was the intent of the conference "Research in Science Education: New Questions, New Directions," held November 6 and 7, 1980, in Boulder, Colorado. Such is the intent of this publication, which is a result of that conference. The conference was organized by Dr. James T. Robinson, Director of the Center for Educational Research and Evaluation (CERE) of the BSCS and was supported by the National Science Foundation and involved over 80 science educators with nine speakers.

The conference planners kept several basic questions in mind:

- What areas of research hold exciting promise for the improvement of science instruction?
- What areas of research attack old intractable problems in a new and different way?
- What areas of research are conceptually innovative, that is, have not been extensively applied to the study of teaching and learning mathematics and science?
- What areas of research stay close to what actually occurs in science and mathematics classrooms?

In answering these questions, three related fields of research were selected:

- Analysis of curriculum materials, especially textbooks.
- Investigations of the science understandings of students and experts.
- Investigations of the contextual factors of science classrooms.

For each of these areas, three researchers were invited to present papers about their investigations. Each was asked to describe a problem he or she is investigating, to discuss the theoretical base underlying the work and how the problem is being investigated, and to give examples of the research findings. The three topics were chosen because these are emerging fields of study that are not being used in science education research or are used in only limited ways. They are avenues to new ways of investigating science instruction that may yield useful knowledge that will influence practice.

The nine researchers spoke to an invited group of participants that included college and university professors and state and school district science supervisors and administrators--a mix of both producers and users of research. The program consisted of four sessions. Three of these were thematic symposia, each of which included presentations by three researchers. During the fourth session, participants met in small groups with individual presenters. Time permitted two such meetings.

Readers of these conference proceedings are invited to share in the excitement and enthusiasm of the participants in the conference. The papers presented at the three symposia, along with several of the papers cited in the references to the last chapter, are suggested as reading for faculty and student seminars about research. Some of the implications of the current ferment in research methodology for science education are suggested in the concluding chapter. These too could be discussed with colleagues focusing on "new questions, new directions" for improving research in the practice of science education.

References

Carey, W. D. Science Education: Rhetoric and reality. Science, 1981, 211, 339.

National Science Foundation and Department of Education. Science and engineering education for the 1980s and beyond. Washington, D. C.: U. S. Government Printing Office, October, 1980.

ACKNOWLEDGEMENTS

The editor is especially grateful to Robert Howe and his colleagues at the ERIC/SMEAC Clearinghouse for Science, Mathematics, and Environmental Education for their encouragement and early agreement to publish this conference report and to the National Science Foundation for their support of the conference and desire to publish the conference report also.

Those who presented papers and whose studies are reported here made the conference a stimulating experience for all. The speakers were involved in the full deliberations of the conference and their stimuli to the participants made the conference a dynamic and challenging event. The participants contributed greatly to the conference with probing questions and salient comments.

The conference could not have happened without the cooperation and assistance of the BSCS and CERE staffs. Special thanks go to Marcia Racheli for handling the conference details and to Barbara Helgerson for handling arrangements. To Mary McConnell, Richard Tolman, and Jo Williams, CERE Research Associates, many thanks for your support and assistance both during the conference and with this conference report.

The final production of the report could not have been completed without the able assistance of Ann Whitcomb, Sue Ann Todhunter, and Marcia Racheli, secretaries; and Eunice Combs, BSCS Art Director.

James T. Robinson
Director
Center for Educational Research
and Evaluation

7

Chapter 1

INTRODUCTION TO THE SYMPOSIUM

Printed materials have long been a central part of teaching and learning in American schools. Yet these materials have been subjected to little rigorous study in their long history of use. Indeed, textbooks suffer even more than other books in the quality of critical reviews they engender in the professional literature. However, in the past decade, important theoretical work in text comprehension and knowledge integration has been developed.

Readability studies have been the method most used for analyzing text and other print materials. Although the severe limitations of readability scores for text passages have been recognized by both researchers and practitioners for many years, this recognition has not, until recently, stimulated the development of alternatives. The recognition by researchers that readibility scores are not properties of print materials but must always be considered interactively with the reader has yet to be understood by many practitioners and by the textbook industry.

Michael Macdonald-Ross (1979), in a major review paper, "Language in Texts: The Design of Curricular Materials," advocated the use of text analysis as a critical part of the formative evaluation of newly developed materials. He acknowledged that the complexity of prose and the lack of good analytic procedures for decomposing text structure have seriously hampered progress in the field of text analysis. There are, however, several lines of research with promising methods of analysis that have been used with a variety of text materials. These lines of research include readability studies, quantitative context analysis, linguistic analysis, and subject matter analysis. In concluding his review of language in texts, Macdonald-Poss indicated that, although research in the past ten years has provided new insights into the structure of texts and pedagogical materials, it has not had much impact on curriculum design. He explained that some of this type of research has reinforced standard procedures of curriculum design, but he also suggests that many research questions of the past have not been fruitful and that new directions should be tried. "There is a notable lack of how-to and when-to/when-not-to information; on the other hand, too much research is done without support of adequate theory."

Three investigators presented papers about text analysis at the conference. Each used a different approach in his or her work. Each was at a different stage of development in the use of the approaches.

Decker Walker was just getting his current research project under way. His argument centered around the usefulness of studying texts in the classroom. He stated that an analysis of print materials must include information about how both teachers and students use texts. Therefore, field experiments may be one of the most useful ways to analyze print materials, if one is concerned with improving such materials in the context of use.

Decker Walker is an Associate Professor in the School of Education at Stanford University. He began his educational career as a science teacher after completing a B. S. degree in physics. Following this, he received an M. A. in the natural sciences and completed a doctoral degree in general curriculum at Stanford University. His major interests have been in curriculum policy and curriculum research, and he conducted major studies of the science and mathematics curriculum development work of the 1960s and 1970s. Professor Walker is coauthor of Case Studies in Curriculum Changes: Great Britain and the U. S. and author of numerous articles and papers in professional journals.

In Chapter 2 of this volume, "Learning Science from Textbooks: Toward a Balanced Assessment of Textbooks in Science," Walker presents some of the complexities he feels must be considered in designing effective field experiments about textbooks and their uses. These studies focus on three dimensions: the textbook itself, how teachers use it, and how students use it.

Authors of texts and developers and writers of curriculum materials can improve student comprehension of their products if they will attend to what is now well substantiated by research. This is the claim of Armbruster and Anderson in Chapter 3, "Analysis of Science Textbooks: Implications for Authors." Their paper considers four characteristics of prose that are commonly mentioned in rhetoric and composition textbooks: structure, coherence, unity, and audience appropriateness. They cite research studies relevant to these characteristics and recommend that prospective writers use the guidelines they derive from these studies to improve student comprehension of text material.

Their paper was presented at the conference by Bonnie Armbruster, who is Visiting Assistant Professor at the Center for the Study of Reading, University of Illinois at Urbana-Champaign. She is currently conducting research on training strategies for the comprehension and recall of expository text. Following undergraduate and interdisciplinary studies in the social sciences, she completed her M. A. and Ph. D. in Educational Research at the University of Illinois. Professor Armbruster is coauthor of a chapter on studying in the Handbook of Research in Reading (in press) and a chapter in Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Artificial Intelligence, Linguistics, and Education.

Armbruster and Anderson's basic premise is that certain characteristics of text can optimize learning, if the learner brings sufficient background knowledge to the situation and engages in certain appropriate activities.

Since 1972, James Deese has been the Hugh Scott Hamilton Professor of Psychology and Commonwealth Professor at the University of Virginia in Charlottesville. After earning his Ph. D. in Psychology at Indiana University, he joined the faculty at Johns Hopkins University. The range of his interests is reflected in such papers as "Thought into Speech," "Towards a Psychological Theory of the Meaning of Sentences," and "Mind and Metaphor." He was principal investigator of a National Science Foundation RISE Project, "Learning from Science and Mathematics Textbooks: Text, Structure, Reading Strategies, and Comprehension."

Psycholinguistics, cognitive structure, the development of meaning, and explanations of the interrelationships among student characteristics and the organization of text materials are of special interest in his recent work. Central to Deese's paper is his assumption that many students never learn to cope adequately with expository text. Using biology textbooks at three different levels (seventh grade, tenth grade and college), he reports on the effects of the structure of the texts on comprehension, on how readers perceive the text, on strategies readers use in trying to comprehend the text, and about certain characteristics of the text itself.

The gist of the papers about text analysis is that careful, rigorously designed investigations can produce useful information about science texts and that such investigations should be made. In these investigations, detailed knowledge of the subject matter, as well as detailed knowledge of useful text analysis methods, may eventuate in the improvement of texts.

Text improvement should affect comprehension, of course, but if the text materials do not engage potential readers in the desire to read on, the research will miss the mark. Current science and mathematics texts are not well known for this latter characteristic. The challenge to research workers and practitioners in science and mathematics is, on the one hand, to develop the knowledge base that will enable such texts to be produced, and on the other, to educate practitioners who will be able to distinguish such texts from others and to choose them for their students.

Reference

Macdonald-Ross, M. Language in texts: The design of curricular materials. In L. S. Shulman (ed.), Review of Research in Education, 6, 1978. Itasca, Ill.: F. E. Peacock, 1979.

Chapter 2

LEARNING SCIENCE FROM TEXTBOOKS: TOWARD A BALANCED ASSESSMENT OF TEXTBOOKS IN SCIENCE EDUCATION[1]

Decker F. Walker, Stanford University

This paper presents designs for field experiments on textbooks and textbook use and discusses some of the complexities that must be considered in designing the treatments and criterion measures. Pilot studies of the types described are under way now in science and social studies.

The Textbook: Much Maligned but Little Studied

Book learning has been an object of even more ridicule in the North American tradition than in Europe. The famous vignette John Dewey reports in Schools of Tomorrow (1915) is typical of the bad press textbooks have received. You may remember that Dewey had just witnessed a class in which students were taught about the interior of the earth, specifically, that it was hot, molten rock. As was then the custom, the teacher upon conclusion of the lesson offered the visitor the opportunity to ask the students a question about the lesson. Dewey accepted the invitation and asked the students what they would find if they dug down into the earth deeper and deeper. Would they get colder or hotter as they dug? There was much discussion; the class was divided. The teacher, in some embarrassment, interrupted the controversy to ask, "Class, what is the state of the interior of the earth?" The class is said to have replied in unison, "A state of igneous fusion."

It's a good story, and illustrates well the perils of superficial teaching and learning. But it's not about textbooks! It's about live teaching, about a flesh-and-blood recitation. Any textbook of the time on geology would have done a better job of explaining the current ideas about geology. The evils attributed to textbooks here, as so often

[1]The work reported here was supported in part by a grant from the Spencer Foundation, administered by the School of Education at Stanford University, and in part by a grant from the Ford Foundation (Grant No. 805-0790). I am most grateful for this support. These agencies are not responsible for the contents of this paper.

elsewhere, are not confined to textbooks, nor is there any good reason to believe that they are any more common with textbooks than with live teaching, television, or other modes of education. There is silliness and superficiality all over the place and textbooks are not an exception.

On the whole, there are several good reasons to look with a respectful eye upon the institution of the textbook in spite of its bad press. First, there is the question of its longevity. There is no need to play amateur historian and try to find the first recorded instance of use of a textbook, but the practice is nearly as old as writing, certainly as old as printing, and has been dominant in American schools since colonial days. Second, there is ubiquity. In spite of attacks by nearly every reformer and reform movment, the practice of teaching with and from textbooks continues to be the dominant mode of instruction in schools. The latest data by Stake and associates (1978) merely confirm what has been true ever since people first began looking at American schools: Textbooks are ever-present and appear to be the chief carriers of the curriculum. On the average, teachers deviate very little from the content coverage and emphasis of their textbooks.

Still, lots of bad practices have been around a long time and lots of them are still all over the place. The computer and television are no longer on the horizon but right down the block to challenge the hegemony of the textbook. Reformers are still in opposition, preferring information programs, laboratory-based programs, community-based programs, anything but textbook-based programs.

The most important justification for the use of textbooks in teaching is that students must, by the time they reach tertiary or higher education, be capable of learning from books on their own, with only the distant guidance of a teacher who assigns, lectures, and tests. This will proably be the case for a long time to come. If at some level students are not able to learn from books, from print, from text, then everybody else in the world who wants to communicate with them--scientists, artists, humanists, journalists--will have to learn to make television or computer programs. It is hard enough to teach all those people to write. More than likely, there will be a market well into the next century for people who can learn abstract subject matter from books. The chief function of teaching from textbooks at the secondary level is to prepare students to learn independently from books. Specifically, the chief function of teaching science from textbooks in the secondary school is so students will become able to learn science from books.

Teachers should not feel ashamed to pitch their teaching toward a textbook, even though such teachers are christened "text-dependent" by those who study teaching. It is possible to be dependent on a text, unable to teach without it. But it is also possible to teach responsibly from a textbook, to teach students how to learn independently from texts by careful, gradual fading of teacher assistance in close study of a textbook. This latter is good teaching. No one concludes that a college professor who teaches from a textbook is thereby incompetent in the subject. The textbook is the primary tool of the teacher's trade.

While a good teacher can teach without a text, good teachers also can work better with a good textbook. Much of the criticism textbooks have received is an ill-considered attempt to glorify the role of the teacher by comparison.

Partly as a result of the bad press textbooks have received, they have been little studied. Until a few years ago, the only serious work of research or scholarship devoted to textbooks (other than histories) was Cronbach's Text Materials in Modern Education (1950). More recently, there has been a spate of publications on textbooks, much of it reviewed in Michael Macdonald-Ross's two-part review in the Review of Research in Education (1977-1978). It is time to turn research attention to the textbook. New methods and ideas now are available that will make such study productive.

Field Experiments as the Ideal Design for Studying Textbooks and Their Uses

A field experiment in this context means simply a study in which a conscious effort is made by the experimenter and the teacher to teach something in an explicitly different way and to compare the consequences of teaching it this way with the consequences of teaching it some other way. Except for these consciously chosen changes in the teaching, the situation is as it would be in that setting if the experiment were not being conducted. Teachers and students are assigned to classes in the usual way, meet in the usual places at the usual times, associate with one another in the accepted ways, give and receive assignments, submit and receive homework, give and receive grades, give and take tests, and so on. The only differences are those that are planned or that follow from the planned ones may be more important in producing the observed effects than the differences planned. This is true in any sort of intervention study, whether in the field or in the laboratory, but the possibilities for unanticipated changes are greater in a field setting than a laboratory setting because of the richness of the context and the relative lack of control of it by the experimenter.

A field experiment offers more convincing evidence of the impact of different text materials and different uses of text materials than more controlled laboratory experiments and also more convincing evidence than nonintervention studies in the field. The single most impressive study in this area is the work of the National Longitudinal Study of Mathematical Abilities (NLSMA), headed by the late Edward G. Begle (1968-1972). The study was extremely complex and ambitious. It was a longitudinal study covering a six-year period and involved hundreds of thousands of children in classrooms around the country. The part of the study that is of greatest interest for us in the present context compared the achievement test results in math of students who had studied under one textbook with the scores of similar students who had studied under a different book.

Figure 1 reproduces the results reported by NLSMA for students who had studied from one particular textbook. The horizontal axis of the figure corresponds to a series of math achievement scales. Each scale consisted of six to ten items selected to be theoretically homogeneous. The first group of scales covered computation, the second comprehension, the third application, and the fourth analysis. The vertical axis is arranged so that the line marked "50" is the fiftieth percentile of the entire population taking that scale. Students who studied from this textbook did slightly worse, on the average, than students in the rest of the population in computation but substantially better on the other scales. The NLSMA data were able to show achievement differences traceable to the use of different textbooks. The differences are apparent in spite of the diversity of communities, schools, classrooms, teaching strategies, student aptitudes, and other variables. Such data as these are quite convincing and extremely useful. For example, if the editors of this textbook were wise, they would look for a way to strengthen the teaching of computation while sacrificing as little as possible on higher outcomes. Over a couple of years' time, they could determine just how much of a sacrifice in higher order outcomes they would have to make to secure a given amount of improvement in computation.

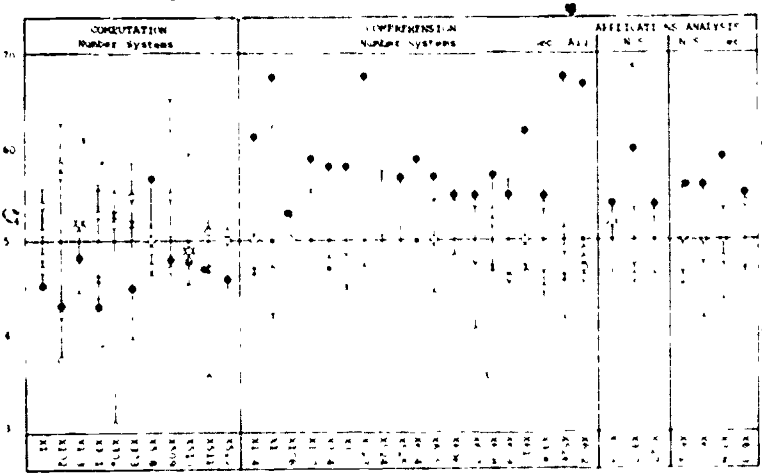


Figure 1. Profile of Achievement for Students Who Studied from One Textbook (darkened circles) Compared to Students Who Studied from Others (open circles), Taken from the National Longitudinal Study of Mathematical Abilities

The field experiment demonstrates, under realistic conditions, the connections between treatments that are fully comparable to those that might actually be implemented and outcomes fully representative of those of the greatest interest to educators and educational policymakers.

This does not deny the value of other types of studies. The work that Andrew Porter and his colleagues (1978) at Michigan State have been able to do by content analysis of texts and tests is extremely valuable, for example. Observational studies of classrooms have added enormously

to the understanding of the activities of teaching and studenting and of the importance of how classroom time is spent. However, many of these studies represent textbooks indirectly in such categories as "seatwork" and "homework." Reading between the lines of these studies, it becomes apparent that much of what is being studied is being guided by an agent, the textbook, which is not represented in the data. In-depth interviews with teachers and students, ethnographic studies of classrooms, attempts to model the planning activities of master teachers, qualitative evaluations in the manner of Elliot Eisner's (1976) connoisseurship and criticism all have something to contribute to our understanding. But if whatever they contribute suggests nothing about the relationship between practices and outcome, it is difficult to see what help it is to the policymaker or decision maker. And if they do suggest relationships between practices and outcomes, it should be possible to detect those relationships in field experiments. These other techniques of study give their greatest benefit when used in conjunction with field experiments.

For these reasons, it is important to design and carry out field experiments to show connections between textbooks and schooling outcomes. Two general issues in the design of field experiments need to be considered: the question of the scale of study and the question of subject selection. Begle's (1968-1972) NLSMA study has literally a cast of hundreds of thousands. If that is what it takes to find relationships between text features and outcomes, then we might as well forget it. Such studies are too time consuming and expensive for the benefits they yield.

Carl Thoresen (1978) has shown how valid and extremely useful experiments can be done with single subjects. In a counseling situation, a counselor may work with a client to help her or him lose weight. Data may be collected on time of eating and amount consumed for two weeks or so before a treatment is undertaken. This establishes a baseline of the client's eating behavior. Once the treatment has been administered, similar data show changes, if any, in eating behavior. And the bathroom scale shows results. At some point, the treatment may be temporarily abandoned to see if the original situation returns. In this way, it is possible to use the client's pre- and post-treatment behavior as a comparison situation and to make quite reliable inferences about the effects of the treatment on this single subject. The conclusion: Field experiments may be done on a quite small scale.

Some theorists of research methodology regard as a true experiment only those studies in which subjects are randomly assigned to control and treatment conditions. Only in this way, they argue, can we estimate the likelihood that differences we find were not already present prior to administration of the treatment. Random assignment of subject to treatment is a sure way to limit this particular source of error, but this is not the only source of error in an experiment nor is it often the worst source. In educational experiments in particular, poorly described, weak treatments and inappropriate outcome measures are often more serious threats to the validity of conclusions that is failure to bound the error of subject assignment by randomization. Furthermore, since students are rarely assigned to classes randomly in the field,

random assignment is a threat to the external validity of a field experiment.

The remainder of this paper will discuss a variety of issues that must be considered in designing field experiments of various sorts to study the connections between textbooks and their educational consequences. It will focus on two broad families of issues: treatments and outcome measures.

What Is the Treatment in Studies of Textbooks?

A textbook becomes a treatment only when it is put to use by teachers and students. The impact of a given feature of a textbook may be altogether altered by the way a teacher uses it. This happens when an inquiry-oriented textbook is taught in a rote fashion. It also happens when a good teacher supplements a bad textbook or makes educational capital out of its shortcomings by having students address it critically as the scientific community would critically examine a scientific book. The most immediate determinant of the outcomes of a textbook considered as a treatment is what the student does to and with it. In the most brute sense, a textbook, unread, teaches nothing. A textbook read in a passive, uncritical fashion does not yield the same results as one read with a mind full of good, relevant questions. A textbook that works well with the student who constantly tries to picture real-world applications of concepts may work poorly with the student who looks mainly for abstract conceptual relationships.

For these reasons, treatments can be considered as having three dimensions: features of the textbook itself, how teachers use the textbook, and how students use the textbook. Direct connections between the textbook itself and student outcomes will be apparent only when the uses to which teachers and students put the textbook are quite similar in spite of variations in personality, aptitude, background, and context; or when the book's influence is the same regardless of how it is used. This latter situation could easily occur if the book is either extremely powerful or extremely weak in its impact. There is also the situation in which the number of teachers and students studied is so large that textbook effects show up in spite of any variation in use. Since the concern here is with designing small-scale field experiments, this possibility will be neglected. Generally speaking, teacher and student use will need to be part of the treatment in field experiments.

The Textbook Itself

Anyone interested in textbooks should spend some time listing all the features that might be different on one page of any text. Making this suggestion is like the ancient Egyptian astrologer-mathematician's asking for two grains of wheat on the first square of the chess board,

four on the second, sixteen on the third, and so on until he lost his head when the kingdom's treasure was about to be exhausted.

Just the physical aspects of the page--margins; type size; type-face; composition; layout; weight, content, and finish of paper; color of paper and ink; and so on--could keep one busy for an hour or so. Then there is content. How does one represent the content of a page of text? As a list of words and phrases? A list of sentences? A hierarchical tree of ideas or concepts? A set of Venn diagrams? And then there are all the organizational features: paragraphs, section headings, sequence of presentation of ideas, subordinations and superordinations of ideas, relationships of temporality, of causality, of inclusion or exclusion, or entailment, and so on, expressed, unexpressed, or embodied in the organization of what is expressed. One can examine the visual aids: charts, graphs, diagrams, and so on. And then there is the language. George Newell (1980) has just reviewed some 50 or so systems researchers have devised for analyzing and categorizing written language. He says he has only scratched the surface. Figure 2 shows some of the features of textbook language he found.

Discourse types (narration, description, argument, etc.)
Topic sentences
Style (colloquial, casual, ceremonial, personal, etc.)
Coordinators (connectors between sentences)
Propositions
Inferences
Phrases, clauses
Referents
Context, tacit or explicit
Semantic structure
Logical structure
Themes
Purpose: Reference, persuasion, expression, creation of beauty
Structure/organization of propositions

Figure 2. Features of Textbook Language, from Some Ways to Analyze Text Language by George Newell

The nonspecialists only have to remember the intricacies of grammar, rhetoric, and composition to get some glimpse of the unlimited field for human ingenuity that even the stodgiest page of textbook prose represents. Finally, one can leap off the page altogether, as psychologists would have us do, to examine the demands, both explicit and implicit, that the text places on students for performances of various types, ranging from simple decoding of the words to the most abstract reasoning with the propositions.

With so many features that might be manipulated in the interest of better learning, where does one begin? To examine them individually would keep one busy for the rest of the century, and even then the interactions would still remain. Yet all these factors are clearly significant determinants of outcomes in some cases. In struggling with this conceptual jungle, several guiding principles have been helpful. First, there is no getting around the content. Content inclusion and emphasis is an absolutely vital feature of any textbook. These can be represented in various ways, and various indicators or measures of inclusion and emphasis may be employed. The question of how content is treated is a practical one; any way that works well enough for the specific purpose at hand is all right. It is important to compare the actual content of a textbook with the content of other books on the subject. Figure 3 shows some of the types of other books usable as comparisons. Differences in content inclusion and emphasis between different text treatments of a topic are absolutely vital determinants of what students will learn from using the text.

A second principle to consider is language as serving two functions in text: access and modeling. In its function of giving access, language presents content to students. It functions best in this mode when it uses forms that are familiar to students, forms that they already know how to decode and comprehend. It also serves to model forms of language appropriate to the topic, not only new vocabulary but also desirable forms of argument, of explanation, of use of evidence, the types of questions considered appropriate in this subject, the most important matters to the particular, specialized language community represented in the text. This two-fold division suggests that language in texts may fail in at least two different ways. On the one hand, it may be opaque to the students. On the other hand, it may fail to expose them to full and adequate models of, in this case, scientific language. Both of these aspects of text language can and should be manipulated experimentally and the consequences investigated.

A third guiding principle is to subsume all the other aspects of texts, except content and language, under a rubric we call presentational features. The work of Alma S. Wittlin (1978) on museum exhibit design has been extremely useful. She has brought together some speculation based upon studies of the brain and the neurophysiology of brain processes with some practical insights drawn from her work in designing museum exhibits. These come together in a most interesting way that suggests how to cut through much of the remaining complexity that textbooks present. Figure 4, on page 14, shows the main ideas in Dr. Wittlin's papers. Dr. Wittlin maintains that to be effective any exhibit must first attract the student's attention (here the danger is an exhibit that is understimulating); present the "message" in a clear, comprehensible way (here the danger is overstimulation, which results in confusion or superficial learning); and maintain attention long enough to complete the lesson. Her work also can be applied to the textbook by allowing one to analyze the content inclusions and emphasis, the language, and the presentational features of textbooks using these principles. These analyses will highlight meaningful differences among the "treatments" in different texts. Stable differences in student outcomes associated with such text differences should be detected.

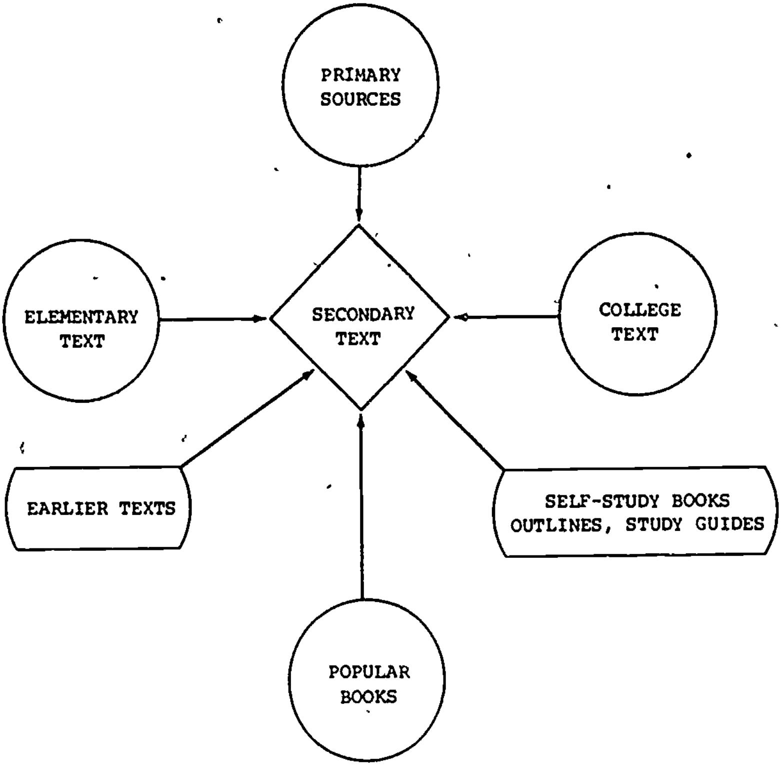


Figure 3. Sources of Comparison for the Content of Secondary Texts

Use of Text by Students

The students' use of texts was mentioned as a second dimension of any treatment. Presumably, this includes prominently what we ordinarily call studying. Studying has been an object of study by educational researchers for several decades. Anderson (1979) has provided a concise review of current thinking and research on studying. Figure 5, on the following page, summarizes Anderson's model, which he has drawn from his review. Because such a vital research tradition already exists in this area and because of the difficulties of finding relationships between outcomes and variables as detailed and microscopic as these, behavior patterns across students will not be studied. Instead, individual students' problems in studying science from the textbook will be characterized. The same text will present different problems to students who approach it differently, and it is expected that students who experience these problems will be able to point them out. The text will be examined with the students, line by line and page by page, and will be compared to the students' test results. The plan is to localize certain

INITIAL AROUSAL	ATTENDING MESSAGE RECEPTION	MAINTENANCE OF ATTENTION
DANGER: Understimulation	Overstimulation	Monotony
TO AVOID THE DANGER: Relevance, interest Dissonance Sensory appeal Appeal to affect	Chunking White space Hi signal noise Planned redundancy Integration of multiple channels Hierarchical organization under key ideas	Change modality Insert questions Vary senses used Drama of ideas

Figure 4. A System of Principles to Govern Museum Exhibit Design, from the Work of Alma S. Wittlin

difficulties quite precisely using such conceptual schemes as those of Wittlin and Anderson as guides in interpreting what students tell the researchers about difficulties they experience.

One difficulty that cannot be localized is motivation. Here the stance taken is that a primary goal of teaching in the secondary school is to make the study of the subject rewarding enough to be self-sustaining in the context of the school and classroom. That is, the satisfactions students derive from studying a textbook should be great enough to sustain the behavior of studying, given the factors built into the learning situation, such as grades, graduation requirements, and public evaluations of individual student performance.

For the time being, the hypothesis is that students who do not find study self-sustaining in this context fail to do so largely because the connections of what they are studying with anything else that is meaningful in their lives--now or in their fantasies, hopes, expectations,

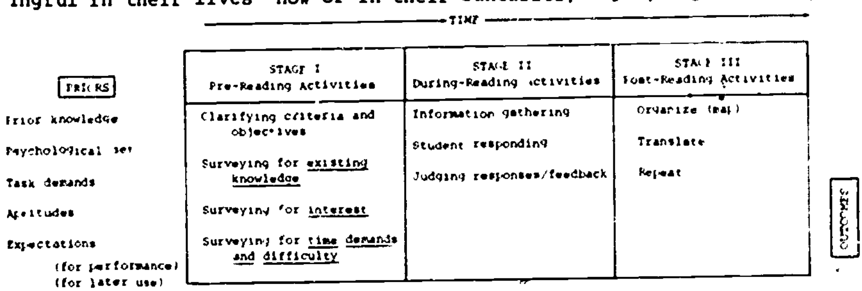


Figure 5. A Model of the Studying Process, Based on Anderson (1979)

and visions of possible future lives--are too few and too weak. Adolescent students are preoccupied with their own versions of adult preoccupations. They have little time or energy left over from their concerns for unrelated matters. It is assumed that those who do manage to bring the study of science into their lives in any significant or meaninful way must have found ways to build it into their agenda, so that progress in learning science contributes to larger life purposes for those individuals. These aspects will be examined in detail in the interviews with students about their difficulties in studying science from their textbooks. What connections do they perceive between science and their current lives or their fantasies of their lives in the future?

In looking at the specific and detailed activities of students studying textbooks, Wittrock's (1979) formulation of learning as constructed meaning will be used as a guide. In his formulation, the learner's primary task is to generate relevant relations between information and experience. According to Wittrock, "Even when learners are given...information...they still must discover its meaning" (page 10). Wittrock identifies several verbal devices, such as metaphors, similies, analogies, examples, and rules, that facilitate construction of meaning and several types of imaginal devices, such as pictures, diagrams, and flow charts. How students use such devices as well as how they manage on their own will be studied. The study will look most closely at what students do and might do when they are stuck, when they just can't get it and are ready to give up. Some ideas have been drawn from an amazing variety of sources, ranging from Polya (1977) to hypnotism, that might be useful when students are at the end of their rope and still unable to extract meaning from the information. Figure 6 presents some examples.

Isolate the problem; localize it.

Formulate the problem; state it as clearly as you can.

Express the troublesome material in some other expressive medium, e.g., draw a picture or diagram.

Look ahead; look behind; are there things you've missed in the context that would help you make sense of this portion?

If you're trying something and getting nowhere, don't persevere in an unproductive strategy: Try something else.

Use your imagination: imagery, fantasy, free association.

Do something outrageous: sing a song about the problem, act it out, make a joke about it; loosen up and get out of your rut.

Make a huge image of the offending passage, put it on your wall and make marginal comments, like graffiti.

DON'T GIVE UP! You may run out of time, but you'll still be working on it when time is called.

Be proud of your efforts, of the fact that you did not give up. Don't scold yourself for "failure."

Figure 6. To the Student: Things to Try When You're Stuck

Use of Text by Teachers

The third dimension of the textbook as treatment is how the teacher uses the textbook. This dimension will be studied in several ways. First, the teacher is viewed as an allocator of resources, particularly classroom and homework time. How the teacher performs this role quite obviously has an enormous influence on what students will or will not do with the textbook and, therefore, on what they will or will not learn. The available time on the textbook content will be mapped to reveal patterns of emphasis in the teachers' allocations. The teachers also set standards of performance. They do this in many ways, most obviously by testing and grading. Patterns of content inclusions and emphasis on teachers' tests will be examined. In addition, teachers can serve in any capacity that the textbook itself can serve: dispensing information, organizing content, making task demands, and so on. Whatever analyses can be done on texts also can be done on the the performance of a teacher: language, visuals, and content.

Teachers also can do at least two other important things beyond the power of the textbook. They can model any and all aspects of human behavior in connection with science or studying science. Also, they can establish a social situation and a social climate in the classroom that can influence what students do and how they feel. These aspects of a teacher's work will be studied in a qualitative way. Sensitive observers informed about the phenomena under study will be used, but they will be unaware of the specific treatment being implemented at the time. These observers will participate in all subsequent analysis and interpretation of all data from the field experiment so that they may suggest findings the researchers might overlook.

Outcome Measures

"Learning science" can mean many things, from learning scientific facts, to learning to behave like a scientist, to learning to love science. Figure 7, on the following page, shows some noncognitive variables identified by Messick (1979) in a recent article. It would be nice to be able to mount a battery of fine, reliable, valid instruments to assess all these dimensions of learning science. On the other hand, to use the typical conglomerate achievement tests as outcome measures would doom the entire effort to failure. In such a test, text differences would be represented by, at the most, two or three items. All the others would cover matters treated substantially the same in the texts being studied.

Therefore, conventional achievement test items that are targeted toward the specific differences being studied will be used. The criterion measure might contain eight or ten items testing materials presented in one or two contrasting paragraphs embedded in two different text treatments. Or they might focus entirely on students' ability to use and extend to other contexts certain language forms introduced in one

VARIABLES	EXAMPLES
Experiential background factors	Exposure to other cultures
Affects	Feelings about school
Attitudes	Feelings plus evaluations, attitudes toward school
Beliefs	About causes of one's successes
Interests	Feelings plus activities
Motives	Impulses toward action, such as need for achievement
Temperament	Energy, dominant mood
Social sensitivity	Empathy
Coping	Preferred ways to handle stress or threat
Cognitive style	Field dependence, independence
Creativity	Fluency of response
Values	Enduring beliefs and preferences

Figure 7. Twelve Varieties of Noncognitive Variables, from Messick (1979)

text but omitted from another. In other words, achievement tests that are a full order of magnitude (ten times) more intensive than conventional tests are planned. Such tests would obviously be unsuited to grading, so they may be administered as self-study aids just prior to the course examination that counts for a grade, or these items could be embedded in the regular examination while a grading arrangement is worked out with the teacher that corrects for the distortions.

Construction of these tailored tests will look beyond comprehensions and recall to longer term retention, ability to apply what is learned to the field of instances presented as appropriate, ability to integrate with other principles, and ability to transfer what was learned to unfamiliar situations. An unbiased assessment of these dimensions of outcomes will be made for all treatments studied.

The dissertation by Yalow (1980) at Stanford that sought aptitude-treatment interactions between verbal and spatial aptitudes and performance on a topic in economics presented with verbal and figural (visual) elaborations has influenced this research. On an immediate post-test, Yalow found that both the group who had verbal elaborations and those who had figural elaborations outperformed the group that had simply been given the basic, unelaborated presentation. On a delayed retention test, however, the group given the basic, unelaborated treatment outscored both of the other groups, showing not only a lower rate of forgetting but even additional learning after the experiment was over. This research plans to test students' knowledge of only a few important and difficult matters within a topic, but it will be tested very thoroughly.

An attempt will be made to estimate the likelihood that students will continue to study science voluntarily. A feature of a textbook that does not have an immediate effect on achievement may nevertheless so impress some students or excite their imagination that they sustain a high level of interest and effort in learning science. Such effects would be extremely powerful when continued over several key years in a student's career.

Conclusion

Many of the issues discussed in this paper deserve more elaborate and detailed treatment. An equal number of issues, almost as important, have been left out. These include establishing collaboration with teachers in the field, developing rapport with teenage subjects so they will open up their hearts and minds about their experiences with textbooks, selecting topics, and selecting student populations for study. The most important matter to be discussed, data, could not be reported. Although these wrestlings with the complexities of design are somewhat superficial and preliminary, the ideas presented may serve as a basis for continuing discussions of analyses of textbooks and their uses.

The point the author would like to leave you with is John Dewey's (1929): Educational practice is the test of the adequacy of any educational science. As we gain greater understanding of educational processes, we will be tempted to look to the laboratory for confirmation. But the real test is in the field where ideas developed in the laboratory or the study or the library are put to the test of practice. Educational researchers must develop the means to make this test of practice as powerful and as sensitive as possible.

Acknowledgement

I would like to acknowledge the assistance of my associates in this work. Without them, none of this could have been done. They are, in no particular order, Sara Guri, George Newell, Don Dake, Barbara McEvoy, Steven Leitz, Carol Stone, Denis Cronin, Robert Curley, and Susan Hanson.

References

Anderson, T. H. Study skills and learning strategies. In H. F. O'Neill, Jr., and C. D. Spielberger (eds.), Cognitive and affective learning strategies. New York: Academic Press, 1979, 77-97.

Begle, E. G., et al. National longitudinal study of mathematical abilities reports. (25 vols.) Stanford University: The Board of Trustees of Leland Stanford University, 1968-1972.

Cronbach, L. J. (ed.). Text materials in modern education. Urbana, Illinois: University of Illinois Press, 1955.

Dewey, J. and Dewey, E. Schools of tomorrow. New York: E. P. Dutton, 1915.

Dewey, J. Sources of science education. New York: Liveright, 1929.

Eisner, E. Educational connoisseurship and criticism: Their form and function in educational evaluation. Journal of aesthetic education, 1976, 10, 175-185.

Elashoff, J. D. and Thoresen, C. E. Choosing a statistical method for analysis of an intensive experiment. In T. Kratochwill (ed.), Single subject research. New York: Academic Press, 1978, 287-310.

Macdonald-Ross, M. Graphics in text. Review of Research in Education, 1977, 5, 194-231.

Messick, S. Potential uses of noncognitive measurement in education. Journal of educational psychology, 1979, 71, 281-292.

Newell, G. Some ways to analyze text language. Working paper no. 1. Ford Foundation Textbook Project, Stanford University, 1980.

Polya, G. How to solve it: A new aspect of mathematical method. Garden City, New York: Doubleday, 1977.

Porter, A. Relationships between testing and the curriculum. Occasional paper no. 9. Institute for Research on Teaching, College of Education, Michigan State University, 1978.

Stake, R. and Easley, J. Case studies in science education, Vol. 2. Washington, D. C.: Government Printing Office, 1978.

Yalow, E. Individual differences in learning from verbal and figural materials. Unpublished doctoral dissertation, Stanford University, 1980.

Wittlin, A. S. Building bridges between behavioral and biological sciences. Occasional paper of the Institute for Neuropsychological Studies, Palo Alto, California, 1978.

Wittrock, M. C. Cognitive movement in instruction. Educational Researcher, 1979, 8, 5-11.

Chapter 3

ANALYSIS OF SCIENCE TEXTBOOKS: IMPLICATIONS FOR AUTHORS[1]

Bonnie B. Armbruster and Thomas H. Anderson, University of Illinois

The basic premise of this paper is that certain characteristics of text can optimize learning, given that the learner has sufficient background knowledge and engages in the appropriate studying activities. First, characteristics of the text that empirical research has shown to affect learning outcomes will be identified. Then these rather abstract features will be translated into concrete recommendations for writing and evaluating the prose in a textbook.

Research Evidence for Effects of Various Text Characteristics on Learning

This paper is organized around four characteristics of "well-written" prose commonly mentioned in rhetoric and composition textbooks: structure, coherence, unity, and audience appropriateness. The role of each of these features will be defined and the evidence and conclusions regarding the function of the feature in the learning process will be reviewed briefly.

Structure

Structure is the feature that has received most attention in the research literature. Very simply, structure refers to how the ideas chosen for inclusion in the text are put together in a linear order.

---

[1]A similar version of the paper, entitled, "Studying Strategies and Their Implications for Textbook Design" was delivered by T. H. Anderson at the Conference on "Designing Usable Text," in Moreton-in-Marsh, England, November, 1980. The Conference was sponsored by the British Open University and the U. S. Army Research Institute. This paper will be published in the proceedings of that conference, also. The preparation of this paper was supported in part by NIE under contract No. US NIE-C-400-76-0116 and in part by NICHD under contract No. HD-06864.

From a fairly extensive body of research, it appears that structure or organization influences the amount as well as the kind of knowledge acquired from reading. With respect to the amount of knowledge acquired, many empirical studies have yielded the following results:

1. Connected discourse is much easier to learn and remember than collections of unrelated sentences or lists of words (Goetz and Armbruster, 1980).

2. Among texts, those that are better organized are better remembered (Goetz and Armbruster, 1980; Shimmerlik, 1978).

3. Providing readers with information about the conceptual organization or superordinate structure of text can facilitate recall. This information can be provided either directly or indirectly. A direct means is through "signaling"--devices such as titles, transitions, and connectives used by authors to emphasize aspects of structure. Information about structure can be provided indirectly by establishing a set of expectations for structure through serial presentations of materials having different content but similar structure (Pearson and Camparell, in press; Shimmerlik, 1978).

4. Readers are likely to recall much more information from text if they identify and use the author's structure than if they use another organizational scheme or no apparent organization (Meyer and Freedle, Note 1; Meyer, Brandt, and Bluth, Note 2).

Structure of text also affects the kind of knowledge acquired from reading by influencing the way information is stored. The better organized the text, the more highly integrated the memory representation is likely to be. Highly integrated memory representations enable learners to consider related facts simultaneously, which is a necessary condition for higher-order cognitive processes such as inferencing, summarizing, and decision making (Walker and Meyer, 1980; Frase, 1972).

In sum, the better the structure of the text, the more likely the reader is to remember the information and to engage in the higher-level cognitive processes that are usually considered to be the important outcomes of a learning situation.

Recent literature in cognitive psychology and artificial intelligence has begun to suggest what is meant by "better" structure. The currently popular notion is that memory is organized in high-level, domain-specific cognitive structures (Iran-Nejad, Note 3), variously called "frames" (Minsky, 1975), "scripts," (Schank and Abelson, 1977) and "schemata" (see, for example, Anderson, 1977: Rumelhart and Ortony, 1977; Rumelhart, 1977). A model of reading comprehension based on these notions views reading comprehension as the process of choosing and verifying conceptual schemata to account for the to-be-understood text. "Better" structure is thus defined with respect to the reader's schemata: The closer the match between the text structure and the reader's cognitive structure, the better the comprehension and recall.

Coherence

Coherence means "a sticking together." With reference to text, coherence refers to how smoothly the ideas are woven together. In a coherent discourse, the relationships among ideas must be clear enough so that there is a logical connection or "flow of meaning" from one idea to the next. Compared to an incoherent discourse, a coherent discourse makes it easier for the reader to perceive the message as an integrated unit. Coherence operates at both global and local levels: that is, at the level of the whole text as well as at the level of individual sentences.

At the global level, a text is coherent to the extent that it facilitates the integration of high-level ideas across the entire discourse. Features that contribute to global coherence include headings and outlines, advance organizers, and visual displays or diagrams. Headings and outlines contribute to global coherence by enabling the reader to see relationships among higher-order ideas. The term, headings, refers to the headings and subheadings embedded in text while the term, outlines, refers to tables of contents or other intact outlines presented as a supplement to the text.

Although headings and intact outlines are commonly included in textbooks, research supporting their effectiveness as text-processing aids is scarse. However, some studies have indicated that headings and outlines can have an important effect on comprehension. For example, Bransford and Johnson (1972) demonstrated that a title presented before a short narrative passage that was deliberately constructed to be ambiguous strongly determined the interpretation given to the passage. The following sample paragraph from the study dramatically illustrates the effect of foreknowledge of theme on coherence and, hence, comprehension.

> The procedure is actually quite simple. First you arrange things into different groups. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, otherwise you are pretty well set. It is important not to overdo things. That is, it is better to do too few things at once than too many. In the short run this may not seem important but complications can easily arise.
>
> A mistake can be expensive as well. At first the whole procedure will seem complicated. Soon, however, it will become just another facet of life. It is difficult to foresee any end to the necessity for this task in the immediate future, but then one never can tell. After the procedure is completed, one arranges the materials into different groups again. Then they can be put into their appropriate places. Eventually they will be used once more and the whole cycle will then have to be repeated. However, that is part of life.

The paragraph seems quite incoherent without the title, "Washing Clothes"; with the title, more of the ideas "stick together" and make sense. Similar facilitative effects for titles have been obtained by Schallert (1976) and Anderson, Spiro, and Anderson (1978).

With respect to more typical textbook material, Dansereau, Brooks, Spurlin, and Holley (Note 4) found that incorporating headings into a 2,500-word passage extracted from an introductory science textbook significantly improved the performance of college students on a delayed recall measure.

Results of research concerning the effect of outlines are inconsistent. While Christensen and Stordahl (1955) failed to find any facilitative effects for a number of organizational aids, including intact outlines, other studies have found positive effects of outlines. For example, Eggen, Kauchak, and Kirk (1978) found that outlines presented along with a 1,000-word text significantly improved the comprehension of fourth, fifth, and sixth graders. Brooks, Dansereau, Holley, and Collins (Note 5) found that a sequence of outline-passage-glossary led to significantly better performance than the passage alone on recall and recognition measures over a basic science text. In sum, although the evidence is not clear cut, it appears that headings and outlines can improve the comprehension and recall of text.

Another device that can contribute to global coherence is an "advance organizer" (Ausubel, 1960, 1978). An advance organizer is a short set of verbal or visual information presented prior to a larger body of text; organizers "are presented at a higher level of abstraction, generality, and inclusiveness than the new material to be learned" (Ausubel, 1978, page 171). Presumably, an organizer serves either of two functions: (a) it provides an organization or structure as an assimilative context that would not otherwise have been present, or (b) it activates an organizational scheme from the reader's prior knowledge that might not otherwise have been used to assimilate the new material (Mayer, 1979). Both functions can be interpreted as ways of ensuring global coherence. Results from the 50 or so studies that tested the effectiveness of organizers indicate that sometimes organizers appear to facilitate learning and at other times they do not (Barnes and Clawson, 1975; Hartley and Davies, 1976; Ausubel, 1978; Mayer, 1979).

Global coherence can also be enhanced by visual displays, diagrams, or charts. Although researchers in the field speak of the "infancy of diagram theory and research" (Holliday, et al., 1977, page 137), some evidence is accumulating that diagrams can be effective and efficient means of communication.

Holliday and his colleagues have conducted several studies involving the use of diagrams in science instruction. In two very similar studies (Holliday, 1975; Holliday and Harvey, 1976), high school science students either read a text description or read the same description along with labeled drawings. In both studies, subjects who read the text plus diagrams attained significantly higher scores on multiple-choice posttests than did subjects who read the text alone.

Holliday (1976) prepared three different versions of instructional materials on the oxygen, carbon dioxide, nitrogen, and water cycles: (a) a block-word diagram (an abstract representation of ideas and the relationships between them similar to a flow chart), (b) a picture-word diagram (the same information as in the block-word diagram but conveyed largely through pictures), and (c) text only. High school biology students studied from either the text alone, the block-word diagram alone, the picture-word diagram alone, the text plus block-word diagram, or the text plus picture-word diagram. Adjunct questions accompanied the materials. Analysis of the results on a multiple-choice test, developed from rephrased and recombined adjunct questions, revealed that the block- and picture-word diagrams alone were superior to the text alone or the text in combination with a diagram. Holliday suggests that flow diagrams increase the chances of the learner forming mental links among verbal labels because a verbal description has been replaced with a condensed and spatially integrated display of block figures and design elements. He states,

> An instructional package consisting of a single flow diagram allows the learner easier and more immediate access to all critical chains and the interrelationships among these chains. This form of instructional condensation into a single "manageable" display also allows the learner to view the total or "big" picture at a glance (1976, page 64).

Within the framework of a course in basic electronics, Gropper (1970) compared the effectiveness of conventional instruction with a programmed lesson consisting of a brief film, a reference booklet comprised solely of diagrams, and a workbook containing problems about information conveyed in the diagrams. Thus, the majority of the content of the programmed lesson was contained in the diagrams. On a criterion test with multiple-choice and constructed response items, the programmed lesson group made significantly higher scores than the control group. In addition, the programmed lesson took considerably less time to complete than the conventional lesson. Gropper attributed the effectiveness of the technique to two properties of diagrammatic presentation: the visual, spatial organization of ideas and the capacity to portray the "big picture." In sum, evidence from the research on diagrams can be interpreted as support for the potential of diagrams to enhance the coherence of the to-be-learned material.

Some features contribute to global coherence of a text; others play a role in local coherence. At the local level, features related to coherence help the reader integrate information within and between sentences. One important local feature is connectives or phrases that function conjunctively. These include linguistic connectives that make explicit the temporal, causal, spatial, or conditional relationships between propositions (Halliday and Hasan, 1976).

Some research has demonstrated a facilitative effect on learning an explicit statement of connectives (rather than leaving the connective implicit and requiring readers to infer it). In a study by Katz and Brent (1968), both first and sixth grade children preferred descriptions of causal relationships where a connective made the relationship

explicit, rather than those where the relationship was implicit. Pearson (1974-75) reported a similar finding: Given a choice as to the surface form for a causal relationship, fourth graders preferred to have the relationship stated explicitly, even though such explicit statement results in a more grammatically complex sentence. In a follow-up study, students were asked to read individual sentences in which the causal relationship was either made explicit or left implicit by omitting the connective. Results indicated that if a sentence was not recalled in a unified form, there was a 50% chance that it would not be recalled at all. Furthermore, in almost two-thirds of the cases in which subjects were asked to read sentences with only an implicit causal relationship, a connective was included in recall.

Pearson concluded that connectives have a strong effect on the salience of causal relationships and may facilitate the integration of ideas in memory. Finally, Marshall and Glock (1978-79) found that explicitly stated logical structures facilitated the recall of "not-so-fluent" community college readers, whereas "truly fluent" college readers were not so affected by the presence or absence of explicitly stated relationships. The authors speculated that the better readers had more knowledge to bring to bear in interpreting the meaning of discourse and filling in needed inferences, whereas poorer readers were forced to depend to a greater extent on information explicitly stated in the surface structure of the text.

Another line of research has examined the effect on comprehension and recall of text coherence as manipulated by reference and repetition. Lesgold (1972) investigated the effect of pronominalization and showed that a shared reference promotes connection of propositions into a single, higher-level memory unit. Similarly, deVilliers (1974) manipulated the use of articles (definite versus indefinite) to influence the likelihood of perceiving a series of loosely connected sentences as a story. Subjects who treated the sentences as a thematically connected story recalled more words and sentences than subjects who treated the sentences as unrelated.

Stone (Note 6), working with adult subjects, also demonstrated that when readers are given relevant information to carry across sentence boundaries, they are able to read naturally occurring text much more rapidly than when those local cohesive ties are broken. Kintsch, et al. (1975), constructed texts containing either few concepts with many repetitions and interconnections or many concepts with fewer repetitions and interconnections. The less-connected passages required more reading time. Finally, Manelis and Yekovich (1976) varied the number of repeated concepts in sentences and found that the less-connected sentences (those with fewer repetitions) took longer to read and that immediate recall of the less-connected sentences was inferior when reading time was fixed.

In sum, features of text contributing to both global and local coherence appear to help readers encode, store, and retrieve the material as a structured, integrated unit.

Unity

Unity refers to the degree to which the text addresses a single topic. A text with unity conveys a sense of closure and completeness with respect to the topic and does not stray from the topic by including irrelevant and distracting information.

Little research bears directly on the question of how unity affects comprehension. However, some empirical support for the notion that unity promotes comprehension can be found in a series of studies reported by Reder and Anderson (1980). In these studies, college subjects studied a chapter from an introductory college textbook in the original form and in summary forms that were about one-fifth the length of the originals. Results on true-false tests from all of the experiments indicated that learning materials from summaries is at least as good as reading the original text and often superior. Also, the acquisition of new material is better if information learned earlier on a related topic was learned by reading a summary. The authors speculated that the effect of summaries was due to helping students focus attention and avoid having to time-share between main points and details. In terms of the feature of unity, the summaries have greater unity than the original text because they contain the pith of the content without the encumbrance of distracting details.

Audience Appropriateness

Audience appropriateness refers to the extent to which the text matches the content knowledge of the reader. The research most related to audience appropriateness has been concerned with the effect of prior knowledge on comprehension. It is quite clear from the literature that relevant topic knowledge has strong effects on comprehension. For example, in a study by Spilich, Vesonder, Chiesi, and Voss (1979), subjects with high and low knowledge of baseball heard a description of a fictitious baseball game and then attempted to recall the text. High knowledge subjects were more likely to recall more information in a more accurate sequence and to elaborate on certain elements.

In another study investigating the effect of prior knowledge, Anderson, Reynolds, Schallert, and Goetz (1977) constructed two ambiguous passages, one concerning either a jailbreak or a wrestling match and the other concerning either a musical soiree or a card game. Music and physical education students read the passages. Their interpretations of the stories were predictable from their backgrounds, and their recall was characterized by background-relevant intrusions.

Another body of research related to the effect of prior knowledge on comprehension was focused on vocabulary, or word knowledge. In their review, Anderson and Freebody (Note 7) conclude from many studies that "word knowledge is strongly related to reading comprehension." A considerable body of research on readability also supports the preeminent role of word knowledge in reading comprehension (see Klare, 1974-75 for a review).

In sum, evidence has been presented to support the claim that text characteristics such as structure, unity, coherence, and audience appropriateness can influence the process of comprehension. They can affect how rapidly text is processed, the amount of information remembered, and the structure of that information. In the next section, the use of these characteristics as guidelines for developing clearly written textbooks will be explained.

An Ideal Textbook

A fantasy--a description of an ideal textbook--is a useful way to begin. The textbook is being used in an introductory biology course as a means of conveying information about the structure and principles of biology to students who are relative novices in the discipline. The students are expected to read and learn from the textbook in order to complete the course successfully.

The preface to the textbook tells how the approach of this textbook compares to the approach used by other introductory textbooks in biology. The preface also gives the overall purpose and scope of the textbook by presenting a brief overview of how the major divisions of the textbook are related to each other. The introduction to the first division does the same for the chapters of the division, as does the introduction to Chapter 1 for the sections of that chapter. From the introduction to Chapter 1, the student could almost predict the title of the first section. Furthermore, the title suggests how the content of that section might generally be organized.

As the students read on, they find that the section _is_ organized as anticipated. In fact, for each heading, subheading, and topic sentence, students are able to predict how the following text will be organized. The text seems to flow naturally from idea to idea; ideas are clearly related to each other, with explicit connectives and referents. The text is "clean" and not littered with details and tangential information. The content does not seem too difficult, and occasional new terms are carefully defined. Particularly complicated text is supplemented by charts and diagrams that clarify the relationships.

Toward the end of the chapter, students notice that they have not been underlining or highlighting anything because the format of the chapter has made important information stand out. Furthermore, it has been very easy to take notes and make an outline. At the end of the chapter, students find that a summary recapitulates the main ideas of each section and again states how the sections are interrelated. A final statement provides a transition to the next chapter. And so students proceed through the textbook, always with a clear sense of where they are now, where they have been, and where they are going.

This fantasy was designed to portray a textbook exemplifying the desirable features of structure, unity, coherence, and audience appropriateness. Producing a textbook exactly as described may not be

feasible now, but some ideas about how to begin can be proposed. The rest of the paper presents these ideas.

Anatomy of a Textbook

The Text Unit

The basic building block of the textbook is the text unit. The text unit consists of the author's purpose, which can be stated in the form of a question that the author is addressing, plus the response to that question. The question and response are composed of idea units, which are approximately equivalent to phrases or pausal units as defined by Johnson (1970). Thus, the idea units within a text unit are either question idea units or response idea units.

Types of Text Units. The structure of a text unit depends on the type of purpose or question addressed. From the analysis of intermediate and secondary level content area textbooks, the conclusion reached is that most informative prose is characterized by only a few structures reflecting a few basic purposes.

Four text units have been identified (purposes and corresponding text structures) that seem to be fundamental to informative text (Table 1, page 30). One purpose of an author is simply to describe a topic, that is, to address the questions, "What is X?" The structure that responds to this question/purpose is a description. A description is a list of the essential features or properties of a topic.

A second purpose of an author may be to exemplify a topic, that is, to answer the question, "What are some examples of A?" An exemplification structure consists of a list of objects, events, or processes, etc., each of which has the defining attributes of the topic.

A third purpose may be to convey the sequence of events in time, or to answer the question, "When did (or should) these events occur in relation to each other?" The structure that responds to this purpose/question is a temporal sequence, in which events are ordered from the earliest-occurring to the lastest occurring.

A fourth purpose may be to explain a topic, or to address the "Why?" and "How?" questions associated with the topic. The corresponding structure is an explanation. An explanation gives the causal or enabling relationships among events or among the various elements in a description. Although most of the relationships in an explanation text unit are causal or enabling, some may be temporal.

Table 1 lists these structures and several types of purposes and questions associated with each structure.

Table 1. Relationship of Structures, Purposes, and Questions

Pattern	Examples of Purpose or Question	
	Imperative Form	Interrogative Form
1 Description	Define A. Describe A List the features of A	What is A? Who is A? Where is A?
2. Exemplification	Give examples of A.	What are some examples of A?
3 Temporal Sequences	Trace the development of A. Give the steps in A.	When did A occur? (in relationship to other events)
4 Explanation	Explain A Explain the cause(s) of A Explain the effect(s) of A Draw a conclusion about A. Predict what will happen to A Hypothesize about the cause of A	Why did A happen? How did A happen? What are the causes/reasons for/effects/outcomes/results of A? What will be the effects/outcomes/results of A?

The Frame

The fundamental text units just described may be embedded in higher-order text structures called frames. A frame reflects the typical, high-probability ways of organizing information and represents a combination of text units. Frames are the prose embodiment of the construct of schemata discussed in the first section of the paper. The writers believe that there are a few general schemata and a larger but finite number of more specific schemata associated with each content area or discipline. These schemata are manifest in content area textbooks as frequently repeated frames.

The most widely used general frames in textbooks seem to be the compare/contrast, problem/solution, and definition/example frames. The purpose of the compare/contrast frame is to answer the question, "How is X similar to/different from Y?" The structure that responds to this question is a compare/contrast. In this structure, the quantitative and qualitative differences between the features or events of two or more descriptions, temporal sequences, or explanations are explicitly noted.

The purpose of the problem/solution frame is to answer the question, "How did solution X help solve problem Y?" The structure that serves to organize the answer to that question is the problem/solution frame. The "introductory" problem/solution frame, which is first introduced in the elementary grades, requires simply that the so-called problem events or characteristics be considered separately from the so-called solutions. A more sophisticated problem/solution frame suggests a slightly different structure: (a) the chain of events leading up to the problem event are considered first, (b) the discovery of the solution event or events are discussed next, (c) the effect of the solution event(s) on the original chain of events is discussed next, and finally, (d) the original chain of events is compared/contrasted with the newer chain of events (including the solution event). The newer chain of events is the preferred one, provided the solution was successful.

The third frequently used general frame is the definition/example frame. This frame helps to answer the question, "Define and give examples of A." The frame simply presents a definition, which is a listing of the major critical features, characteristics, or traits, plus several examples and, perhaps, relevant nonexamples.

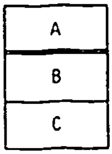
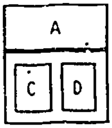
In addition to such general frames, textbooks also include content-specific frames. An example of a content-specific frame in science is the "scientific theory" frame described by Dansereau (1980). The frame includes six major categories of information related to scientific theories: Description, inventor/history, consequences of the theory (how the theory influenced humans), evidence for the theory, other similar or competing theories, and an open category of extra information. Another frame might be "process," which includes categories of information on the function(s) of the process, an explanation of how the process works, and where the process takes place.

Each of the categories of information in a frame corresponds to a basic text unit. That is, each category contains an implicit question and suggests the kind of information appropriate as a response to that question. For example, in the "scientific theory" frame, one text unit is composed of the implicit question, "How has this theory influenced people?" and the response to that question assumes the structure of an explanation. It is surely not the case that all text units are subsumable in frames; however, it appears that frames are an important and common construct.

Mapping: Representing Text Units and Frames Diagrammatically

The writers have devised a technique for representing text units and frames in a visual display or diagram. The technique is called "mapping." Mapping uses symbols as a prose-free way to designate the relationships that convey essential meaning in text. Table 2 presents the basic relationships used in mapping and their corresponding symbols.

Table 2. Basic Relationship Symbols in Mapping

Relationship		Symbol
B and C are characteristics or properties of A	(characteristic)	
C and D are examples of A	(example)	
A occurs before B	(temporal)	$A \longrightarrow B$
A enables B	(enabling)	$A \rightsquigarrow B$
A causes B	(causal)	$A \Longrightarrow B$
A is less than B A is greater than B A is similar to B A is identical to B	(compare/ contrast)	$A < B$ $A > B$ $A \approx B$ $A = B$
Negation	(negation)	X
and or but	(logical connectives)	-and- -or- -but-

These relationships are usually identified in text by a few "key words" or other standard linguistic devices.

An important feature of mapping is that maps take on distinctive shapes depending on the type of text unit or frame being represented. Table 3, below, and Table 4, on page 34, show the mapped representations for the text units and some of the frames discussed previously.

Table 3. Types of Text Units and Their Corresponding Maps

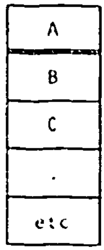
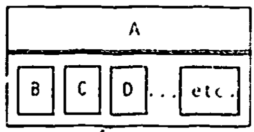
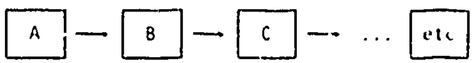
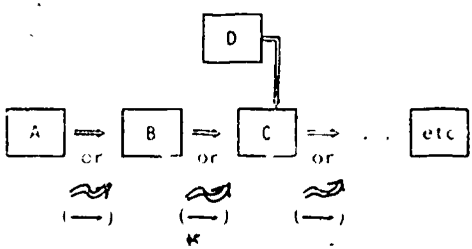
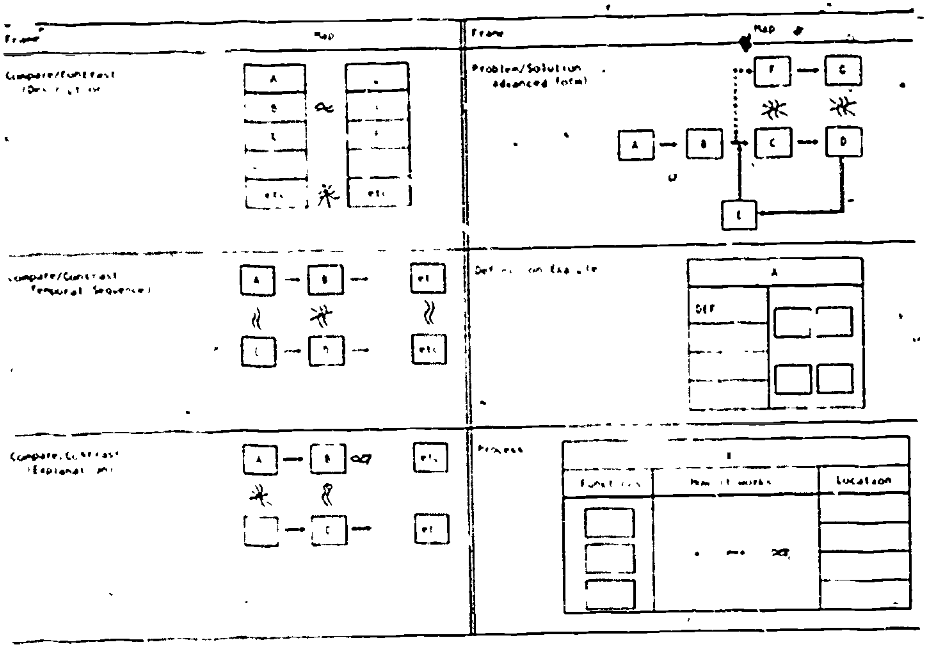
Text Unit	Map
Description	
Exemplification	
Temporal Sequence	
Explanation	

Table 4. Some Examples of Frames and Their Corresponding Maps



The Textbook Hierarchy

A textbook is a hierarchical arrangement of text units and frames. The author begins with some very broad purposes or questions. The responses to these questions in turn suggest other component subquestions. And so the hierarchy expands in a top-down fashion, with each question giving rise to a new frame or text unit.

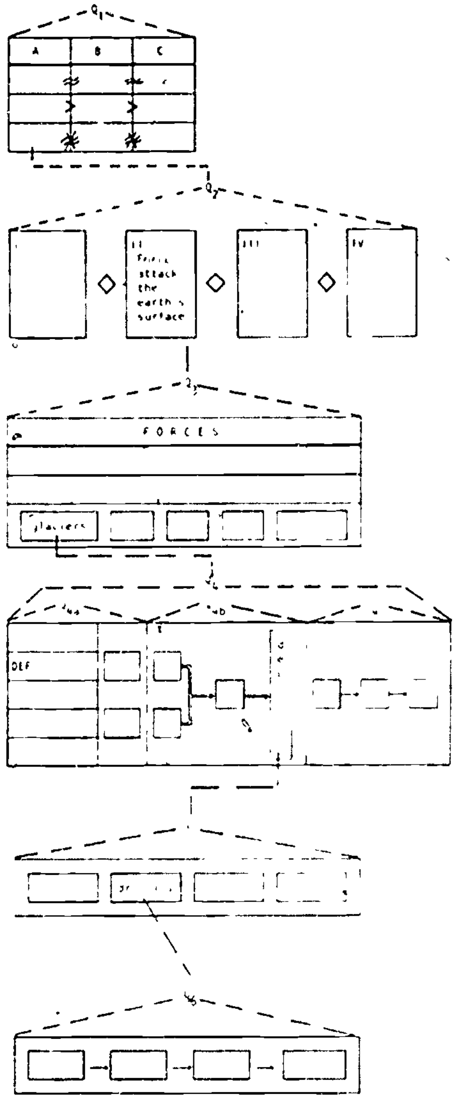
Table 5 presents a partial textbook hierarchy for a hypothetical introductory earth science textbook. The question at the top of the textbook hierarchy is "compare and contrast this textbook with other earth science textbooks." A point of contrast between this textbook and other textbooks is a function of the particular schema or organizing framework the author brings to the content area. In the hypothetical example, this schema is conveyed in a frame that includes the question, "Explain the dynamic forces of the past and present that have shaped the earth." Since the precise nature of that frame is not known, the relationships connecting the component text units have not been indicated in Table 5.

One of the response idea units within the frame of this example suggests the question, "Describe the forces that attack the surface of

Table 5. A Partial Hierarchy for a Hypothetical Introductory Earth Science Textbook

Questions	Hierarchy

- Q_1 Compare and contrast this textbook with other earth science textbooks
- Q_2 Explain the dynamic forces of the past and present that have shaped the earth
- Q_3 Describe and give examples of the forces that attack the [illegible] face of the earth
- 4 Des[illegible]
  - 4a [illegible]
  - 4b [illegible]
  - 4c [illegible] the events [illegible]
- 5 [illegible] examples [illegible] glacier
- Q_6 Explain how drumlins were formed



41

the earth." The response to this question is a frame consisting of description and exemplification text units. One of the response idea units of the exemplification text unit spawns the question, "Describe glaciers." This question is handled by a frame composed of the following frames and text units: definition/example; explanation; and temporal sequence. A response idea unit in the explanation text unit suggests the question, "Give some examples of deposits left by glaciers," and thus an exemplification text unit. One of the response idea units within the exemplification structure in turn suggests the question, "Explain how drumlins were formed," and the corresponding explanation text unit.

The preceding sections defined text units and frames as the basic building blocks of a textbook and explained how text units and frames are hierarchically arranged to form a textbook. The final section of the paper offers some guidelines for evaluating and writing text units/ frames and textbooks.

Guidelines for Evaluating and Writing Text Units and Frames to Achieve Structure, Unity, Coherence, and Audience Appropriateness

1. Each title, heading, and topic sentence should readily suggest (a) a unique purpose or question of the type listed in Table 1, or (b) a frame.

Examples of some inappropriate titles or headings: "New genetic types and agriculture," "How fast are impulses?" and "Following the inheritance of two traits at once."

Examples of appropriate titles, headings, and topic sentences:

"Two kinds of skeletal systems are found in organisms." (The text goes on to compare/contrast exoskeletons and endoskeletons.)
"An engineering problem: How is water transported in plants?" (The text gives three explanations for water transport in plants.)
"Early classification systems." (The text gives examples and descriptions of early classification systems, according to a frame.)
"The respiration process." (The text describes the function, how the process works, and where it is located, according to a frame.)

12

2. The structure of the response to the question should match the type suggested by the question.

dendrite — branch of a nerve that carries impulses to the nerve cell body

axon — long nerve fiber that carries impulses away from the nerve cell body

synapse — a small gap between nerve cells

A poor example: "<u>How do nerves carry signals</u>? Each nerve in the body is really a bundle of very fine nerve fibers. Nerve fibers are too thin to be seen except with a microscope. Each nerve fiber is part of a nerve cell.

"The nerve cell has a very irregular shape. It has a number of small branches sticking out, almost like the branches on a tree. These are called dendrites (den' drit). The word dendrite comes from a Greek word for 'tree.'

"At one end of the cell is a particularly long branch. It is called an axon. This axon is a single nerve fiber, as shown in the drawing.

"The message that travels along the nerve cell is a small electrical charge. It usually starts at the dendrites, goes through the cell body, and then moves out along the axon. This electrical charge is called the nerve impulse.

"Nerve cells are strung together in chains. The branches of the axon of one cell come close to the dendrites of another cell. The parts of the different cells don't quite touch. There is a small gap between them that is called a synapse (si'naps)." (Note how the title/question is never answered. Instead, various components of the nervous system are described.)

Examples of close correspondence between titles and text structure:

"<u>A description of the eyes</u>. The eyes are nearly spherical in shape. Muscles are attached to the eyes and to bony sockets in the skull. These muscles enable the eyes to be moved in all directions. The center of the front of the eye is clear and colorless. This part of the eye is called the cornea. A short distance behind the cornea is a circular, colored band, the iris. In the center of the iris is a circular opening, the pupil. Immediately behind the iris and pupil is a clear, elastic structure, the convex lens. Muscle tissue and tissue composed of fibers hold the lens in place. The retina is a thin layer of nerve cells that coats almost all the inner surface of the eyball. The nerve cells of the retina are especially sensitive to light. The retina is connected to the optic nerve, which in turn is connected to the cerebrum. Clear fluids fill the internal chambers of the eye and help maintain its shape."

"<u>Vision explained</u>. Light rays coming from an object pass through the cornea, enter the pupil, and reach the lens. The lens bends the light rays, which then form an image on the retina. This image is formed in the same way that a camera lens forms an image on film. The image on the retina produces impulses that reach the brain by way of the optic nerve. When the impulse arrives at the special area of the cerebrum that controls vision, we become aware that we see something."

One way to check whether the structure of the response matches the question is to map the text unit or frame, using the following procedure:

(a) Predict the structure of the text unit or frame from the title, heading, or topic sentence.
(b) Select the corresponding map shape from Table 3 or 4.
(c) Map the text unit or frame.
(d) Compare the shape of the map produced in Step c with the map shape selected in Step b.
(e) If the shapes match, for the most part, the structure is appropriate. If the shapes do not match, the author apparently used a structure that was not indicated in the title, heading, or topic sentence; or began with the appropriate structure, but then introduced a new text unit or frame that was not indicated by a title, heading, or topic sentence.

A simple example: "Dikes are a kind of igneous intrusion. Dikes are flat sheets of igneous rock that cut across the rock layers they intrude. They were formed when magma was forced into vertical cracks in the rocks. They come in all lengths and thicknesses. Their rock is usually basalt or diabase. Sills are also a kind of igneous intrusion. Sills are sheets of igneous rock that are parallel to the layers they intrude. They were formed when magma was forced along bedding plates between rock layers. Sills can be hundreds of meters thick and many kilometers long. They, too, are usually basalt or diabase."

(a) From the topic sentence, it appears that the text unit will be a description responding to the question, "Describe dikes."
(b) The corresponding map shape is:

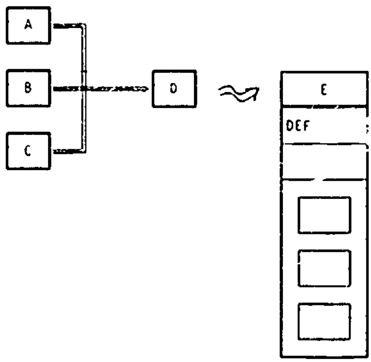
A
B
C
etc

(c) The map of the text unit is:

Dikes		Sills
a kind of igneous intrusion	≈	a kind of igneous intrusion
flat sheets of igneous rocks	≈	sheets of igneous rock
cut across the rock layers they intrude	≉	parallel to the layers they intrude
formed when magma was forced into vertical cracks in the rocks	≉	formed when magma was forced along bedding plates between rock layers
come in all lengths and thicknesses	≈	can be hundreds of meters thick and many kilometers long
rock is usually basalt or diabase	≈	rock is usually basalt or diabase

(d) The shapes of the maps in Steps b and c do not match.
(e) Therefore, the conclusion is that the author did not indicate in the topic sentence the structure actually used. A more appropriate topic sentence might have been, "Dikes and sills are two different kinds of igneous intrusions." Alternatively, the subtitle, "Dikes and Sills: A Comparison," could have been used.

Of course, the text of the example was simple enough so that the disrepancy between indicated and actual structures was easy to perceive without actually mapping the text unit. However, structural changes sometimes are subtle enough that they are not readily apparent unless they are portrayed graphically. For example, authors sometimes signal an explanation, which begins smoothly enough but then changes when the author decides to define and give examples of some event in the causal sequence. The mapped representation may appear as follows:



The map makes it obvious that two text units or frames are involved; both should be clearly indicated in the text by titles, headings, or topic sentences.

3. The paragraph should be the prime vehicle for conveying a text unit. That is, the question component of a text unit is suggested in the topic sentence (or in the immediately preceding heading) and the response units are presented in the body of the paragraph.

Example: See the example paragraphs in Guideline 2.

4. Since a frame is composed of two or more text units, more than one paragraph is usually required to convey it.

Example: "Two kinds of skeletal systems are found in organisms." (The text goes on to describe the exoskeleton in one paragraph and the endoskeleton in a second one.)

An obvious exception to this rule is the definition/example frame in which both the definition and the examples are often represented in one paragraph.

Example: "The Hook. Waves and crosscurrents may drive the end of a spit towards the shore. Such a spit--with a curved end--is called a hook. Sandy Hook in New Jersey is a well-known hook. Others are Cape Cod in Massachusetts and Rockaway Beach on Long Island, New York. Although currents built them, they were raised above sea level by the action of waves and winds."

5. Each idea unit in the text unit should contribute to the answer to a question. An idea unit that is not clearly contributing to an answer but still seems to fit should be signaled as such by using phrases such as, "Incidentally,...," "As an aside,...," "If you are interested,...." A large number of irrelevant idea units should be edited out or should form the basis of another text unit.

Example: "Describe a conditioned response. This type of behavior is learned when the nervous system links together two events that occur at nearly the same time. When a conditioned response has been formed, the response that ordinarily would have been the reaction to one stimulus becomes the reaction to a second stimulus. The second stimulus becomes a substitute for the first one. Although forming a conditioned response usually requires repetition of the two stimuli, it does not necessarily require that one be conscious of what is going on. You might be interested to know that Ivan Pavlov, A Russian scientist, was the first to experiment with the conditioned response. He worked primarily with dogs in his laboratory."

6. Enough relevant idea units should be included in the text unit to form a complete answer to the question (taking into account the prior knowledge of the readers.)

Example of a text when too little information is given: "Describe how blood is transported through the body. The numerous arteries that branch off the aorta carry blood to various organs and systems. Then veins return the blood to the heart. Among these pathways are those to the digestive organs, the limbs, the head, the kidneys, and the walls of the heart."

Example of a more appropriate text: "Describe how blood is transported through the body. Blood is transported by the circulatory system, which is composed of a network of tubes (arteries, veins, and capillaries), a pump (the heart), and a fluid (blood) that moves. The heart pumps the blood by a series of contractions and expansions from its chambers into the arteries. Arteries take blood to all organs of the body. Blood enters an organ through arteries and their branches. Arteries divide and subdivide until they enter networks of capillaries. Blood also exits the organ through capillaries that are connected to veins. Veins then conduct the blood back to the heart."

7. The most direct way of signaling the relationship among idea units should be used.

Example of indirect ways of signaling a relationship: "In the evening, the light fades. Photosynthesis slows down. The amount of carbon dioxide in the air spaces builds up again. This build up of carbon dioxide makes the guard cells relax. The openings are closed."

An improved example: "What happens to these processes in the evening? The fading light of evening causes photosynthesis to slow down. Respiration, however, does not depend on light and thus continues to produce carbon dioxide. The carbon dioxide in the air spaces builds up again, which makes the guard cells relax. The relaxing of the guard cells closes the leaf openings. Consequently, the leaf openings close in the evening as photosynthesis slows down."

8. Clear referents for pronouns and indefinite noun phrases should be provided. Also, when an identical idea is referred to several times in a text unit, the same noun or noun phrase should be used.

Example: "Photosynthesis in the leaves starts with the morning sunlight. Carbon dioxide was made during respiration all night. It was stored in the air spaces around the spongy cells. This carbon dioxide is used as the leaf starts to make sugar. When a <u>certain</u> <u>amount</u> has been used up the guard cells respond. They become stiff, swell out, and make openings in the leaf." (In this passage, students become confused as to whether <u>certain</u> <u>amounts</u> refers to sugar or to carbon dioxide.)

9. Technical terms or other difficult vocabulary words should only be introduced when learning their meanings is an intrinsic part of learning the content. When such vocabulary is required, clear definitions should be provided.

Example: "A plateau is an area of horizontal rock layers that has high relief. Relief is simply the difference between the highest and lowest points of a region. There is no fixed amount of relief for a plateau. As a rule, however, a plateau's relief is 1,000 meters or more. Its high points may be well over 1,000 meters above sea level. Its low points are the bottoms of its canyons and steep river valleys." (Note how and where the concept of relief is defined.)

10. Analogies, metaphors, and other figurative language from the reader's knowledge of the world should be used frequently, but only to supplement meaning, not to carry the meaning of the text unit. Also, these figurative language devices should be signaled, so the reader will recognize the need to shift from comprehension processing in a literal mode to that of an inferential one.

Example: "In each mitochondrion these enzymes are organized like a complicated assembly line. The final product of the assembly line is cellular energy in the form of ATP. The raw materials are foodstuffs and oxygen. The by-products are carbon dioxide and water." (The analogy of the mitochondrion being like an assembly line is clearly stated and explained.)

Guidelines for Evaluating and Writing Textbooks to Achieve Structure, Unity, Coherence, and Audience Appropriateness

1. At all levels of the hierarchy, the reader should know the author's purpose or question and the relationship among coordinate elements that respond to that purpose or question.

(a) The reader <u>may</u> be told how the textbook is similar to or different from other textbooks. If included, this information should be given in the preface to the textbook.
(b) The reader <u>should</u> be told the author's frame for the con-tent area and how that frame is reflected in the organization of the major divisions (parts) of the textbook. This information should be given in the preface to the textbook.
(c) The reader <u>should</u> be told how the chapters within each part are related to each other and to the title of the part. This information should be given in the introduction to each part.
(d) The reader <u>should</u> be told how the sections within each chapter are related to each other and to the title of the chapter. This information should be given in the introduction to the chapter.
(e) The reader <u>should</u> be told how the response idea units of the text unit are related to each other and to the question idea unit(s). This information should be given in headings, subheadings, or topic sentences.

2. Frames should be used whenever appropriate. Also, frames should be used consistently; that is, with the same categories presented in the same order. Frames should be described <u>before</u> they appear in the text, and the categories should be clearly signaled <u>within</u> the text. An example of a likely frame in biology is a "system" frame. Most biology textbooks include a good deal of information on the systems of the human body--circulatory, digestive, respiratory, etc. The descriptions of systems contain similar categories of information; for example, structure (including component parts), function, and comparisons/contrasts in different organisms. We are suggesting that "systems" be treated as a frame in the textbook. The frame would be described before the first system is introduced. The text on systems would contain the same categories of clearly signaled information presented in the same order.

3. Each text unit should be related to at least one other text unit. Consequently, the main flow of prose is the progression from one text unit to another related one. Often, the author may wish to include text units that are only slightly related to the main flow of prose. If these text units are included, they should be located in boxed-in areas, or perhaps appendices. Examples of such text units are those that

(a) teach skills that are necessary for understanding a later text unit (such as map reading or direction finding using a compass).
(b) relate content area ideas to familiar knowledge of students.

(c) lend some authenticity to certain ideas in the textbook (such as excerpts from Freud's letters or Darwin's diary).
(d) describe people and personalities involved in the science.
(e) have high interest value because of their unusual and/or attractive features (like including a unit about the "Venus fly-trap," a plant that feeds on insects, in a chapter about how animals feed on plants.)

These boxed-in text units can be a valuable resource to textbook authors. They can use them in a variety of ways to improve the overall countenance of a textbook, without increasing the risk of making the basic prose more difficult to comprehend. It is common knowledge that the prose of many content area textbooks is brutally boring, even when written clearly, and that many students will understandably shy away from reading them frequently or for long periods of time. The solution to the problem of how to make textbooks more enticing and fun to read lies not in the manipulation of the basic text units, but rather in the cleverness that authors and editors can use to discover and develop intrinsically motivating boxed-in units.

4. A textbook should provide only as much detail as is required by the abilities of the students to understand it and the needs of the students to learn it. If the abilities and needs of the students vary greatly, provide text units over the same topic at several clearly signaled levels of detail.

Example: "Note the two levels of detail used to explain the hydrogen pathway in the biology chapter, Appendix A."

5. When the textbook content is so complex that a prose text unit is cumbersome, other visual presentations, such as charts, diagrams, tables, and photographs, should be used to supplement the prose.

The preceding guidelines for evaluating and writing text units and frames and textbooks are proposed to help ensure that text has the characteristics of structure, unity, coherence, and audience appropriateness. The writers believe that, in the hands of clever and conscientious authors and editors, textbooks based on these guidelines may approximate the "ideal textbook" of the fantasy presented in an earlier section of the paper. That is, a textbook written with these recommendations in mind may be easily read and understood by students. The assumption is made that learning will increase both quantitatively and qualitatively and that this learning will accrue <u>without</u> the heavy use of studying strategies. Such a textbook may help the teacher work <u>with</u> the textbook rather than <u>around</u> it. That is, the teacher can use class time supplementing the textbook rather than translating or interpreting it. Finally, it is hoped that when learning from science textbooks becomes easier, students will develop a more positive and receptive attitude toward the disipline.

The appendix to this paper contains an example of a short chapter that was written to exemplify some of the guidelines we have proposed in this paper. The content of the example is based on Chapter 7 of <u>Biological Science: A Molecular Approach</u> (BSCS Blue Version, Fourth

Edition), which deals with the role of oxygen in the evolution of life processes. From this, we selected content that appeared to reflect the author's main purposes and cast this content into a textbook hierarchy with its component text units and frames. The intent of this example chapter is to illustrate some of our guidelines, specifically, the use of (a) a "process" frame and its clear signaling in the text, (b) "layered" presentation of content to accommodate the varying needs of the audience, (c) diagrams to supplement cumbersome prose descriptions of complicated processes, and (d) boxes or marginal notes for content that does not directly respond to the author's purpose or question. The example is not meant to function as a "stand alone" text, nor does it include everything that should be said about cell respiration. Furthermore, there is no suggestion that the chapter exemplifies other desirable features of text, such as interest and aesthetics. However, it is hoped that the chapter conveys some sense of the role of structure, unity, coherence, and audience appropriateness in the comprehensibility of text.

Reference Notes

1. Meyer, B.J.F., and Freedle, R. The effects of different discourse types on recall. Princeton, N. J.: Educational Testing Service, 1978.
2. Meyer, B.J.F., Brandt, D. M., and Bluth, G. J. Use of author's schema: Key to ninth graders' comprehension. Paper presented at the meeting of the American Educational Research Association, Toronto, March, 1978.
3. Iran-Nejad. The schema: A structural or a functional pattern (Tech. Rep. No. 159), University of Illinois, Center for the Study of Reading, February, 1980.
4. Dansereau, D. F., Brooks, L. W., Spurlin, J. E., and Holley, C. D. Headings and outlines as processing aids for scientific text. Manuscript submitted for publication.
5. Brooks, L., Dansereau, D. F., Holley, C. D., and Collins, K. W. The use of outlines and glossaries as pre and post organizers in prose processing. Paper presented at the Annual Meeting of the Southwestern Educational Research Association, Houston, Texas, February, 1979.
6. Stone, V. F. The effect of textual cohesion on the comprehension of connected discourse. Unpublished manuscript, 1979.
7. Anderson, R. C., and Freebody, P. Vocabulary knowledge and reading. (Reading Education Report No. 11), University of Illinois, Center for the Study of Reading, August, 1979.

References

Anderson, R. C. The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro, and W. E. Montague (eds.), Schooling and the acquisition of knowledge. Hillsdale, N. J.: Erlbaum, 1977.

Anderson, R. C., Reynolds, R. E., Schallert, D. L., and Goetz, E. T. Frameworks for comprehending discourse. American Educational Research Journal, 1977, 14, 367-381.

Anderson, R. C., Spiro, R. J., and Anderson, M. C. Schemata as scaffolding for the representation of information in connected discourse. American Educational Research Journal, 1978, 15, 443-450.

Ausubel, D. P. In defense of advance organizers: A reply to the critics. Review of Educational Research, 1978, 48, 251-257.

Ausubel, D. P. The use of advance organizers in the learning and retention of meaningful verbal material. Journal of Educational Psychology, 1960, 51, 267-272.

Barnes, B. R., and Clawson, E. U. Do advance organizers facilitate learning? Recommendations for further research based on an analysis of 32 studies. Review of Educational Research, 1975, 45, 637-659.

Bransford. J. D., and Johnson, M. K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 717-726.

Christensen, C. M., and Stordahl, K. E. The effect of organizational aids on comprehension and retention. Journal of Educational Psychology, 1955, 46, 65-74.

Dansereau, D. F. Learning strategy research paper. Paper presented at the NIE/LRDC Conference on Thinking and Learning Skills. University of Pittsburgh, October 8-12, 1980.

deVilliers, P. A. Imagery and theme recall of connected discourse. Journal of Experimental Psychology, 1974, 103, 263-268.

Eggen, P. D., Kauchak, D., and Kirk, S. The effect of hierarchical cues on the learning of concepts from prose materials. Journal of Experimental Education, 1978, 46(4), 7-10.

Frase, L. T. Maintenance and control in the acquisition of knowledge from written materials. In J. B. Carroll and R. D. Freedle (eds.), Language comprehension and the acquisition of knowledge. Washington, D. C.: Winston, 1972.

Goetz, E. T., and Armbruster, B. B. Psychological correlates of text structure. In R. J. Spiro, B. C. Bruce, and W. F. Brewer (eds.), Theoretical issues in reading comprehension: Perspectives from cognitive psychology, artificial intelligence, linguistics, and education. Hillsdale, N. J.: Erlbaum, 1980.

Gropper, G. L. The design of stimulus materials in response-oriented programs. AV Communication Review, 1970, 18, 129-159.

Halliday, M.A.K., and Hasan, R. Cohesion in English. London: Longman's, 1976.

Hartley, J., and Davies, I. K. Preinstructional strategies: The role of pretests, behavioral objectives, overviews, and advance organizers. Review of Educational Research, 1976, 46, 239-265.

Holliday, W. G., Brunner, L. L., and Donais, E. L. Differential cognitive and affective responses to flow diagrams in science. Journal of Research in Science Teaching, 1977, 14, 129-138.

Holliday, W. G., and Harvey, D. A. Adjunct labeled drawings in teaching physics to junior high school students. Journal of Research in Science Teaching, 1976, 13, 37-43.

Holliday, W. G. The effects of verbal and adjunct pictorial-verbal information in science instruction. Journal of Research in Science Teaching, 1975, 12, 77-83.

Johnson, R. E. Recall of prose as a function of the structural importance of the linguistic units. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 12-20.

Katz, E., and Brent, S. Understanding connections. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 501-509.

Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., and Keenan, J. M. Comprehension and recall of text as a function of content variables. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 196-214.

Klare, G. R. Assessing readability. Reading Research Quarterly, 1974-1975, 10, 62-102.

Lesgold, A. M. Pronominalization: A device for unifying sentences in memory. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 316-323.

Manelis, L., and Yekovich, F. R. Repetitions of propositional arguments in sentences. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 201-312.

Marshall, N., and Glock, M. D. Comprehension of connected discourse: A study into the relationship between the structure of text and information recalled. Reading Research Quarterly, 1978-1979, 16, 10-56.

Mayer, R. E. Can advance organizers influence meaningful learning? Review of Educational Research, 1979, 49, 371-383.

Minsky, M. A framework for representing knowledge. In P. H. Winston (ed.), The psychology of computer vision. New York: McGraw-Hill, 1975.

Pearson, P. D. The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. Reading Research Quarterly, 1974-1975, 10, 155-192.

Pearson, P. D., and Camparell, K. B. The comprehension of text structures. In J. Guthrie (ed.), Reading comprehension and education, in press.

Reder, L. M., and Anderson, J. R. A comparison of texts and their summaries: Memorial consequences. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 121-134.

Rumelhart, D. E. Schemata: The building blocks of cognition (CHIP Rep. No. 79). La Jolla: University of California, Center for Human Information Processing, December, 1977.

Rumelhart, D. E. and Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, and W. E. Montague (eds.), Schooling and the acquisition of knowledge. Hillsdale, N. J.: Erlbaum, 1977.

Schallert, D. L. Improving memory for prose: The relationship between depth of processing and context. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 621-632.

Schank, R., and Abelson, R. Scripts, plans, goals, and understanding. Hillsdale, N. J.: Erlbaum, 1977.

Shimmerlik, S. M. Organization theory and memory for prose: A review of the literature. Review of Educational Research, 1978, 48, 103-120.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., and Voss, J. F. Text processing of domain-related information for individuals with high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 1979, 18.

Walker, C. H., and Meyer, B.J.F. Integrating information from text: An evaluation of current theories. Review of Educational Research, 1980, 50, 421-437.

THE PROCESS OF RESPIRATION

In this chapter, the process of respiration will be discussed within the "process" frame. That is, the order of presentation will be *functions, evolution, sequence,* and *location.*

Functions of Respiration

Respiration is an essential life process that has two main functions:

1. It breaks down the complex organic molecules of food (glucose, carbohydrates, fats, and proteins) into simpler molecules that can be used as building blocks for other life-sustaining processes.

2. In breaking down the organic molecules, it releases the chemical energy stored in the bonds of these molecules and stores it as ATP, a form of chemical energy readily usable by cells. One molecule of glucose releases 36 ATPs of energy.

The Evolution of Respiration

The life processes of fermentation and photosynthesis evolved before respiration. (See the summaries for Chapters 5 and 6 if you're a little shaky on the evolution of fermentation and photosynthesis.) The gradual addition of oxygen to the atmosphere by the process of photosynthesis began over two billion years ago. As the oxygen level of the atmosphere kept rising, living cells developed a very efficient way of obtaining energy from organic compounds, using the oxygen. This efficient way of producing energy is called *respiration*.

The Sequence of Respiration

Because of its complexity, only two stages of respiration will be described here: the carbon pathway and the hydrogen pathway. This section contains two versions of the carbon pathway and the hydrogen pathway. The first is a shorter, simplified version. The second is a longer, more detailed version.

The Carbon Pathway: The Simple Version

Function: The major function of the carbon pathway is to break the large carbon compounds of foodstuffs into smaller carbon compounds that are used in other processes to help maintain cell life.

How It Works: The first phase of the carbon pathway is simply fermentation, a process described in Chapter 5. It starts with a potential foodstuff (glucose, carbohydrate, fat, protein) and yields two molecules of pyruvic acid, a three-carbon compound. The pyruvic acid is first decomposed to one molecule of carbon dioxide and a two-carbon compound. The two-carbon compound becomes an "active" form of acetic acid. This acetic acid next enters the Krebs cycle. In the Krebs cycle, acetic acid (the two-carbon compound formed at the end of fermentation) yields carbon dioxide, hydrogen, and four- and five-carbon compounds.

The Hydrogen Pathway: The Simple Version

Function: The major function of the hydrogen pathway is the gradual release of energy in the form of new ATP molecules. Each molecule of glucose releases 36 ATPs of energy.

How It Works: The hydrogen pathway starts when hydrogen atoms are released by some of the reactions of the carbon pathway. The hydrogen atom (or more precisely, its electron) is then transported along a complex network of enzymes called the respiratory chain, until it combines finally with oxygen. As the hydrogen electrons are passed along on their journey toward oxygen, energy is released in the form of new ATP molecules in several discrete stages.

It is important to the cell that the energy is only gradually released through the respiratory chain because a sudden release of energy might damage the structure of the cell.

A good analogy to the process of respiration is a complicated assembly line. Just as workers control an assembly line, enzymes control the process of respiration. The raw materials at the beginning of the assembly line are foodstuffs and oxygen. The final product is cellular energy. The by-prod .:ts are carbon dioxide and water.

The Carbon Pathway: The Detailed Version

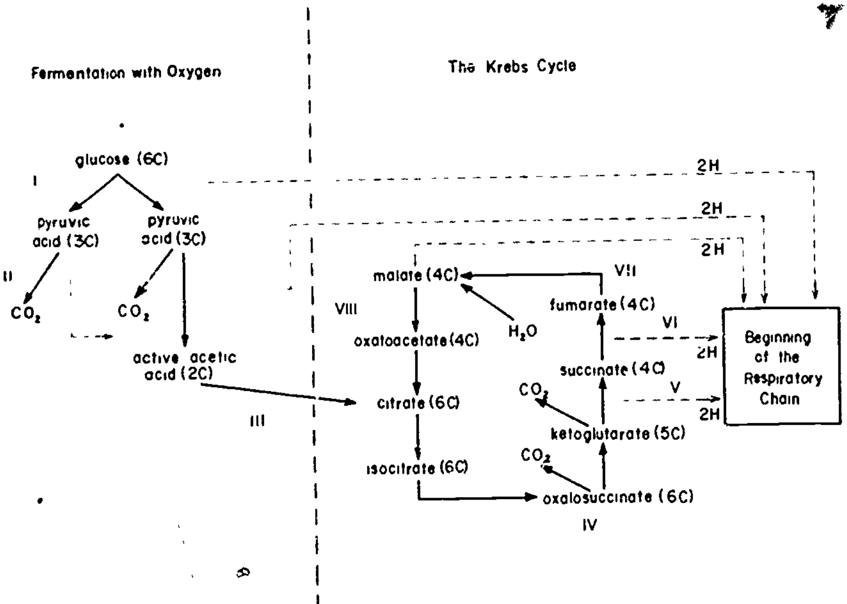


Figure 1.

Stage I: In this example (Figure 1), the process begins with one molecule of glucose. Each molecule of glucose breaks down into 2 molecules of pyruvic acid (3C) plus 3 hydrogens.

Stage II: Each pyruvic acid molecule is decomposed into one molecule of carbon dioxide (1C), one molecule of acetic acid (2C), and 2 hydrogen atoms. The remainder of the discussion traces the reactions from only one of these pyruvic acid molecules.

Stage III: Acetic acid combines with oxaloacetate (4C) to give citrate (6C).

Stage IV: After changing to isocitrate (6C), and then to oxalosuccinate (6C) while yielding 2 hydrogens, the change is then to ketoglutarate (5C) plus one free molecule of carbon dioxide (1C).

Stage V: Next, the ketoglutarate (5C) splits off a carbon dioxide (1C) molecule to give succinate (4C).

Stage VI: Succinate yields 2 hydrogens plus fumarate (4C).

Stage VII: Fumarate (4C) plus a water molecule yields malate (4C).

Stage VIII: Malate (4C) yields 2 hydrogens plus oxaloacetate (4C), which is ready to combine with acetic acid (2C) and begin the cycle again starting with Step III.

The Hydrogen Pathway: The Detailed Version

Figure 2.

Stage 1: Figure 2 shows that NAD receives two hydrogen atoms (2 electrons accompanied by 2 protons) from foodstuffs that were generated by one of the reactions in fermentation and/or the Krebs cycle.

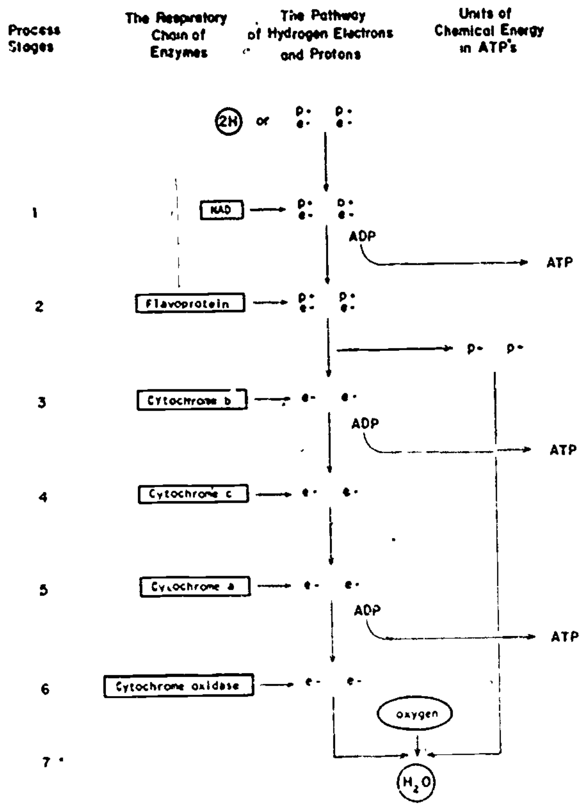
Stage 2: The two hydrogen atoms are transferred to a flavoprotein whose active group is riboflavin (vitamin B). In this transfer, one ATP molecule is usually formed from ADP.

Stage 3. The pair of electrons held by the flavoprotein is now transferred (without the protons) to two molecules of cytochrome b. The proton pair go into the surrounding solution.

Stage 4: Two molecules of cytochrome b transfer a pair of electrons to two molecules of cytochrome c, and during this transfer one more molecule of ATP is formed.

Stage 5: Two molecules of cytochrome c transfer a pair of electrons to two molecules of cytochrome a.

Stage 6: The electron pair from two molecules of cytochrome a are transferred to two molecules of cytochrome oxidase and one molecule of ATP is produced--a total of three molecules.



Stage 7: Finally, two electrons and two protons combine with one atom of oxygen to form one molecule of water. In the process, three ATP molecules have been produced. The respiratory chain of enzymes remains intact to transport the next two hydrogens.

Note that cytochromes can carry only one electron at a time, whereas NAD and flavoprotein can carry two electrons at a time. For this reason, two cytochrome molecules are needed to transfer a pair of electrons.

Location of Respiration

The entire process of respiration takes place in special compartments in the cell, called mitochondria. Mitochondria contain the enzymes that control the many step-by-step reactions of respiration. In each mitochondrion, these enzymes are organized in small particles that have the same pattern and composition.

Some More Facts About Mitochondria

There are many mitochondria in a single cell. Some cells have only 10 to 20; other cells have as many as several thousands. A liver cell, for example, contains about 1,000 mitochondria. A typical mitochondrion is shaped like a sausage, about two or three three micrometers (μm) long, and about one μm thick.

Chapter 4

TEXT STRUCTURE, STRATEGIES, AND COMPREHENSION IN LEARNING[1]

James Deese, University of Virginia

Scientists more than any other group of people are likely to be aware of the fact that literacy is not simply a matter of degree. Scientific literacy requires special knowledge and, above all, skill in and understanding of mathematics. But even scientists are not likely to be aware that scientific literacy requires, in addition to the knowledge peculiar to a particular discipline, the ability to grasp the structure of scientific text. The project reported here is based on the assumption that many students never learn how to cope with expository text.[2]

Text Grammar

Part of the problem, of course, in reading expository prose lies in the knowledge required to understand any given domain. But another problem is created by the need, in comprehending what is read, to relate the propositions that form the body of knowledge in a particular domain to one another. The pattern of interrelations among propositions is what is generally described as a text grammar (Fredericksen, 1975; Rumelhart, 1977; Kintsch and van Dijk 1978). Groups of propositions must be synthesized into and summarized by macrostructures if the text is to be comprehended (Fredricksen, 1975).

It is unfortunate that the term, text grammar, has come to be used to describe structures of knowledge, for it has caused many students of the problem to ignore or neglect the fact that a given domain of knowledge can be presented in many different ways. In many disciplines,

[1]The work reported here was supported under Grant 00233 under the joint sponsorship of NSF and NIE.

[2]This project is a cooperative effort between investigators from the Department of Psychology and the McGuffey Reading Center at the University of Virginia. The work reported here is the responsibility of the following people (alphabetical order): James Deese, Thomas H. Estes, John A. Rotondo, Wayne L. Shebilske, and M. Elizabeth Wetmore.

particularly scientific disciplines, textbooks on a given subject differ not so much in what they contain as in how they present what they contain. For this study, the particular way in which something is put, the sentences and how they are arranged, is called the surface structure of text. Not only does the student have to grasp the structure of knowledge in a particular field, he or she must do so by abstracting it from the special way in which the author has presented it.

Knowledge must always be realized in particular ways, and alternative realizations serve different purposes. These purposes may be consciously derived, but more often the intuitions of the moment rather than some rational plan as to how knowledge ought to be presented guide an author. Usually, the purpose behind a textbook is to present some knowledge so that persons unfamiliar with the context of that knowledge can understand it.

Reading, even in this multimedia age, is still the most important way in which people acquire knowledge of a complicated sort. While psychologists and educators have attended closely to the mechanics of the reading process, to eye movements and the perception of words and letters, they have attended little to the forms of discourse. What work exists has been mainly concerned with the correlates of style, sentence length, and word length, rather than with style itself.

Textbook Characteristics

The most obvious characteristic of textbooks is that they are written. However obvious that fact, it is necessary to say it at the outset, for there are great differences between written and spoken language. Written language, as Hirsch (1977) reminds us, cannot make use of gesture, facial expression, and the tacit understanding between speaker and hearer. Punctuation provides only an approximation to prosody.

Another aspect of text is that it is always somewhere on a continuum between writing whose purpose lies entirely within itself and writing whose purpose is to communicate something. Literary text must be judged by very different standards than text whose purpose is to communicate knowledge. There is the whole discipline of literary scholarship, the concerns of which center very closely to one end of that continuum. With the rare exception of people like Hirsch (1977), very few literary scholars have paid attention to the other end.

Children come to school the virtual masters of spoken or expressive language. Educators, in an effort to make the transition from speech to print as easy as possible, try to see that children's initial encounters with print involve expressive language. They almost invariably employ a combination of (1) children's own language in dictated stories, a technique aptly called the language-experience approach, and (2) simple narratives in texts called basal readers, anthologies of very simple

stories. These accounts and stories move the reader gradually to an ability to deal with narrative text. Children must internalize "grammars" for text of this kind, strategies of story comprehension, perhaps like those described by Rumelhart (1977) and Bransford (1979). After several years of elementary reading instruction, often following years of having been read to, children come to understand stories because the actions of the stories fit the plans they are able to generate.

As children move up the grades, the narratives they read change. It is not just a new content but a new style they must cope with. Yet, for the most part, formal training in reading is with narrative text of concrete description. About the time children begin to read a radically different kind of text, text containing abstractions, formal instruction in reading ceases. Some students learn strategies on their own for dealing with this kind of text; but many students, even some who are fluent readers of narrative, experience difficulty with textbooks.

This point was brought to the author's attention by a case history. It is important enough to warrant recounting. This case was reported by Carol Seal, Director of the Reading Laboratory at Warrenton Junior High School in Warrenton, Virginia. Ms. Seal administered an informal reading inventory to Ann, a 14-year-old eighth grader who complained that "I can read the assignments, but when I have to explain something, I can't. Then the teacher will go over it, and when she explains it, I understand it." Ann did well on the informal reading inventory. In fact, she scored about 90% comprehension on a passage from Steinbeck's <u>The Grapes of Wrath</u>. She could tell stories in her own words with precision. In contrast, when she read two selections from her basic science text (one on starfish and one on fossils), she showed persistent evidence of misunderstanding until the passage was explained orally.

There are many students like Ann, even in college, who never learn the strategies of coping with expository text. Furthermore, the surface structure that textbook writers impose on what they describe may compound the difficulties of those students. With these possibilities in mind, the researchers set out to investigate the following aspects of reading: (1) the effect of the structure of the text on comprehension, (2) how readers perceive the structure of text, (3) strategies readers use in trying to comprehend, and (4) certain characteristics of text itself.

Three Biology Textbooks

For reasons that had to do more with finding a practical point of departure than anything else, this work began with some textbooks in biology. Originally, the plan was to investigate some mathematics texts also. It was soon found that more than enough problems arose in coping with the biology texts, thus plans to deal with mathematics texts were abandoned.

Three textbooks at three different levels, seventh grade, tenth grade, and college, were chosen.[3] The books are all standard, well-known texts, though teachers in the Charlottesville public school system reported that the seventh grade text is probably more difficult than ones usually used at that grade level. The data strongly support that possibility. Although all parts of these books were examined, the research concentrated on passages from the three books that deal with the same material--speciation. Almost all of the data reported here concerns how students reacted to these three passages.

The experimental design and testing procedures will not be described in detail here. That will be done in a monograph dealing with the total scope of the research project. Seventh and tenth graders from the Charlottesville public school system and first year students at the University of Virginia were tested. Virginia undergraduates were tested on all three passages, and their data served as a reference to how skilled readers cope with these texts. The University of Virginia is highly selective as public institutions go, and it is assumed that the Virginia undergraduates are, on the average, skilled at dealing with abstract, expository text.

Students did a number of different things in various combinations. They were given two different kinds of recall tests. In one case, they were asked to write down all they could remember from the excerpt after they had studied it in much the way they would study their own texts. In the other, they were given specific questions to answer. The questions were based on the topical headings in the texts; and assuming the adequacy of the headings, all of the material in each passage should have been covered by the questions. This procedure is described as cued recall. In addition, some students took multiple-choice examinations based on the material. These are all more or less traditional ways of testing comprehension. In addition, the procedure decribed by Johnson (1970) and Brown and Smiley (1978) was followed. Students were asked to break up the passage into naturally occurring idea units and then to rate the importance of these units in the context of the whole passage.

The rating technique used in this study is a little different from that used by Johnson and Brown and Smiley. Students were told to indicate whether a given segment was important or unimportant to the passage as a whole. They were then asked to rate their confidence in that judgment. These two judgments were combined to construct a ten-point rating scale. Because these subjective segments usually contained more than one proposition, all the texts were reduced to base propositions, or propositions containing a single idea. In this way, the critical portion of each important idea could be determined.

---

[3]The textbooks are: Abraham, N., Beidleman, R. G., Moore, J. A., Moores, M., and Utley, W. J., Interaction of Man and the Biosphere: Inquiry in Life Science, 2nd ed., Chicago, Rand-McNally, 1975: Curtis, H., Biology, 2nd ed., New York, Worth Publishers, 1975; Biological Sciences Curriculum Study, BSCS Green Version: High School Biology, 2nd ed., Chicago, Rand-McNally, 1968.

Students were asked to fill out a study habits inventory and to provide certain demographic information. It was suspected that some of the material in the text was common knowledge to at least some of the students. Thus, groups of students at each level were asked to take tests without having read the passages. The results of this testing provide some estimate of the degree to which various aspects of the text would be known to the average student without reading.

The procedure for determining how many subjectively determined units each passage contained is also different from those used by earlier investigators and deserves attention.[4] Students were asked to place slash marks between units they thought were separate. Different people, of course, used different numbers of slash marks and placed them at different locations. The mean number of slash marks made by the readers was determined. That number was the most representative for the number of subjective units the passage contained. The passage was divided into the number of units, using the slash marks on which there was the highest agreement. While readers differed in the number of units they thought a passage contained, the good readers (college students) seldom differed in where they put the slash marks; rather, some readers divided more finely than others.

Experimental Results

The results are not completely analyzed, but some major features have emerged. First, as in previous investigations (e.g., Johnson, 1970; Brown and Smiley, 1978), a correlation was found between the importance of idea units and the probability that essential information in that unit will be recalled. Instead of emphasizing this correlation, as earlier investigators have, we point to its small size. For the tenth grade passage, rated and recalled by college students, the correlation was .39. It was .48 for the same students reading the college passage. These correlations account for 16% and 23% of the variance, respectively. These correlations may be low because even college students have difficulty identifying important ideas and then using study techniques that enable them to recall the ideas. Furthermore, for reasons that will also be discussed, we strongly suspect that the fault in part lies in the text and that the deficiencies are worse in the tenth grade book than in the college text.

Less important than the magnitude of the relation are the exceptions to it. Figure 1, on page 58, shows the regression of the proportion of college students recalling the essentials of a subjective idea unit upon the rated importance of that unit for the tenth grade text.

---

[4]John A. Rotondo is responsible for devising this technique. He has further developed the method into a clustering analysis for the subjective partitioning of text, to be published shortly.

This particular scattergram is exhibited because the passage from the tenth grade text exemplifies all of the problems generated from combining confusing text and inadequate skills of students in discovering the structure of text.

For convenience, the scattergram has been divided into nine regions. Notice the three points labeled 20, 28, and 21 in region IV. The recall of these ideas was above average. Unfortunately, these segments of text also are distinguished for their silliness. The author evidently forgot the wisdom in the old admonition: "Don't tell a small child not to stick a pea up her or his nose." The author tells the reader not to do something that it would never occur to anyone but a demented reader to do. He or she tells the reader how not to classify animals. He or she tells us, at great length, that animals should not be classified alphabetically. In so doing, the author points out that it would be absurd to arrange a grocery store alphabetically, so that abalones, apples, and almonds would all occur together. Nearly half of the students who read the passage remembered these two ideas, even when most of them could not recall the really essential message of the passage. Notice, however, that the college students could identify these ideas as being unimportant (any idea with a rating of five or less is unimportant). Idea number 28 is even more absurd and perhaps for that reason vivid and easy to remember. It points out that it would be biologically foolish to classify by likeness in color. For example, grouping bluebirds, bluegrass, bluefish, blue crabs, and blue spruce together.

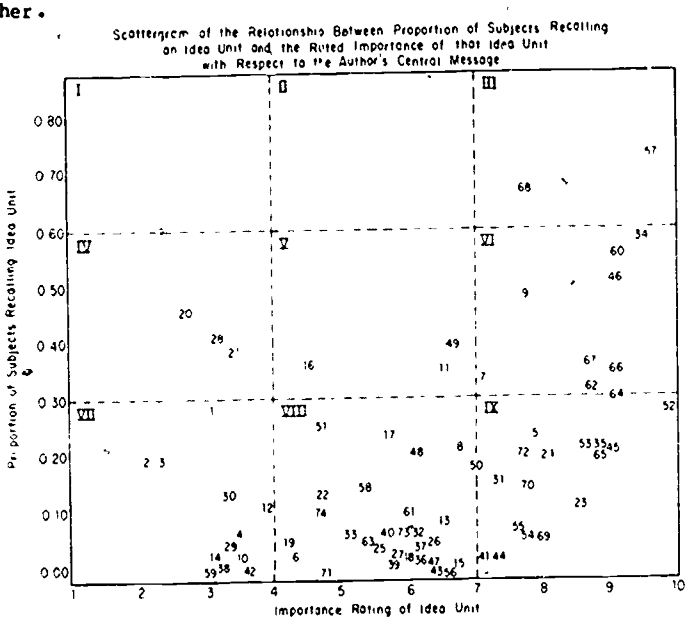


Figure 1: The Relationship between Proportion of Subjects Recalling an Idea Unit and the Rated Importance of That Idea Unit

Equally interesting and important were the large number of ideas grouped in area IX. These ideas were viewed by the students as important but were not well recalled, such as the fact that something more than structural similarity is required to group individuals together into species. These propositions were almost literally buried under examples.

In addition to tabulating the particular ideas (and propositions comprising them) recalled from the text, a system was developed for determining the quality of recall. This was necessary because some students do organize the propositions into what Fredricksen (1975) calls macrostructures. In short, some students are skilled enough to develop summary statements. Also, authors either purposefully or inadvertently leave certain propositions implicit. Sometimes students do write down things merely implied by what they have read. Therefore, simply matching ideas or propositions with the original would not accurately represent the quality of recall.

The college students worked with all three texts, and once again it was apparent from the quality of recall that the tenth grade text was by far the worst. The authors illustrated sexual isolation with the fact that wild mallard ducks and wild pintails live together but do not interbreed. The text then illustrated geographic isolation by pointing to Alaskan brown bears and polar bears. In neither case was the principle at issue named, and a majority of those who recalled the material lumped these two examples together as if they illustrated a common principle. A multiple-choice question on the material confirmed the belief that the students were simply confused by what they read about these two examples.

It was not surprising that the seventh grade and tenth grade readers were less skillful than college students. But they were also far less skillful at the ability to segment text sensibly into ideas and to judge the relative importance of those ideas. This was determined in two ways. First, both the tenth graders and the seventh graders showed more individual variation both in how they segmented text and in their judgments of importance. Such a result suggests that they had not yet come to a common standard as to how the text is structured.

We obtained more direct evidence for the tenth grade passage. The segmentation by college students coincided almost perfectly with the structural description of the surface text. The tenth graders, on the other hand, often segmented in a way that joined separate formal structures and divided others. They were far less perceptive about the structure of the text.

Before turning to the analysis of text, one additional finding requires brief discussion. Some students were tested for recall by giving them questions in the form of topical headings. When the text was clear, the college students recalled more with the topical-heading questions than they did when they were told to remember everything they could. The interesting point is that the difference for unimportant ideas was insignificant ($p = .48$), while the difference for important ideas was highly significant ($p = .01$). Providing cues enabled students

to recall more that was important without increasing the proportion of unimportant items recalled. This is true, however, only for the well-structured college text and college students. Nearly everyone wrote more when asked to answer questions in the form of topic headings, but for the seventh and tenth graders, the extra written material was irrelevant and sometimes erroneous. Even the college students working with the tenth grade passage did not do better with questions than with free recall.

The college students were not only more skilled at studying and reading than the seventh and tenth grade students, they were self-consciously so. They admitted more often to using devices intended to grasp the structure of text. These included skimming, trying to relate the author's ideas to things learned elsewhere, and trying to put the text in their own words. These and other techiques specifically endorsed more often by the college students appeared to be largely self-generated. Less than 10% of the students admitted to having taken a course in study skills or having read a book on the subject. Furthermore, fewer (44%) of the college students admitted to having been helped by teachers in studying than did the tenth graders (60%) or the seventh graders (54%). Of the college students, 83% said they largely had evolved their own methods, while 42% of the seventh graders and 53% of the tenth graders admitted to having evolved their own techniques.

Many more analyses have been completed. However, a vast portion of the data are without statistical summaries. Rather than describe the results further, the remaining pages will be devoted to the analysis of text and the relations between the structure of text and the research findings about how people go about grasping that structure.

Text Structure

The structure of a text is formed by the way in which propositions relate to one another. The notion of proposition is one of those deep primitives of intellectual inquiry. For purposes of what follows, a proposition is defined as a minimally meaningful segment of language. It is something that can be judged true or false. The word "ghosts" by itself cannot be classified as true or false, but the statement "ghosts exist" can be judged by someone to be true or false (or plausible or implausible). Propositions are embedded in the syntax of the language and, to be exhibited separately, they usually need to be removed from that syntax. There are various ways of doing that. A modified system of base phrase markers from a syntactically based transformational grammar was the method selected. Such a choice would reduce the sentence, "The old fence is sound," to two base propositions: "The fence is old," and "The (same) fence is sound." It is a modified system because it introduces certain restrictions.

A word needs to be said about the function of transformations. Chomsky (1957) described their function as that of interrelating base phrase markers. That is so; but there is a more fundamentally functional way to describe their use. Transformations enable one to state some proposition and, simultaneously, comment upon it. It allows one to tell the listener what one thinks is important and unimportant. It lets the person to whom one is talking know whether the information needed was received. Most important, transformations enable one to comment subjectively on what we say, as we say it.

In a separate project, this author examined in detail a very large number of sentences spoken by educated Americans. The most common form of complex sentence construction in such speech is one in which there is a sentential complement in the verb phrase; for example, "I believe that it's going to rain." To appreciate the subjective force of that construction, consider the series: "I suspect that it's going to rain," "I feel that it's going to rain," "I believe that it is going to rain," "I know that it is going to rain," and "I am certain that it is going to rain."

Not all relations among propositions are syntactic, i.e., within sentences. Parts of sentences are related to parts of other sentences, and whole sentences are related to other whole sentences. It is not clear why some propositional relations are within sentences and some are between, but the major limitation may be human memory. The author once rewrote a whole newspaper story as one sentence. It was incomprehensible, largely because it was impossible to remember something in it long enough enough to have it connect with a phrase to which it was related later in the discourse.

For whatever reason, certain relations among propositions are within sentences, syntactic, and certain others are between sentences. What are ordinarily called text grammars deal chiefly with the semantics of relations among propositions. There is something in discourse analogous to syntax, the way in which the particular form of discourse is organized. One critical feature of the surface structure of discourse is the extent to which context generally is necessary to place a proper interpretation on a particular proposition. Occasionally, a proposition may be correctly interpreted out of its context, but most of the time it cannot be.

This feature can be illustrated, using a paragraph from Willa Cather's novel <u>A Lost Lady</u>. The last sentence of the paragraph is "Every stick of timber was tortured by the turning lathe into something hideous." Few people have any difficulty understanding the sentence, but without context, it is almost impossible for them to interpret it correctly. The most common context people give is that of a description of some incompetent cabinetmaker's workshop.

The opening sentence of that paragraph is, "The Forrester place, as everyone called it, was not at all remarkable." Few literate persons take this for anything but what it is, the description of a house. And most people can make a number of correct inferences about the house, e.g., the Forresters are people of substance. The paragraph goes on to

tell about the porches about the house, and when we reach that last sentence we know positively that it describes the gingerbread posts on the porches. The first sentence is relatively free of contextual constraints. That is to say, the correct interpretation of it does not depend on anything else. The last sentence, in isolation, while easy to understand, is likely to be misinterpreted.

The relation of contextual constraint is one of dependency. A dependent proposition explains, modifies, elaborates, or restricts another proposition or a term in another proposition. Most commonly, propositions relatively free of constraint occur early in discourse, and those that depend directly upon it follow soon. Discourses serve different purposes, and there are countless variations on this statistical average. Shaggy dog stories often depend on having the proposition on which the whole story depends for its interpretation delayed until the end. These relationships are not semantic. Semantic relations of the sort specified by text grammars are invariant across different surface structures. It is a truism of rhetoric that something can be told in many different ways. Educators often say to students, "Tell it in your own words." This does not mean to use an idiosyncratic vocabulary, as the phrase might imply. We are telling them to produce some invariant semantic content in a way that best serves the purposes of the moment as they perceive them.

The context of a text produces a tree of hierarchical dependencies. The relationship among propositions in such a tree is transitive rather than reciprocal. That is to say, propositions that explain, modify, or elaborate do so about a proposition introduced at a higher node in the tree. The hierarchy is a strong one. A given proposition will not depend on two or more propositions located on different branches.

Linguists are used to talking about such hierarchical trees. A more familiar way to put the matter is to say that text can always be outlined. There are major entries and minor entries, which elaborate or alter the major ones. It is in this respect that the author modified the rules for breaking text into its constituent propositions. The rules were made to conform to traditional rules of outlining.

It is easier to outline some text than others. Most written text readily falls into an outline, but usually it is necessary to invent propositions implied rather than stated by the author. This is not necessarily a deficiency. Text in which every proposition is made explicit is clumsy and tedious. A good writor knows how to let the mental processes of readers work.

Figure 2 is a paragraph from a well-known textbook in experimental psychology. This paragraph illustrates many characteristics of the structure of text. This is a paragraph in which the opening sentence containing the proposition highest in the hierarchy is not the topic of the paragraph. What an English teacher would call the topical sentence is sentence number three. It contains the proposition containing the message of the paragraph, namely, that human behavior is consistent. Notice that the sentence cannot stand alone. This can bes be illustrated with the outline of the text presented in Figure 3, on page 64.

Certain business organizations have as their major task the establishment of credit ratings. Individuals as well as businesses ordinarily need some minimum credit rating in order to borrow money or open charge accounts. This whole enterprise is based on the simple but critical assumption that there is consistency in man's behavior. If the history of the financial dealings of a man shows that he regularly met his financial obligations, it is assumed that in the future he will continue to respond in the same manner. If, on the contrary, a man's history shows that he has frequently made late payments to his debtors, has often changed his place of residence to avoid being easily contacted, or has had merchandise repossessed, it is assumed that he will be a poor credit risk in the future. Both men have behaved in a consistent manner in the past, and it is presumed that each man's responses in the future will parallel to some extent those of the past: One man will continue to assume only the financial obligations he can handle, the other will not.

Figure 2. The Opening Paragraph from Experimental Psychology, 2nd ed., by Benton J. Underwood (New York, Appleton-Century-Crofts, 1966).

The phrase, "this whole enterprise" refers to the tasks of the certain business organizations described in the first sentence.

The paragraph is a kind of parable that illustrates the principle of consistency of behavior through some contrasting examples. The point is readily grasped by students. In some preliminary investigations, it was found that students nearly always get the point of passage. When they recall it they often reorganize the information to make the proposition, "man's behavior is predictable," the organizing statement.

Karmiohl (1979) presented the reactions of students to this and a variety of other segments of text in a master's thesis. He discovered that, under appropriate circumstances, people readily identified the third sentence as being the most important in the paragraph, and they were more likely to remember it than other sentences. However, that could be altered.

Examining the paragraph again, one can see that it could well be an exerpt from a chapter on consumer credit from a textbook in consumer economics. In that case, the perception of what is important would be altered. Karmiohl asked students to read the passage as if it were from a class in business. In this case, the relative importance assigned to the first three segments of text was altered as was the relative proportion of persons recalling each segment. A final point from Karmiohl's thesis has to do with the development of a structure as people read. When skilled readers read through this paragraph (and other like it) for the first time, they assume that the first few sentences are the most important ones. After the whole paragraph has been read, this idea may change. While one's perception of text is always subject to change as one reads, the process is greater in material like this paragraph, in which the essential message is introduced as a by-product of something else.

1. Organizations have as task
    1.1 Their (task)
    1.2 Major (task
    1.3 Business (organizations)
    1.4 Certain (organizations)
    1.5 [ORGANIZATIONS INCLUDE INDIVIDUALS]
    1.6 (Task is) establishment of credit ratings
        1.6.1 Individuals ordinarily need credit rating
            1.6.1.1 Some (credit rating)
            1.6.1.2 Minimum (credit rating)
            1.6.1.3 (Credit rating is) in order to borrow money
            1.6.1.4 (Credit rating is) to open charge accounts
        1.6.2 As well as businesses (ordinarily need credit rating)
            1.6.2.1 (Some credit rating)
            1.6.2.2 (Minimum credit rating)
            1.6.2.3 (Credit rating in order to borrow money)
            1.6.2.4 (Credit rating to open charge accounts)
        1.6.3 [TASK IS AN ENTERPRISE]
        1.6.4 This whole enterprise is based upon the assumption
            1.6.4.1 Simple (assumption)
            1.6.4.2 But critical (assumption)
            1.6.4.3 (Assumption is) that there is consistency in man's behavior
            1.6.4.4 [HISTORIES OF TWO MEN ILLUSTRATE ASSUMPTION]
                1.6.4.4.1 [FOR ONE MAN]
                    1.6.4.4.1.1 If history regularly shows he met obligations
                        1.6.4.4.1.1.1 His (obligations)
                        1.6.4.4.1.1.2 Financial (obligations)
                    1.6.4.4.1.2 It is assumed (BY SOMEBODY) that he will continue to respond
                        1.6.4.4.1.2.1 (respond) in the future
                        1.6.4.4.1.2.2 (respond) in the same manner
                1.6.4.4.2 On the contrary [FOR THE OTHER MAN]
                    1.6.4.4.2.1 If a man's history shows that he regularly made payments
                        1.6.4.4.2.1.1 Late (payments)
                        1.6.4.4.2.1.2 (Payments) to his debtors
                    1.6.4.4.2.2 (If a man's history shows that he) has changed place of residence
                        1.6.4.4.2.2.1 Often (changed place of residence)
                        1.6.4.4.2.2.2 (Changed place of residence) to avoid being easily contacted
                    1.6.4.4.2.3 Or (if a man's history shows) he has had merchandise repossessed
                    1.6.4.4.2.4 It is assumed (BY SOMEBODY) that he will be a risk
                        1.6.4.4.2.4.1 Poor credit (risk)
                        1.6.4.4.2.4.2 (Risk) in the future
                1.6.4.4.3 Both men have behaved in a manner
                    1.6.4.4.3.1 Consistent (manner)
                    1.6.4.4.3.2 (Manner) in the past
                1.6.4.4.4 It is presumed (BY SOMEBODY) that responses will parallel to some extent those of the past
                    1.6.4.4.4.1 Each man's (responses)
                    1.6.4.4.4.2 (Responses) of the past
                1.6.4.4.5 One man will continue to assume only the obligations
                    1.6.4.4.5.1 Financial (obligations)
                    1.6.4.4.5.2 (Obligations) that he can handle
                1.6.4.4.6 The other will not (continue to assume only the obligations)
                    1.6.4.4.6.1 Financial (obligations)
                    1.6.4.4.6.2 (Obligations) that he can handle

Figure 3. The Dependency Structure for the Excerpt from Underwood. Implicit terms and propositions needed to complete the structure are given in capital letters.

70

Structural Style

Style refers to many aspects of text but is limited here to deal only with dependency structures.[5] Even so, the varieties of these must be almost as numerous as texts themselves, but there are certain characteristics of structural style that are particularly important. Among these is the variation in width and depth of a discourse. Wide structures characteristically result from the recitation of a lot of irrelevant facts about some single topic, such as municipal statistics or an entry in Who's Who. Deep structures occur when some single branch is pursued in very great detail. Newspaper stories tend to be like this.

Another characteristic of particular importance to scientific text is the location of the critical information. While any classification of text results in an oversimplification, to a considerable degree it is easy to classify scientific texts as one of two kinds. Some local text structures are deductively structured while others are inductively structured. In a deductive structure, a general principle is stated, usually at or near the outset, and the rest of the text is an explanation and exemplification of that principle. In an inductive structure in its pure form, examples bring the readers to some general principle, which may be deeply buried in the dependency structure. In the selections used in this research project, the passage from the college text was mainly deductive, while the tenth grade passage was mainly inductively structured. The seventh grade passage was largely descriptive, but to the extent that general principles were arrived at, they were inductive.

A number of sections of textbooks in various fields were examined. It was found that inductive structures predominate in lower level and introductory texts. This is probably because authors feel they must introduce some abstraction or some new idea to their readers by reference to the familiar and the concrete. There is a danger here, because it is much harder to find the general principle when it is not clearly marked, as it sometimes is not in inductively structured text. In some text, the general principle may be left unstated, and the reader must arrive at it by inferential reasoning. Such a Socratic approach may be a strong teaching device when a living teacher is present to correct misconceptions. However, it may leave many a less-skilled reader ignorant of the major information he or she is to derive from the text.

At its worst, the inductive style is sometimes accompanied by irrelevant details and a syntactic style that is positively misleading to those who have not mastered translating the surface structure of text into general principles. The passage from the tenth grade text begins:

---

[5]In connection with a project on children's narratives, Carole Menig-Peterson and Allyssa McCabe have developed a working system for classifying stylistic differences in text structure. It will be published in a forthcoming book.

"Every day of the week, and especially on holidays, crowds of curious people throng to zoos." The student who is still struggling to master stylistic variations in expository text may believe that the rule that says the most important idea comes first always applies. Such a student will be misled by this bit of irrelevancy, even if it is intended to make the text more interesting. If he or she has been taught by an English teacher to use certain grammatical features as clues to text structure, the student may imagine the clause following "especially" to be very important. Such mistakes may seem ludicrous to skilled readers, but they do occur in those who are trying to master expository style.

All of the burden should not be on textbook writers, though there certainly are differences in the quality of writing. Students must learn to cope with all kinds of styles, and at a higher level, they must deal with extremely dense and abstract ways of putting things, ways that make strong demands on inferential capability and general knowledge of the subject under discussion. The selection studied in detail is perhaps a model of what not to say at a level in which readers are still struggling with making sense of text structures.

No firm prescriptions can be made at the present time about instruction in reading in the higher grades. It is possible that instructional adaptation of some of the methods that have been used in this and similar research to determine how readers perceive text would lend themselves to instructional methods. In any event, severe deficiencies have been uncovered in how text is sometimes structured. The difficulties that may arise when these deficiencies interact with students having poor techniques of text analysis have been explored briefly.

Contemporary Textbooks

A final word about some other characteristics emerging in contemporary textbooks, particularly those at the high school level and introductory college level is required. Two points need to be made. One concerns writing style, and the other concerns the whole concept of text.

Some years ago, there began an effort to discover the correlates of text that is easy to read and to understand. This work began with E. L. Thorndike, but it received broader fame if not notoriety upon the publication of Rudolph Flesch's book, The Art of Plain Talk. Flesch's book had a salutary effect, for it made us aware of unnecessary wordiness and jargon in scientific writing. There is a danger associated with the whole "readability index" movement. Readability indices provide, at best, correlates of good style. When authors begin to write to the indices, unexpected side effects occur. The words become shorter, the sentences simpler, and the human interest greater. This can happen at the expense of coherence of the text. No text could be made of shorter sentences than text composed of sentences modeled after base phrase markers. But such a text would be unreadable.

The second point is more general. A trend has emerged over the past 25 years. Books are attempting to become something they are not. Books cannot compete with other media, except on their own terms. A book with text printed in two colors, with boxes, inserts, cartoons, and the like is a poor imitation of self-paced instructional devices or the limited attention-span demands of television. In recent years, the author has taken an interest in textbooks and has examined a number of them in areas ranging from biology to economics, history, political science, and psychology. Some of the better ones use headings, tables, and graphs to assist the reader in sensible ways, but others appear to be designed more to produce a vivid product that sells rather than to help the student come away with a coherent body of knowledge.

But texts in the social sciences, particularly the softer sides of psychology, are among the worst. In one of the best introductory textbooks in psychology, there is a picture of a man playing tennis. There is no caption. Apparently, it serves only to break up what would otherwise have been a nearly unique half-page of uninterrupted print. In another well-known introductory text, there is a picture of a man and a woman talking. The caption says, "This dialogue illustrates the fact that the basic function of language is communication." Another picture shows a young woman playing a flute. The caption for this one is worse, "The cerebellum controls the intricate muscular movements of the musician." Aside from the fact that both of these captions are of dubious accuracy, the major effect of them is to interfere with the development of a serious, intellectually strong grasp of a body of knowledge.

Readers who are going to be the scientists, lawyers, and professional persons of all kinds in the future must learn to understand dense prose, prose in which what modifies what is hard to discover, and what needs to be inferred is not easy to determine. For someone who has not been prepared for this intellectual exercise, it is an impossible task.

References

Bransford, J. D. Human cognition: Learning, understanding, and remembering. Belmont, Cal.: Wadsworth, 1976.

Brown, A. L., and Smiley, S. S. Rating the importance of structural units of prose passages: A problem of metacognitive development. Child Development, 1978, 48, 1-8.

Cather, W. A lost lady. New York: Knopf, 1923.

Chomsky, N. Syntactic structures. The Hague: Mouton 1957.

Flesch, R. F. The art of plain talk. New York: Harper and Brothers, 1946.

Fredricksen, C. H. Representing logical and semantic structures of knowledge acquired from discourse. Cognitive Psychology, 1975, 1, 371-458.

Hirsch, E. D. The philosophy of composition. Chicago: University of Chicago Press, 1977.

Johnson, R. E. Recall of prose as a function of the structural importance of the linguistic units. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 12-20.

Karmiohl, M. Text-based vs. reader-based methods for the determination of conceptual structure and the prediction of the process and the quality of reading comprehension. Master's thesis, University of Virginia, 1979.

Kintsch, W. and van Dijk, T. A. Recalling and summarizing stories. Unpublished manuscript, University of Colorado, 1976.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge and S. J. Samuels (eds.), Basic processes in reading: Perception and comprehension. Hillsdale, N. J.: Erlbaum, 1977.

Steinbeck, J. The grapes of wrath. New York: Viking Press, 1958.

Chapter 5

INTRODUCTION TO THE SYMPOSIUM

To find out how people know science, what they understand about science, investigators predominantly have administered paper-and-pencil tests to groups of subjects. This method continues to prevail at the practical level of the classroom. Teachers judge what students in their classes "know" by means of normative, competency-based, mastery, or some other type of testing.

When achievement is the variable, much science education research also depends on paper-and-pencil instruments. Experimental and control groups are tested, and the results are subjected to various statistical analyses in search of statistically significant differences. Questions about whether the students "really understand" an idea, a problem, a principle, or a concept may be discussed by teachers and researchers, but such questions have seemed intractable to researchers. The concept "understanding" has been considered too vague and ambiguous for research purposes. Yet, the commonsense question of "real understanding" continues to haunt those concerned with science teaching.

The concept of understanding has been better defined by the work of Piaget, and, more recently, by many investigations in a new field of study termed cognitive science. Piaget demonstrated the power of using concrete physical experiments as stimuli for getting children and adolescents to think. The task is not as easy as it may appear on the surface, for the questioning procedures used and the interaction of physical objects, subject, and researcher must be planned and executed carefully to obtain data that can be interpreted. One critical contribution of Piaget's work that relates to the subject of this symposium is his insistence that the investigator interact directly with a child or adolescent if the object is to learn how thinking occurs. Paper-and-pencil approaches are too limited for answering questions about how knowledge grows and develops in humans. Interacting with individuals was a revolutionary approach to investigating science understanding. Piaget's work was ignored in the United States for many years, but its value has been recognized, and thoughtful investigators have developed their own skills in interview methodology to the point where they have confidence in the data they collect and the interpretations they make.

Directly interacting with individuals as they explain phenomena or solve problems has enabled investigators to ask more specific questions about human thought processes. "Think aloud" experiments have demonstrated the value of qualitative research that is grounded in theory.

The theoretical framework for the research reported in Chapters 6, 7, and 8 had its origins in work in computer sciences, artificial intelligence, and cognitive development. These chapters illustrate the importance of gathering data directly, rather than indirectly, from people as a way of developing new insights into what it means to understand a particular subject matter. They also demonstrate the usefulness of asking questions that involve the subjects in content problems. Another aspect of the approaches presented here is their reliance on a few experts and novices to find answers to questions, rather than on random samples from large groups of subjects.

Currently, most researchers using the procedures presented here have selected mathematics and physics as the subject matter for their studies. Hopefully, future research will include investigations of learning and problem-solving difficulties in other areas of the natural sciences. Such research should add to the methodology found useful in the studies described at the symposium.

Each of the three investigators who presented papers at this symposium used a different approach. Their work has commonality, however, in being theory-based, in combining subject matter knowledge with psychological expertise, and in their concern with rigor in the use of qualitative methods.

Chapter 6 describes investigations into the understandings of freshman physics students with regard to selected aspects of cause and effect. In this chapter, Jack Lochhead discusses student confusions about concepts of cause-and-effect relations, especially those involving rate of change. The structure of student concepts is investigated through methods that include interviews, paper-and-pencil essay questions, and carefully designed examinations.

Jack Lochhead is Director of the Cognitive Development Project, Department of Physics and Astronomy, University of Massachusetts at Amherst. Since the early 1970s, he has been investigating problem solving in mathematics and physics and the cognitive development of students. Professor Lochhead received a B. S. and an M. S. in Physics before completing a doctoral degree in Educational Research and Statistics at the University of Massachusetts. Among his numerous publications are <u>Cognitive Process Instruction</u> and <u>Developing Mathematical Skills: Computation, Problem Solving, and Basics for Algebra</u> (in press).

Paul Johnson and his colleagues at the University of Minnesota have used the "think aloud" experiment to study "heuristic" problem solving. Heuristic problem solving can be studied only when problem solvers are confronted with problems requiring an untaught or unlearned combination of physics principles for their solution. This research involved individuals knowledgeable about physics teaching and learning and the problems presented in text, lectures, assignments, and examinations at particular universities, as well as experts in conducting psychological research using "think aloud" procedures and methods of analysis. The

analysis itself requires collaboration of those knowledgeable in physics problems and solutions with those knowledgeable in analyzing "think aloud" data.

Johnson and his colleagues found "holes" in physics instruction, as well as two interesting kinds of problem failure--"garden paths" and "missing bridges." Their findings will stimulate new questions and new directions for research in science education. The paper appearing as Chapter 7 in this section was presented at the conference by Johnson.

Paul E. Johnson is Professor of Educational Psychology at the University of Minnesota and is a faculty member of the University's Center for Research in Human Learning. He has advanced degrees from Johns Hopkins University and has written many professional papers and technical reports for various professional journals. Psychology of School Learning; Learning: Theory and Practice; and Verbal Training Research and the Technology of Written Instruction are among the books he has written. Professor Johnson's interests include human learning and cognition, with particular emphasis on the study of human expertise in complex problem-solving environments and decision-making processes in technical fields and professional practice.

The final paper (Chapter 8) in the series on investigating science understandings uses information-processing models from psychology to explore problem solving in physics. Jill Larkin has been concerned with such questions as, "What makes someone an expert in physics?" and "What skills and cognitive processes do experts exhibit in arriving at problem solutions that novices do not?" Her investigations include "think aloud" experiments, but her analysis involves concepts of production systems and condition-action units. Verification of the utility of the analytic schema of production systems and condition-action units is accomplished by constructing and testing computer programs that can solve the physics problems solved by experts and novices. Her work is providing information about the knowledge and processes essential for expertise in physics problem solving. It also shows the usefulness of psychological models in understanding learning and problem solving in complex domains.

Jill Larkin is Assistant Professor in the Psychology Department at Carnegie-Mellon University. She earned a Bachelor's degree in Mathematics from Harvard University and an M. A. in physics and a Ph. D. in Science and Mathematics from the University of California, Berkeley. She taught high school mathematics in both the United States and Ethiopia. After serving as an Assistant Research Physicist and Lecturer in the Physics Department and the Group in Science and Mathematics Education at Berkeley, she was invited to join Herbert Simon and his associates in cognitive science research at Carnegie-Mellon University.

Professor Larkin is coauthor of two volumes of Principles of Physics for the Physical and Biological Sciences, as well as a number of papers in professional journals. She has also contributed papers to several volumes reporting research on cognitive processes. Her paper discusses a sequence of work addressing the role of understanding in solving physics problems. Differences in the problem-solving approaches

of experts and novices in science have been identified. These differences are beginning to indicate what knowledge is essential for expertise in physics problem solving.

The three chapters in this section represent new directions in science and mathematics education research. They illustrate the unique contributions that the careful use of nontraditional research strategies can contribute to the complex task of unraveling the threads of human thought about technical subject matter.

Chapter 6

THE COMPOUNDING OF CAUSE AND EFFECT, CHANGE AND QUANTITY

Jack Lochhead, University of Massachusetts

Introduction

One of the most basic and essential of all scientific concepts, namely, rate of change, will be discussed in this paper. Although a relatively simple idea, it is often misunderstood by students and, worse, frequently mishandled by teachers and textbooks. It is not the intention of this paper to answer the question of how this concept should be taught; rather, it is to draw attention to the confusion it causes students, to summarize some of what is now known about that confusion, and to suggest ways in which teachers can go about investigating the problem themselves.

The importance of rate is that it is a particularly simple way of describing a type of relationship between two variables. Thus, to understand students' understanding of rate one must first consider their knowledge of relationships. R. L. Gray (1975) and Piaget (1952) have pointed out that a relationship implies an invariant. In many relations the invariant is a directly observable, concrete quantity. For example, consider the relationship between the various members of the pine family and the number of needles in each cluster. The Parry Pinyon is unusual because it has four needles per cluster. Thus, an invariant characteristic for this species is the number 4. It can be directly determined by counting the needles in a cluster. This type of relationship is quite easy to grasp, and children, from an early age, are quite good at learning such relations. A large part of our scientific knowledge consists of these relations. Unfortunately, knowledge of this kind is highly specific, rarely generalizable, and of minimal use in making predictions.

Another type of invariant is associated with functional relationships in which a variable quantity, such as the height of a tree, is related to a second variable quantity, perhaps time. In these cases, the invariant may not be directly observable, and it may be impossible to conceive of it as a concrete quantity. If the height (<u>h</u>) is related to the age (<u>a</u>) by a simple linear equation, $h = ka$, then the invariant <u>k</u> is the ratio of height to age. A growth of one foot per year certainly cannot be directly observed, nor is it possible to

point to it in the same sense that we were able to point to the four needles in each cluster of the Parry Pinyon. For this reason, these relations are abstract and beyond the grasp of immature thinkers.

The above discussion can be summarized using Piaget's distinction between concrete and formal operations. The invariants in concrete relationships can be represented as tangible objects. Those in formal relationships cannot. Considerable mental activity is required to reason about formal relations. Studies by Karplus (1979), Renner (1972), Lawson (1974), and Arons (1976) have shown that most high school students lack the conceptual building blocks to think clearly about abstractions such as rate of change. A major challenge, therefore, for high school and college science is to help students develop the reasoning necessary for conceptualizing such formal relationships.

Misconceptions About Change

How could one go about investigating the difficulties associated with learning a concept such as rate of change? The first step is to place the problem in some theoretical framework. This has already been done by referring to the Piagetian perspective. Through it, one is able to understand that rates, as formal concepts,[1] require certain prerequisite skills. Piaget and Inhelder (1958) describe these prerequisites. Arons (1976), Fuller (1977), Gray (1979), Lawson (1976), and Renner (1972) consider methods for teaching students to become formal thinkers. In this paper, it will be assumed that the reader is at least partially familiar with the above references, since the task of summarizing them goes well beyond the scope of the paper.

With a theory in place, the next step is to collect data concerning student performance. Initially, this step is likely to be anecdotal. The investigator's interest in the subject sprang from casual observations made while teaching. It was found that many calculus students often could not solve problems such as:

> Two pipes can be used to fill a tank with water. When both pipes are turned on, the tank fills in 9 minutes. When only pipe A is used, the tank is filled in 15 minutes. How long will it take to fill the tank if only pipe B is used?

Two of the more common incorrect responses are shown below. Notice that neither makes a clear and unambiguous use of the concept of rate.

Response Method 1. Both take 9 minutes so, if they were the same size, one would take 18 minutes. A pipe that takes 18 minutes can fill

[1]Some rates are not really formal. For example, we may want to calculate the number of cakes in eight packages where cakes come at the rate of two per package.

half the tank in 18/2 = 9 minutes, so 2 pipes can fill it in 9 minutes. But the problem says pipe A fills the tank in 15 minutes. 15 is 3 less than 18 so, to balance pipe A, pipe B should take 18 + 3 = 21 minutes.

Response Method 2. This method starts like Method 1 but considers the ratio of 15 to 18, not the difference. To get 15 to equal 18, we need to multiply by 18/15. (18/15 x 15 = 18.) Therefore, if pipe B takes $\underline{T}$ minutes, then 15/18$\underline{T}$ should give 18. This is so the times will be symmetrical around the average time of 18 minutes. If 15/18T = 18, then $\underline{T}$ = 21.6 minutes.

In order to investigate the source of such confusions the investigator began to write homework and exam questions that would test students' ability to think about rates. Examining incorrect student responses allows one to see how they think about the problem. One common confusion is between the amount of a quantity and its rate of change. Thus, we get several types of answers.

The question: A car travels the 90 miles from Amherst to Boston at 30 miles per hour. Draw a speed versus time graph that represents the trip. Students often respond with a graph of distance versus time (Figure 1).

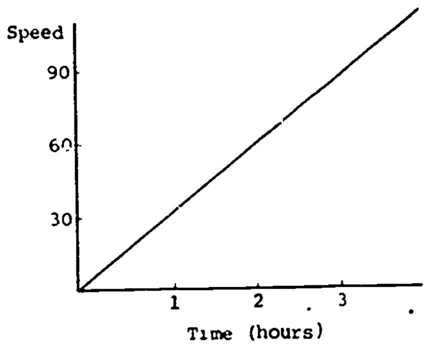


Figure 1. Graph of Speed versus Time (actually distance versus time) Produced by Some Students

The question: Sketch speed versus time and acceleration versus time graphs for a cart that is shot by a rubber band across a table top but rolls to a stop due to friction. Students often respond with identical graphs (Figure 2).

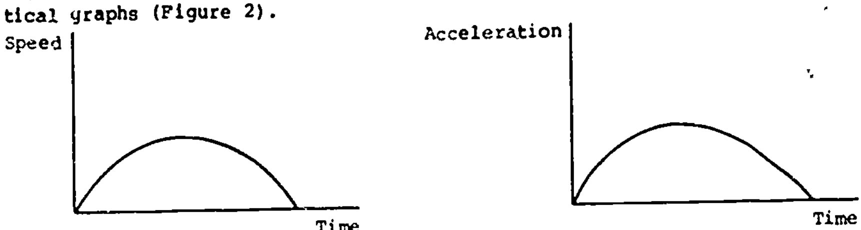


Figure 2. Graphs of Speed and Acceleration versus Time Produced by Some Students

Similarly, rate of change can be confused with quantities other than the one whose change it measures.

The question: Draw the shape of a graph of speed versus distance that would describe the following two-hour bicycle trip. You start off along the level, then you come to a long, steep, uphill section and finally to a downhill section that is twice as long as the uphill part. Many students respond with a graph of altitude versus distance (Figure 3).

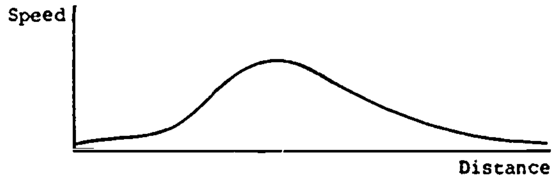


Figure 3. Graph of Speed versus Distance (actually altitude versus distance) Produced by Some Students

Of course, the students' graphs by themselves do not form convincing evidence for their failure to distinguish between a quantity and its rate of change. Extensive interviews with students are needed in order to distinguish problems of graphing from those stemming from a conceptual confusion about rates. In the subject area of velocity and acceleration, extensive interview studies have been conducted by Clement (1979), Champagne and Klopfer (1979), and Trowbridge and McDermott (1980a,b). Sample transcripts help to illustrate the students' confusion. In the Trowbridge and McDermott experiments, two balls were rolled down parallel tracks such that one ball passed the other. Students were then asked if the two balls ever had the same speed. Some typical student responses were:

S1: Somewhere around in here (indicates region near first passing point) they must be going about the same speed, because ball B passes ball A. So while ball B is speeding up, ball A is slowing down. There's got to be a point where they're going about the same speed.

S2: Well, it's hard to say. It seems like they would...Yeah, they are, because like when I'm driving on the freeway...You know how you hate to have someone directly on the side of you; he might have been behind me, that he's caught up, so he must be going the same speed, even though he keeps on passing me. So at the time when we're together, we're probably going the same speed.

In another task, two balls were rolled down parallel tracks such that they never passed each other. However, one ball accelerated while the other slowed down; thus their velocity versus time graphs crossed. A typical investigator-student dialog was as follows:

I: Let's see whether these two balls ever have the same speed.

S: No.

I: How could you be sure that they didn't have the same speed?

S: Because they never met; they were never lined up to each other.

Similar experiments were carried out on acceleration. A typical investigator-student dialog was as follows:

S: Wouldn't they have the same acceleration at the point they have the same velocity?

I: Why do you think that would be true?

S: Because your acceleration is that <u>delta V</u> over <u>delta t</u>. And at the point where you have the same velocity, you have the same <u>delta t</u> and the same <u>delta V</u>.

Interview responses such as these are strong evidence that some students confuse position and velocity or velocity and acceleration. The full interviews from which these responses were excerpted are even more convincing, since they allow the interviewer to probe the student's understanding from a variety of perspectives.

Clinical interviews can provide detailed insights into student thinking. However, except in the case of Trowbridge and McDermott (1980) who interviewed over 300 subjects, clinical interviews are usually conducted with only about a dozen subjects. To determine whether the misconceptions revealed in such interviews are widespread, some form of group testing is needed. This can be done in a variety of ways. The investigator recently gave his calculus class the following quiz question:

> During the past 6 months the rate of inflation has dropped from 18% to 10%. Explain what this change means in terms of the value of a dollar.

Out of 26 students, eight correctly stated that the value would continue to decrease but at a slower rate. Seventeen said the value of the dollar would increase. One seemed to think both might happen:

> "In terms of the dollar, if inflation has dropped, the value of the dollar has increased, or at least its decrease in value has slowed down."

Other group test studies have been conducted with populations as large as several hundred students (Clement, 1980; Trowbridge and McDermott, 1980a,b). These have shown that the amount/rate confusion abounds among college-level, science-oriented students. Questions dealing with the more abstract forms of rate, such as acceleration, yield error rates in the range of 80% to 90%. The effect of two semesters of calculus-based physics on this type of confusion is negligible. Thus, the confusion is common, and it is also extremely difficult to overcome.

It is also possible to collect data on misconceptions from sources other than students. A recent American Motors advertising campaign seemed to exploit this amount/rate confusion when they claimed that their car had a larger gas tank than that of their competitors. It was thus able to go a greater distance between fill-ups. While no mention was made of the miles-per-gallon rating, one wonders how many customers felt that this figure would also be higher.

A less humorous case is shown in Figure 4. This graph is taken from a biology pamphlet on the circulatory system. While the writer probably knew what he was trying to say, he nonetheless labeled the vertical axis "changes in volume" rather than "volume." This kind of error is particularly serious, since it strengthens the students' tendency to confound amount and its rate of change. The record shows the volume changes in cubic milliliters in the left ventricle. Note how the ventricular volume decreases during ventricular systolic ejection into the aorta.

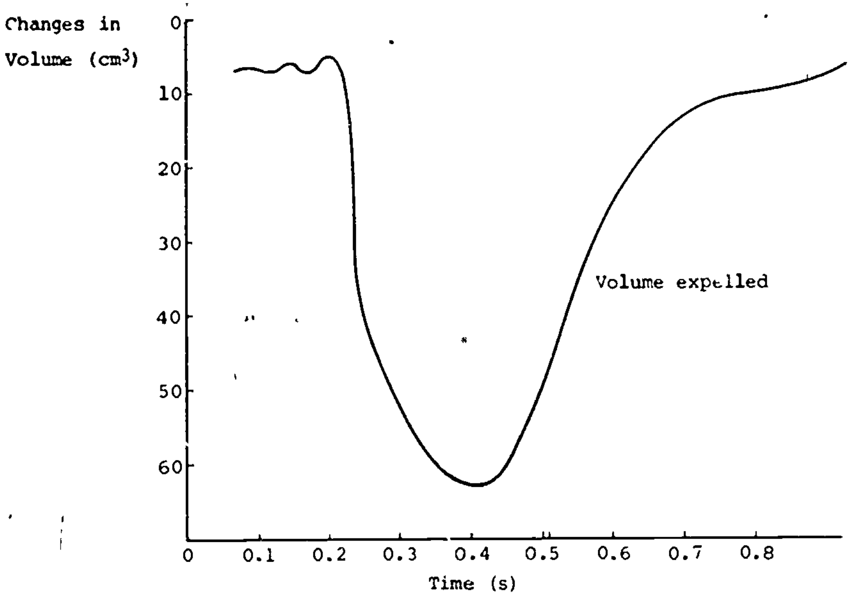


Figure 4. Example of Mislabeling in a Textbook that Results in Confusion Between Amount and Rate of Change

Cause and Effect as Change

In some student responses there seems to be a slightly different type of confusion associated with certain cause-and-effect relationships. For young children, cause and effect are often inseparable from association. Thus, they will say "Christmas causes snow," or "the trees swaying is what makes the wind blow." This view of association is carried into adult life in such areas as interpreting the meaning of a statistical correlation between two variables or determining the political beliefs of a person who has been known to associate with communists or conservatives.

In contrast to this primitive view of cause and effect is the statement that the presence of a change agent causes change in some quantity. While the associative view is essentially static, i.e., there is no need to imagine a process evolving in time, the change agent concept is dynamic. This is often called an <u>operative</u> perspective since it involves a change <u>operator</u> that <u>transforms</u> an initial state into some final state. Thus, cause and change (rate of change) are linked. In physics, for example, force is the change agent of velocity. When a force acts on a body it causes that body's velocity to change. For the student who is neither clear on the meaning of rate nor fully free from the associative concept of cause, the relationship between force and velocity is exceptionally confusing.

The following are student answers to the quiz question, "What is the cause-and-effect relationship between force and velocity?"

"As one increases the other increases, and as one decreases the other decreases."

"A change in force causes a change in velocity."

"Force causes velocity to occur. So if force increases, so does velocity, be it negative or positive."

"The stronger the force on an object, usually the velocity is greater."

"When the force increases, velocity increases and when force decreases, velocity decreases. In other words, force is the cause and velocity the effect."

"If the force is constant, the velocity will remain constant. If the force is increasing, velocity will increase. If the force is decreasing, the velocity will decrease."

"Force is the cause which gives the object its velocity."

"Force causes velocity, but when force stops, the effect may go two different ways: it may either speed up or slow down."

These answers were given by introductory physics students directly after a lecture on force in which the "force causes velocity" misconception was addressed directly! Students had been explicitly told that the correct equation was $\underline{F} = \underline{m}\ \underline{dv/dt}$ and that $\underline{F} = \underline{kv}$ was not true. These equations were explained with examples and informal verbal descriptions as well as the standard formal explanations.

Other studies of student conceptions concerning the relation of force and velocity show that misconceptions continue well into the science student's undergraduate career. For example, Clement (1980) reports that among junior level engineering students, 70% demonstrated the "force causes velocity" confusion.

Among students who had taken a special course designed to attack such misconceptions specifically, the error rate dropped to 55%. The task of teaching students to avoid such confusions is formidable indeed! Yet, as suggested in the next section, there may be no adequate substitute for a thorough grasp of cause and effect-rate relationships if even the simplest rate of change examples are to be understood.

Some Further Confusion

So far, it has been shown that a great many apparently different areas of student confusion can be viewed as specific examples of the more general amount/rate confusion. The confusion may even extend to many cause and effect relations. This raises the question of whether the amount/rate confusion may extend to other situations not normally viewed as being an amount and its rate of change, but which are in fact so related.

Arnold Arons (1978) suggests that over one-third of the college sophomores in the life sciences cannot distinguish surface area from volume. Could this confusion have roots similar to those of the confusion between acceleration and velocity? While surface area is not normally conceptualized as the rate of change of volume, it can be viewed that way and is so when we calculate volumes through integration. Mass and density are another pair of frequently confounded variables. Could students fail to distinguish density from mass because density describes the rate at which mass increases with respect to an increase in volume? At this point, most of the solid evidence on the amount/rate confusion has been described. We know the confusion exists. In certain specific areas, such as velocity and acceleration, it has been studied exhaustive detail. Amount/rate confusion is related to the general problem of formal thinking and a few partially effective, but frustratingly incomplete methods for helping students overcome it have been developed. What is lacking is a clear theoretical understanding of why it is so incredibly difficult to teach students to distinguish between rate and amount.

The following is some speculation on what could be the beginning of a new insight into this problem. It begins with the apparently

unrelated problem of students' ability to translate between English text and algebraic statements. For several years students' responses to a rather simple problem have been investigated:

> Write an equation using the variables $\underline{S}$ and $\underline{P}$ to represent the following statement: There are six times as many students as professors at this University. Use S for the number of students.

Among technical students, nearly 40% fail this question; among nontechnical students, the rate may be as high as 90%. Based on extensive clinical interviews and many different written tests, it has been concluded that, while there are several different sources of error in the Students and Professors problem, a primary source has to do with students' tendency to view things statically (Clement, Lochhead, and Monk, 1981; Clement, Lochhead, and Soloway, 1980).

In the correct equation, $\underline{S} = 6\underline{P}$, the number 6 must be seen as an operator that increases the number $\underline{P}$ so that it becomes equal to a larger number, $\underline{S}$. The Piagetian literature is full of examples of the difficulty children have conceptualizing such dynamic transformations. But the real source of the error among college students may be much simpler.

The printed page is simply not a good medium for representing dynamic processes. Research at the University of Massachusetts has been able to show that students perform far better on the Students and Professors type problems when they are phrased in the context of computer programming. Programming languages, while constrained to the frozen frame of the printed page, nevertheless convey more of a sense of operation and transformation than do algebraic expressions.

An even greater sense of process and change is possible if one goes beyond the limitations of paper to the dynamic presentations possible with computer graphics. Through this medium it may be possible to faithfully represent and describe the images one uses when thinking of rates. Imagine, for example, the process by which one builds a sphere through the integration of surface area shells:[2]

$$4\pi \underline{r}^2 \underline{dr} = 4/3\pi \underline{R}^3$$

No collection of static textbook illustrations can capture the concept, whereas a relatively simple computer graphic display could.

In a different domain, consider the equation: $\underline{F} = \underline{ma}$. Here the static presentation of the printed page is further reinforced by disguising the rate of change of velocity under the letter $\underline{a}$. DiSessa (1980) recommends an attempt to capture the dynamics of this equation by

---

[2]Note that the information presented on this page gives no hint to what sort of dynamic process is being discussed. The reader must bring that knowledge to the text. For a more effective method of conveying that process, see page 223 of Fleming and Kaput in Appendix I.

referring to force as a source of momentum flow. This suggests computer graphic displays for physics in which streams of momentum could cascade from each point of force application.

This is not to suggest that an understanding of rate is only possible via computers. Rather, it is suggested that some of the confusion students feel may stem from the limitations of formal representations. By becoming more aware of these limitations, teachers may be able to find methods for circumventing them. It is also important to remember that the dynamic image hypotnesis outlined previously is pure speculation. Further investigations in the classroom and in the clinical laboratory are needed to test its validity. It is hoped that some readers may help in that venture.

Conclusion

Some research on the tendency of students to confuse amount and rate of change has been reviewed. It has been suggested that, in some circumstances, this confusion may be connected to confusions between cause and effect. These confusions are often based on the more primitive notion of association and reflect a failure to conceptualize the process elements of a cause-and-effect relationship. Finally, it has been speculated that the underlying problem may be the students' failure to apply dynamic imagery.

Except in the area of velocity and acceleration the amount/rate confusion has received relatively little study. Teachers can add to the meager data pool by designing quiz and test questions probing various aspects of this difficulty.

The deepest insights are likely to come from "teaching experiments" in which course material based on a theoretical perspective, e.g., the dynamic imagery hypothesis, is shown to be either more or less effective than previous approaches. The failure of current instruction to have a serious impact on student learning in this area suggests that significant improvements should be possible.

Acknowledgement: Robert L. Gray and Seymour Papert have had such a pervasive influence on my own thinking in this area that I can no longer distinguish their ideas from mine. If this paper contains any useful insights, it is largely due to them.

References

Arons, A. Cultivating the capacity for formal reasoning: Objectives and procedures in an introductory physical science course. American Journal of Physics, 1976, 44, 9.

Arons, A. Personal communication, 1978.

Champagne, A. B. and Klopfer, L. E. Factors influencing learning of classical mechanics. Technical Report, LRDC, University of Pittsburgh, 1979.

Clement, J. Student preconceptions in introductory mechanics. Invited paper, national meeting of the American Association of Physics Teachers, Chicago, January, 1979.

Clement, J., Lochhead, J., and Monk, G. Translation difficulties in learning mathematics. To appear in the American Mathematical Monthly, Spring, 1981.

Clement, J., Lochhead, J., and Soloway, E. Positive effects of computer programming on students' understanding of variables and equations. Invited paper, Association for Computing Machinery national conference, Nashville, October, 1980.

DiSessa, A. Momentum flow as an alternative perspective in elementary mechanics. American Journal of Physics, 1980, 48, 5.

Fleming, D. J., and Kaput, J. J. Calculus with analytical geometry. New York: Harper and Row, 1979.

Fuller, R. The ADAPT book. ADAPT Program, Lincoln: University of Nebraska, 1977.

Gray, R. L. Toward observing that which is not directly observable. In J. Lochhead and J. Clement (eds.), Cognitive process instruction. Philadelphia: Franklin Institute Press, 1979.

Inhelder, B., and Piaget, J. The growth of logical thinking from childhood to adolescence, New York: Basic Books, 1958.

Karplus, R. Proportional reasoning and control of variables in seven countries. In J. Lochhead and J. Clement (eds.), Cognitive process instruction. Philadelphia: Franklin Institute Press, 1979.

Lawson, A. E., and Renner, J. W. A quantitative analysis of responses to Piagetian tasks and its implications for education. Science Education, 1974, 58, 4.

Lawson, A. E., and Wollman, W. Encouraging the transition from concrete to formal cognitive function—an experiment. Journal of Research in Science Teaching, 1976, 13, 5.

Lochhead, J., and Gray, R. L. A heuristics laboratory for encouraging structural growth in college students. Lecture, Piaget Society, May, 1975.

Piaget, J. _The child's concept of number_. London: Routledge and Paul, 1952.

Renner, J. W., and Stafford, D. G. _Teaching science in the secondary school_. New York: Harper and Row, 1972.

Trowbridge, D. C., and McDermott, L. C. An investigation of student understanding of the concept of velocity in one dimension. _American Journal of Physics_, 1980, _48_, 12.

Trowbridge, D. C., and McDermott, L. C. An investigation of student understanding of the concept of acceleration in one dimension. _American Journal of Physics_, in press.

Appendix I

SECTION 6.6: VOLUMES OF SOLIDS OF REVOLUTION*

Objectives: Set up and evaluate definite integrals that describe the volumes of solids generated by revolving plane regions about an axis, using either the disk or the shell method, as appropriate.

The Disk Method

We can think of a solid sphere as being generated by revolving the region enclosed by a semicircle about its diameter. Similarly, a solid right circular cone results from revolving the region enclosed by a right triangle about one of its legs (see Figure 6.46). In this section we shall deal with more general solids generated by revolving a given plane region about a given line. Such a solid is known as a *solid of revolution*, and the line is called the *axis of revolution*. As in our treatment of area problems, we assume the region to be bounded by the graphs of functions.

Let us now determine the volume of the solid generated by revolving about the x axis the region below the graph of $y = f(x)$ from $x = a$ to $x = b$, where $f(x)$ is continuous and nonnegative on $[a,b]$; see Figure 6.47.

As we know, the area under the graph can be approximated using vertical rectangles of width Δx_i (see Figure 6.47) based upon a partition P of $[a,b]$. When each rectangle is revolved about the x axis, a disk is generated (see Figure 6.48). The volume of a disk of radius R and thickness h is $\pi R^2 h$, so the volume of a typical disk generated by such a rectangle is

$$\pi[f(x_i)]^2\Delta x_i$$

Since the sum of the rectangular areas approximates the area of the region being revolved, the sum of the volumes of the disks approximates what we shall define as the volume of the solid generated. Thus, we say the approximate volume of the solid is

$$\sum_{i=1}^{n} \pi[f(x_i)]^2\Delta x_i$$

But this is a Riemann sum, and if we take finer and finer partitions P of $[a,b]$, we get better approximations. Hence it makes sense to define the actual volume V of the solid of revolution to be the limit, as $\|P\| \to 0$, of such Riemann sums. But, by definition this limit is a definite integral.

Definition $$V = \lim_{\|P\|\to 0} \sum_{i=1}^{n} \pi[f(x_i)]^2\Delta x_i = \int_a^b \pi[f(x)]^2 dx$$

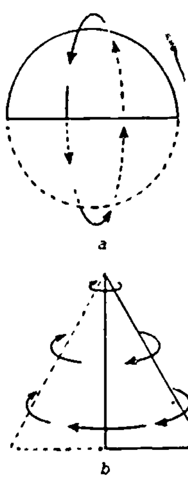


Figure 6.46

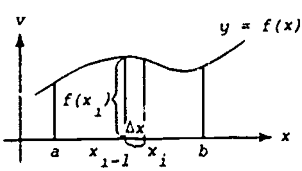


Figure 6.47

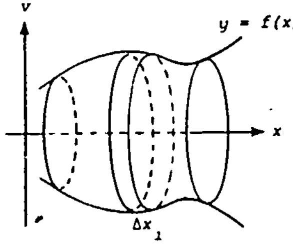


Figure 6.48

*Fleming, D. J. and J. J. Kaput, Calculus with Analytic Geometry, Harper & Row, New York, 1979.

91

Chapter 7

SCIENTIFIC REASONING: GARDEN PATHS AND BLIND ALLEYS[1,2]

Paul E. Johnson, Andrew Ahlgren, Joseph P. Blount, Noel J. Petit,
University of Minnesota

Introduction

This paper represents a brief progress report on the status of work done over the past year by a team of psychologists and physicists at the University of Minnesota. The focus of this work has been the investigation of scientific reasoning and, in particular, heuristic thinking. It is assumed that individuals prefer to use the least cognitively demanding process for solving problems (Neeches and Hayes, 1978; Hayes-Roth, 1978). For much problem solving by experts, the process consists of matching a given problem with a category for which solution methods are known (Chase and Simon, 1973; Hayes-Roth, 1978). For much problem solving by novices, the process consists of searching memory for a relevant formula that contains the desired unknown (Heller and Greeno, 1978; Simon and Simon, 1978). In neither of these cases does the individual engage in what we have termed heuristic thinking.

Typical analyses identify two phases of the problem-solving process: representing the problem situation, and applying physics principles to that problem representation in order to generate a solution procedure (Larkin, McDermott, Simon and Simon, 1980). Representation involves translating the problem statement into a "canonical form" that is accessible to physics principles. For many problems, the translation to the canonical form may entail identifying the problem as a standard prototype, e.g., harmonic oscillator, free fall at constant _g_, block

[1]The work reported here was sponsored by the Joint NIE/NSF Research Program on Cognitive Processes and the Structure of Knowledge in Science and Mathematics (SED79-13036).

[2]Research of the sort described here simply cannot be done by one investigator working alone. At Minnesota we have been fortunate to have the collaboration of able colleagues in physics, science education, and psychology. We would especially like to acknowledge the contribution to the work reported here of Scott Fricke, Clifford Malcom, Theodore Petroulas, Wolfgang Rothen, and James Werntz.

on a frictionless plane that has standard solution methods. More generally, translation consists of equating aspects of the problem situation with canonical objects--idealized conceptual entities, such as massless rods, frictionless surfaces, and point masses, whose behavior can be represented clearly by a few principles. The canonical representation is not tantamount to solution. The representation can be inappropriate or flatly wrong, and even a good representation may not be tractable, e.g., the three-body problem.

Once a representation is created, problem solving proceeds by the application of domain principles in order to generate a solution procedure (Greeno, 1979). In some instances, single principles, such as uniform linear motion, will be sufficient to derive a solution; while in other cases, combinations of principles, rotational as well as linear motion, will be required. It is the latter case, where multiple principles are required, that is of special interest to us.

In order to be applied efficiently either singly or in combinations, principles of physics need to be selected in such a way that they fit the assumptions and data of a given problem. When several principles are involved, the coordination of their application becomes important. Prototypical problems are those for which single principles suffice, or ones for which the coordination of principles is taught explicitly in introductory textbooks, e.g., the simultaneous use of momentum and kinetic energy conservation in an elastic two-body collision.

In more advanced texts, procedures are identified for handling additional complexities, and algorithms are derived for dealing with general cases, e.g., the Lagrangian and Hamiltonian equations in advanced mechanics. Problems that require a novel coordination of principles for their solution, even though all the requisite principles may be known, provide opportunities for the study of heuristic thinking.

In order to study properties of heuristic thought, it is necessary to have tasks that elicit this activity. If one were to use problems that are familiar to individuals, then only the approaches of formula searching or prototype matching would appear. Whenever a problem has attributes requiring an untaught or unlearned combination of principles for its solution, one has an opportunity to see heuristic problem solving in action.

Problem Clusters: A Methodology for the Study of Heuristic Thinking

Initially, the investigators had thought to do a more or less exhaustive sampling of the problem-solving literature in classical mechanics as a means of ensuring the representativeness of our problem-solving tasks. A case file was developed of problems based upon an analysis of standard textbooks (e.g., Halliday and Resnick, 1970) as well as graduate examinations in the field (masters examinations for the

Universities of Minnesota, Chicago, and Saskatchewan). The resulting set of about 500 problems was sorted into a taxonomy based on major physics principles employed and a systems analysis of the context of the problems.

Among the problems inspected were some problems that Cohen (1975) called "dragons." Initially, it appeared that the Cohen problems merely represented situations in which "tricks" must be discovered in order for a solution to be obtained. Upon reexamination of these problems, however, particularly the MILKO (Cohen, 1975) problem, the conclusion reached was that they represented areas of mechanics in which the physics principles necessary for solution had not been explicitly taught. This speculation about such a "hole" in physics pedagogy was verified by examination of several texts in physics and engineering mechanics.

In addition to representing a situation in which the combination of principles required for solution had not been explicitly taught, Cohen's MILKO problem also contained a definite "garden path" to an incorrect answer. When MILKO was tried with several professors of physics, each one quickly came up with the same incorrect answer. Each readily concluded that the problem situation was familiar and easily solvable by a single, well-known principle. The misconception was so strong that, especially with more expert individuals, the correct answer was resolutely rejected when it was explained. A gradual series of hints, in the form of related problems presented to the subject, was necessary to enable one finally to "see" the way in which her or his existing knowledge needed to be modified in order to solve the problem.

The discovery of the "hole" nature of the Cohen problems led to an attempt to generate such problems in other areas of physics and to generate for each hole a set of associated problems that would (a) establish the extent of the holes and (b) lead problem solvers to understanding the nature of the holes. Tentative hole-centered sets have been generated in analytical mechanics, electrodynamics, quantum mechanics, and thermodynamics. Each problem set begins with an apparently simple problem that shows a high failure rate. This is followed by more transparent problems that draw the subject's attention to aspects of the initial problem that are overlooked or misunderstood. It often happens that the successful solution of one of these follow-up problems sends the subject scurrying back to correct her or his solutions to the earlier problems. The series ends with a test of understanding by asking for a statement of a general principle involved in the class of problems. It is particularly interesting that experts often spontaneously express such understanding, sometimes rushing ahead to generate the whole set of problems themselves, unprompted.

Another type of hole problem does not involve a garden path, an easy route to an incorrect solution, but rather involves a chasm--there appears to be no way to relate different principles that are relevant to the solution. Whereas, for the "garden path" hole, one can proceed directly to a solution with inappropriate principles, the "chasm" or "missing bridge" hole leaves one unable to proceed at all. In the "missing bridge," as well as the "garden path" variety, the essence seems to be a hole in the subject's exposure to similar problems. The

"holeness" of a problem may, of course, be only in the head of the individual problem solver. Some holes appear to be in the way physics is taught; in the press to cover all ground, some bare spots must be left. A few holes may turn out to be actual holes in physics--as the anomalous heat capacities of gases were found to be at the end of the 19th century.

It is important to recognize that what is being focused on here is not a situation in which the principles necessary for solving a problem are unknown to the problem solver, but rather a situation in which all the necessary ingredients are present but they have not been combined before in the required way. The experienced problem solver copes with this situation by falling back upon tricks, caveats, and other heuristics, while the less expert individual often does not recognize any deficiency or does not know how to proceed. In some cases, such as MILKO, the problem may be such that experts as well as nonexperts are led to believe the problem can be solved with simple principles rather than with a combination of principles. In this class of problems, there is an especially interesting opportunity to study expert error, the process of recovery from error, and memory modification as a result of problem-solving feedback.

In addition to the problems themselves, a rationale has been developed for making inferences about the behavior of subjects who work on them. The raw data of the research consist of tape-recorded, thinking-aloud transcripts of problem-solving sessions in which the subjects attempt to solve problems while verbalizing their cognitive activities. These transcripts are coded for evidence of one or more lines of reasoning as described below, as well as the heuristics and strategies that reflect how a given line of thought may be implemented by a particular subject on a given occasion.

Lines of reasoning are defined as any meaningful or comprehensible (in terms of physics principles) sequence of problem-solving steps that lead from some starting point to a recognizable goal or end state. These lines of reasoning constitute idealizations that are imperfectly realized in the problem-solving behavior of a given subject. The investigators are interested in determining which line of reasoning is employed by each subject in doing a given task as well as how a given line of reasoning is executed by each subject. The analyzed data of the research consists of reconstructed records of problem-solving behavior based upon lines of reasoning applicable to particular tasks.

In order to make clear the methodology employed, problem clusters will be considered that were designed by the team of investigators during this past year. In each case, the rationale will be presented for solving problems from which possible lines of reasoning for each cluster are developed.

Task Materials: The Minnesota Clusters

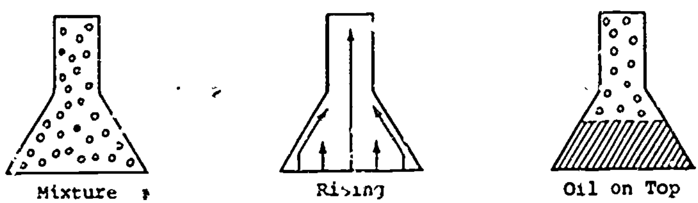
The problems used to explore scientific reasoning are grouped in clusters. The initial problem is the raison d'etre for the cluster. It reflects what is believed to be a "hole" in physics problem-solving experience that the subsequent problems are designed to help fill. Each cluster deals with an area of physics and each problem in a cluster helps lead the subject to explore the area further. The problems in a cluster become progressively more involved and lead to a full use of the principles in the area of physics under investigation. This section gives examples of each of the clusters developed thus far. The next section presents data from subjects attempting to solve the problems in one of the clusters.

Hydrostatics

The first cluster is based on Cohen's MILKO[3] problem. The cluster centers around the pressure change occurring as a two-liquid mixture separates in a milk bottle. The shape of the container is important to understanding the pressure change. This is because the pressure at the bottom of the column of fluid depends on the heights and densities of the liquids above. If the cross-sectional area varies with height, the ratio of heights of the separate states may be different from the ratio of volumes of the two fluids. In the case of a milk bottle shape, the pressure at the bottom decreases as the fluids separate. The major problems in the MILKO cluster are as follows:

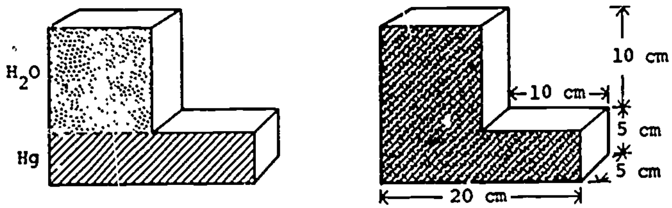
1. (MILKO) A milk bottle is allowed to stand so that the cream rises to the top; this occurs without any change in total volume. Does the pressure near the base of the bottle change? Why?

2. (ERLENO) An Erlenmeyer flask is filled with a mixture of corn oil and water. The mixture is allowed to stand so that the oil separates out on top; this occurs without any change in total volume. Does the pressure near the base of the flask change? Why?



[3]The initial problem in the MILKO cluster can be found in The Art of Snaring Dragons by H. A. Cohen, 1975. The remaining problems in this cluster are expansions of the original problem.

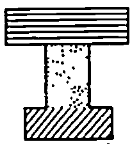
3. (CALCULO) Consider two identical L-shaped glass tanks containing 1,000 grams of water and 13,546 grams of mercury (see figure). Would the pressure be different at the bottom of the two tanks--and if so, by how much?



4. (REVERSO) You have solved a problem where two mixed fluids separate and pressure decreases. Now design a situation where the pressure will increase upon separation.

5. (BOWLO) Milk is poured into a bowl and allowed to stand so that the cream rises to the top; this occurs without any change in total volume. The pressure changes at the lowest point in the bowl. Tell whether the pressure increases or decreases and explain why.

6. (CANDLEO) Consider the physics of a candle flask (a novelty item sold in stores; see diagram). Three different oils float on top of each other. If the oils were thoroughly mixed and no change in volume occurred, how would the pressure at the bottom of the flask change?



The subject may pursue one of several lines of reasoning in solving the cluster of problems, but most revolve around four distinct ideas:

- Remembering the "Pascal Principle" that pressure (for a homogeneous liquid!) depends only on depth.
- Calculating pressure as force per unit area (P=F/A). Here the force is calculated from the weight of the material in the container.
- Expressing pressure as a function of density and height. Here the pressures in the mixed and separated states are calculated and compared.
- Calculating pressure from the proportions of liquid in the vertical central column above the bottom of the container, then comparing the pressures in the mixed and separated states.

Whether or not the effect of container shape is considered, the above principles may or may not lead to a correct answer. To get the correct solution, one must: (a) consider pressure as a function of density and height and, by using either inequalities or actual values for densities and relative volumes, determine the difference in pressure between the two states; or (b) realize that, for certain shapes, the surface areas exert a downward force on the fluid and that downward force will change as the relative distribution of liquids (and, thus, fluid density) changes, or (c) consider that the sides force a horizontal concentration of one of the fluids during separation, so that the relative proportions of liquids over the bottom of the central free-surface column changes.

A correct solution might follow logic similar to this: Consider the column of liquid in the center of a bottle with inward sloping sides. In the mixed state, the pressure at the bottom of the column is: $P = \rho gh$, where ρ is the average density of the liquid, g the acceleration of gravity, and h the height of the liquid. As the mixture separates, the less dense material rises. Since the top of the container is narrower than the bottom, when the fluids are fully separated the height of light material in the central column is greater relative to its volume than the height of the heavier material. Thus, the average density of the liquid in the column decreases. Since g and h remain constant, P decreases as the liquids separate.

In the first problem, one must use the effect of the container shape, which causes the cross-sectional area to vary with height, to solve the original problem. A subsequent problem in the cluster asks for a shape that would cause the opposite pressure change. If the effect of shape is understood, the shape needed to increase the pressure is easily given. A final problem in the cluster asks for a generalization of the effect of shape if more than two separating liquids are used in a complex shape. This problem further explores the subject's understanding of the line of reasoning he or she has used. There is no easy generalization, and here again specific shapes, densities, and volumes must be known to complete the problem.

As near as can be determined, the MILKO problem does not appear in college-level physics texts. The emphasis in hydrostatics in the typical text is contrary to this problem cluster in that homogeneous fluids are the only cases shown. For example, Halliday and Resnick (1970, p. 293), authors of a popular college-level physics text, emphasize Pascal's Vases, pointing out the irrelevance of shape to the pressure in a container. Sears and Zemansky (1970, p. 178) imply that they are discussing a homogeneous fluid, but emphatically state the pressure depends only on the weight of the liquid above the region of interest, and "any vessel, regardless of its shape, may be treated the same way." Ironically, Sears and Zemansky (1970) call this the "hydrostatic paradox," which they fail to clarify by qualifying the discussion as dealing only with a homogeneous fluid. Additionally, many introductory college texts do not deal with hydrostatics at all.

Analytical Mechanics

A second cluster of problems is titled KICKO.[4] This problem cluster centers on the interaction of rotational and translational forces and inertia when a body is set in motion. The KICKO problems are as follows:

1. A thin, rigid rod of weight W is supported horizontally by two vertical props at its ends. One of these supports is kicked out. Find the force on the other support immediately thereafter.

2. A winch is used to lift a load of mass $\underline{m}$. The winch is a solid drum of radius R and mass M. If the engine applies a torque, τ, to the winch drum, what is the acceleration of the load?

3. Two spheres are connected by a thin, rigid rod (of negligible mass). The spheres have equal mass and each is supported by a prop. One of these supports is kicked out. Find the force on the other support immediately thereafter.

4. Three particles are connected at the ends of a "T" made of thin, rigid rods. A large mass is attached to the base of the T with small equal masses at each end of the cross on the T. The system is supported by a prop under the large mass and a frictionless pivot (hinge) at the juncture of the T. The prop is kicked out. Find the force on the pivot immediately thereafter.

5. A log of length ℓ and weight W rests on a frictionless slab of ice. A chain is attached to one end of the log and attached to a truck. A stake is driven in behind the opposite end of the log to prevent it from "kicking back." Through the chain, the truck applies a force F to the log. Is the stake necessary and, if so, what is the force on the stake (at the instant the truck starts up)?

6. A uniform, thin, rigid rod supported at a 30° angle below the horizontal by a frictionless pivot (hinge) at one end and a vertical thread at the other end. The thread breaks loose. Find the vertical force supplied by the pivot immediately thereafter.

[4]The initial problem in the KICKO problem cluster can be found in the 1978 graduate school examination in the Physics Department at Saskatchewan University, Saskatoon, Saskatchewan, Canada. The remaining problems in this cluster are expansions of the original problem.

The introductory problem is a rod set at each end on a point support: If one support is abruptly removed, what change occurs instantaneously to the force on the support? When fully analyzed, the force is found to be reduced as the rod is starting to fall and rotates. It is not obvious to most subjects that rotation must be included in the analysis. Subsequent problems in the cluster are designed to clue them to the rotational issues. The first problem can be solved as follows: Before removing the support, the total torque and force on the rod must be zero. This implies that each pivot provides a force of one half the weight of the uniform rod. When one support is removed, the total force on the rod is $F = F_p - mg$. The total torque on the rod is $\tau = mg\ell/2$ where ℓ is the length of the rod. The rotation of the rod about the remaining pivot point is governed by the equation: $\tau = I(d\omega/dt)$ and ω (the angular speed of rotation) is related to v (the velocity of the center of mass) by: $v = \omega\ell/2$. Thus, $\tau = mg\ell/2 = 2Idv/\ell dt$, and we can substitute $I = \ell/3m^2$ (for a thin rod) to get:

$$mg\ell/2 = I/3m^2(2/\ell)dt/dt = 3/4g$$

and, since $F = m(dv/dt) = F_p - mg - 3/4mg = F_p - mg$, then $F_p = 1/4mg$. The force instantaneously drops from $1/2\ mg$ to $1/4\ mg$, a factor of $1/2$.

The solutions to problems in this cluster are similar to those above. They may also be obtained easily using the more formalized Lagrangian or Hamiltonian mechanics.

In all the problems in this cluster, one must calculate both forces and torques to determine the linear and rotational motion of the body. Once this is done, the acceleration of the center of mass and the rotation of the body can be used to solve for a numerical value of the forces or accelerations. This cluster emphasizes the close ties between linear and rotational motion.

The source of the initial problem in the cluster is the graduate school examination of the University of Saskatchewan, Department of Physics, 1978. This question is also modified and simplified for our use. Many college-level texts discuss similar problems, but typical problems deal with ladders leaning on a wall and masses at the ends of thin rods. (See, for example, Freier, 1965, pp. 121-133; Halliday and Resnick, 1970, p. 203.) Sears and Zemansky (1970, p. 130) emphasize that such problems must account for linear and rotational motion, but never ask the question as it is presented in the KICKO cluster. Most high school level texts (e.g., Project Physics Course, Rutherford, et al., 1970) deal only remotely with torques and rotational motion, so it is expected that undergraduates will be novices in the analysis of these problems.

Electrodynamics

A third cluster titled CURRENTS[5] asks subjects to differentiate between currents in wires and beams of charged particles. Subjects are asked to determine the forces on current-carrying wires when the observer is (a) stationary with respect to the wire and (b) stationary with respect to the moving electrons in the wire. Then the forces are calculated for a moving beam of electrons in the wire and in a reference frame moving along with the cluster beam. The problems in this cluster are as follows:

1. Two wires, spaced d centimeters apart, each carry currents I in the same direction. Is there a force felt by a segment of one of the wires? If so, what is the direction of the force?

2. An observer is moving along the wires mentioned above at the same velocity as the electrons (their drift velocity). Thus, the observer sees the electrons at rest in her or his reference frame. What force is felt by the wire in this reference frame? Why?

3. Two parallel electron beams of current I are emitted in a dual beam oscilloscope from separate electron guns. Do the beams deflect toward or away from one another? Why?

4. Now an observer travels at the same velocity as the electrons leaving the guns of a dual beam oscilloscope. What force does this observer see affecting the electrons' motion?

5. You have now worked on four problems. What are the physics principles involved? How are the problems related and how are they different? Hint: Are there differences between the beam of electrons and a current of electrons? What forces act on each electron in the two cases?

For the current-carrying wires, the force on the wires is magnetic. The magnetic field of one wire or the other, each carrying current separated by r, is $B = I/r$. The force on a unit length of current-carrying wire is $F = IB$. If the currents are in the same direction, the wires attract, and vice versa. If the observer is stationary with respect to the moving electrons, the current is still present. (The positive ions are moving by the observer in the opposite sense as the electrons.) Again, the wires attract if the currents are in the same direction, and vice versa.

In the case of beams of electrons, the magnetic field still influences the motion of the electrons. However, the electrons of each beam repel electrostatically. (The wire is electrostatically neutral in

---

[5]The initial problem in the CURRENTS cluster can be found in Berkeley Physics Course: Electricity and Magnetism (Vol. II) by E. M. Purcell, 1965. The remaining problems in this cluster are expansions of the original problem.

charge.) Thus, beams traveling in the same direction repel by the force $F = q_1q_2/r^2$. This force dominates the magnetic force between the beams.

In the current-carrying wire problems, the interaction between the wires is chiefly magnetic. (The net charge on the wire remains virtually zero.) In the beam problem, the interaction is both electric and magnetic, with the magnetic interaction being very small except for beams traveling at speeds near the speed of light. The beam problem can be seen most easily by going to a reference frame in which the electrons are stationary. In this frame, the interaction is electrostatic repulsion only. The forces remain nearly the same in any other inertial reference frame except near the speed of light, where relativistic corrections to time and length cause the force to be different from that in the electron's rest frame.

These problems are easily solved from magnetic field, electric field, and Lorentz force calculations. However, as in previous clusters, subjects must realize the difference between various forces and when each force is effective.

Subjects tend to mistake a beam of charged particles for a neutral current. This is the misconception that the CURRENTS cluster explores. When subjects realize the difference between currents and beams, they then can sort out the applicable force laws. However, if the difference is not understood, the calculation leads to contradictory forces (i.e., the magnetic force attracts, the electrostatic force repels) and subjects often are left in a quandary. They see two identical (as far as they can tell) physical settings leading to forces of opposite direction.

The problems can be solved by the following schemes:

- Calculating forces on the wire and beams from $B = \oint I d\ell$ and $\vec{F} = \vec{I} \times \vec{B} d\ell$ and $E = \int_{\infty}^{\infty} q_1/r^2 (d\ell)$ and $\vec{F} = qE$. Here the current is calculated in each separate reference frame.
- Calculating as above, but transforming the fields and currents from a single reference frame, using a four-vector transformation.
- Calculating as above, but neglecting the positive ion currents when moving with the current-carrying electrons.

The CURRENTS cluster was inspired by the treatment of the <u>Berkeley Physics Course, Volume II, Electricity and Magnetism</u> (Purcell, 1965). This is the only attempt found on the introductory level that shows how fields and currents are transformed as one changes inertial frames of reference. A similar discussion occurs in Feynman's (1960) "Lecture on Physics," and a more advanced treatment is given in <u>Classical Electrodynamics</u> by Jackson (1975). The force exerted on a current-carrying wire is a typical textbook problem (Halliday and Resnick, 1970, p. 522), and most texts have problems asking for the force on a beam of electrons in a magnetic and electric field (ibid., p. 553; Sears and Zemansky, 1970, p. 43?). However, no case was found of two beams interacting (except in the Berkeley text). Thus, it was found that most subjects

are not familiar with the transformations of electric and magnetic fields to different reference frames.

Quantum Mechanics

The fourth cluster of problems deals with a subject normally treated in chapters on "modern physics." Rudolf Mossbauer (in Eyges, 1965) found that gamma rays from nuclear decay would not be absorbed and re-excite an identical nucleus. This fails to occur because a small amount of excitation energy is lost in the recoil kinetic energy of the decaying nucleus. This slight energy loss of the photon is enough to make its energy less than needed to excite another identical nucleus. Four questions comprise the DECAY[6] cluster:

1. A free ^{57}Fe-excited nucleus decays with a lifetime of 10^{-9}s by emitting a 14.4 Kev γ ray. Can this γ ray excite another free ^{57}Fe nucleus from its ground state to the excited state? Why?

2. Recall the previous problem; if the source and target nuclei were in a lattice of a crystal, could such resonant absorption occur?

3. Which of the following reactions are kinetically possible?
   (a) $\gamma + e^- \rightarrow e^-$
   (b) $\gamma \rightarrow e^- + e^+$
   (c) $e^- + e^+ \rightarrow \gamma$
   (d) $\gamma + e^- \rightarrow e^- + e^- + e^+$

4. A nucleus of mass M at excitation energy E_o emits a photon. Calculate the frequency of the photon.

The DECAY cluster asks under what conditions the resonant absorption can occur. The absorption will occur only if both nuclei are in lattices of many nuclei and the recoil energy loss or emission and absorption is vanishingly small. As one subject said, this was a very exciting phenomenon when discovered, but its physics was fully understood in a matter of months. The Mossbauer Effect has many modern-day applications, but little active research is being done.

The initial question of the DECAY cluster was used in the University of Saskatchewan, Physics Department, comprehensive exam of 1978. Most modern physics tests refer to the Mossbauer Effect, but no examples were found of the calculation required in the DECAY cluster (e.g., Frier, 1965, p. 561; Feynman, 1960, pp. 42-51). Neither Sears and Zemansky (1970) nor Halliday and Resnick (1970) discuss the Mossbauer

---

[6]The initial problem in the DECAY cluster can be found in the Eyges, L. article, Physics of the Mossbaur Effect, American Journal of Physics, Vol. 33, 1965, pp. 790-802. The remaining problems are an expansion of the original problem.

Effect, though it would seem to be a nice problem in classical mechanics, as shall be shown. Weidner and Sells (1973, p. 342) discuss the Mossbauer Effect in their Introductory Modern Physics but again fail to calculate the energy loss of the photon in detail.

The third question is asked to prompt the subject's use of conservation principles. The interaction described in Question 3 can occur only if momentum and energy can be conserved simultaneously. The subject is required to use these conservation laws in Question 3, just in case he or she has not used them previously to analyze the problems of Questions 1 and 2.

To obtain a complete answer, one must calculate the energy loss due to nuclear recoil and show that, if the entire crystal recoils, the energy loss is small. The calculation can be done classically, assuming correctly that the recoil speed of the nucleus is much less than the speed of light; or relativistically, which requires slightly more complex mathematics. In either case, the answer becomes: $E = E_o (1 - E_o/2Mc^2)$, where E_o = excitation energy of the nucleus; M = mass of the recoiling nucleus. If a single nucleus recoils, one uses M^2c = 2 GEV/nucleon, and E_o = 10 Kev, and the energy lost will exceed the resonant band width of most long-lived nuclear states. If the entire lattice recoils, M becomes 10^{22} times larger and $(E_o/mc)^2 = 10^{28}$, becoming insignificant.

The problem appears more difficult than it is. It can be solved classically by energy and momentum conservation, so long as the energy of the photon is written as $E = pc$, where p = photon's momentum. The basic solution is outlined as follows: The decay of the nucleus conserves momentum and energy. Assuming the nucleus recoils at much less than the speed of light, then

momentum before = momentum after,

$$\emptyset = p + p_{nucleus}$$
$$\emptyset = E/c + p_n$$
$$p_n = -E/c$$

energy before = energy after, $E_o = E + p^2/2m$

substituting for the momentum of the photon, $E_o = E + E^2/2mc^2$.

Here we can solve by the quadratic equation or assume $E_o \sim E \cdot mc^2$, and write $E = E_o - E_o{}^2/2mc^2 = E_o|1 - E_o/2mc^2|$.

As seen above, $E_o/2mc^2$ is small, meaning the assumption $E_o \sim E$ is good.

A second, more exact, solution includes the relativistic effects by writing total energy as $E_t2 = m^2c^4 + p^2c^2$ for the nucleus. The solution is slightly more complex, and the result is only slightly more exact.

This cluster explores the subjects' ability to use classical approximations to relativistic problems. Most students are not comfortable with these approximations and find difficult solutions to the simple problems.

Thermodynamics

The cluster titled THERMO[7] deals with a simple problem of heating air in a building.

1. We wish to calculate the change in internal energy of the gas in a room filled with air. Initially, the room is at atmospheric pressure and 0°C. The air is then heated to 20°C and air is free to flow to the outside. Is there a change in internal energy of the gas?

2. In a problem similar to the last one, we begin with a cylindrical, frictionless, massless piston above a one liter cylinder of air. The air is heated from 0° C to 20°C. How much energy does this use? What is the change in internal energy of the gas? Is there a difference between the input energy and the change in internal energy of the gas?

3. Typical heating systems cycle over 3°F to 5°F, heating the room's air after it has cooled off significantly. Can you suggest a more efficient way of heating the air? That is, how can we minimize energy consumption in space heating?

4. Can you design a more efficient shelter to minimize energy consumption?

The subject is asked to calculate the change in energy of the air in the room as it is heated from 0°C to 20°C. At first glance, the problem appears trivial: Calculate $\Delta U = mc\Delta T$, where m is the mass of air in the room, c is the heat capacity of the air, ΔT the change in room temperature, and ΔU the change in energy. One must assume the gas remains either at constant volume or pressure. (The heat capacity differs for the two cases.) Then, substituting reasonable values for the unknowns, the answer is obtained.

However, the gas must expand as it is heated. Thus, some of the air leaks out of the room and one must account for this loss of air when calculating the change in internal energy of the room. The correct solution to this problem would use the heat capacity at constant pressure. Also, the volume of the heated gas must be calculated. A possible solution would be as follows:

---

[7]The initial THERMO problem was derived from Sommerfield's discussion of the second law of thermodynamics in Thermodynamics and statistical mechanics, Academic Press, 1956. The remaining problems in the cluster are expansions of the original problem.

Assuming an ideal gas, $pV = (m/M)RT$

p = pressure
V = volume
m = mass
M = molecular weight of air
R = gas constant
T = absolute temperature

The change in volume for $0^{o}C$ to $20^{o}C$ is $p_0V_0/P_{20}V_{20} = T_0/T_{20}$ and, at constant pressure, $p_0 = p_{20}$, $V_0/V_{20} = T_0/T_0$. The change in energy of the gas in the room must be reduced by the above fraction, since only V_0/V_{20} of the air remains in the room.

Since the gas is assumed ideal, the internal energy is a function of temperature only. Thus, $U_0 = m_0CT_0$; $U_{20} = m_{20}CT_{20}$, where C is the specific heat of the gas and M_0 and m_{20} are the mass of the gas in the room at that temperature. Since the mass of the gas in the warmer room is reduced by V_0/V_{20}, we get:

$$U = U_0 - U_{20} = m_0CT_0 - m_{20}CT_{20}$$

$$= C[m_0T_0 - m(V_0/V_{20})T_{20}] = Cm_0[(T_0 - T_0/T_{20})T_{20}] = 0$$

Thus, no additional energy is added to the gas in a room as it warms. The heating pushes gas out of the room. All of the heating energy is essentially lost.

The solution above is simple and straightforward. The principles invoked are:

1. An ideal gas obeys the equation state: $pV = nRT$.

2. The internal energy of an ideal gas depends on temperature only (i.e., $U = C_mT$).

An alternate solution may be: Calculate the heat needed to heat the room from $0^{o}C$ to $20^{o}C$ ($Q = mcp\Delta T$). Then subtract the work done against the surrounding as ($W = p\Delta V$) to find the change in internal energy of the gas. The fraction of the gas that escapes the room is subtracted, resulting in no change in internal energy of the gas that remains within the room. This would be more complex algebraically, but would still result in no change in the energy of the room.

The second question asks to calculate the energetics of expanding a gas at constant pressure. This emphasizes the concept that energy (heat) added to a gas causes it to expand. The solution is straightforward and shows a significant difference between the heat input and the internal energy of the gas--the difference being the work done by the gas on the surrounding atmosphere. The third question is asked in order to explore the subject's understanding of the thermodynamics of an ideal gas.

The THERMODYNAMICS cluster was derived from Sommerfeld's (1956) discussion of the second law of thermodynamics in his book Thermodynamics and Statistical Mechanics. On pages 40 and 41, Sommerfeld poses

the question, "Why do we have winter heating?" After a short analysis he states, "The energy content of the room is thus independent of temperature, solely determined by the state of the barometer." We have not found this question asked or discussed in any other thermodynamics text. Thermodynamics sections of introductory college texts ask questions similar to Question 2 (e.g., Halliday and Resnick, 1970, pp. 371-374), but the system is always a closed system. Sears and Zemansky (1970, p. 247) emphasize thermodynamic systems that are closed, but do not discuss the consequences of being an "open" system. Thus, it is expected that most subjects will be uncomfortable with their answer to the first problem until they do the second problem.

One important aspect of the simple solution given above is that internal energy of an ideal gas is a function of temperature only. This principle is emphasized in Halliday and Resnick (1970, pp. 382-384), Sears and Zemansky (1970, p. 268), and most other texts. One expects, however, that it will not be evoked as a simplifying principle by most subjects. This is because it is not an intuitive response to calculations of energy changes of gases.

Some Results and Tentative Conclusions

Twelve subjects, four faculty in the Department of Physics at Minnesota, four graduate students in physics, and four undergraduates majoring in physics, are serving as subjects in the project research. These subjects are given the problems in each cluster in a two-hour, problem-solving session. As stated earlier, results of these session are tape-recorded, transcribed, and analyzed in terms of lines of reasoning employed. Some of the results for the MILKO cluster are presented next.

Lines of Reasoning

The lines of reasoning employed by subjects in working problems in the MILKO cluster can be characterized in terms of solution framework and solution parameter. The solution framework consists of the general representation that is given to the phenomenon stated in the problem. This representation is of two types:

1. Each problem is attacked as a single system throughout the problem-solving process--this is called a Type 1 solution framework. In the MILKO problems, this means that subjects do not distinguish between the mixed and separated states of the fluids in the containers in their reasoning steps or in their calculations.

2. Each problem is attacked as a series or sequence of states with the emphasis on the transition or comparison between states--this is called a Type 2 solution framework. In the case of the MILKO problem, this

means that subjects consider the mixed and unmixed states of the fluids separately in their reasoning steps and in their calculations.

The solution parameters employed in solving a given problem consist of the physics principle or principles, often equations, upon which the answer depends. In the case of the MILKO problems, two distinct parameters have been identified. One parameter is based on the concept of pressure defined as hydrostatic height, $P = \rho \underline{g} h$. The other parameter is based on the concept of pressure defined as force and formalized as $P = F/A$.

The parameter based on pressure defined as hydrostatic height can be further subdivided into two types:

The pressure at the base is expressed as the sum of the densities and heights of the corresponding fluids, $P = \Sigma \rho_i \underline{g} h_i$. This is called a Type 1 solution parameter.

The pressure at the base of a container is expressed as an instance of Pascal's general principle, namely, that the pressure in two points at the same level of liquid is constant. This is called a Type 2 solution parameter.

The parameter based on the pressure defined as force can be further subdivided into two types:

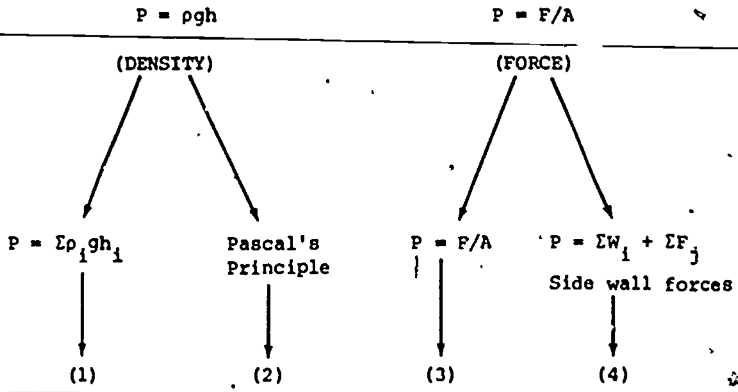
The pressure at the base is the weight (force of the liquid above the base) divided by the area. This is called a Type 3 solution parameter.

The pressure at the base is the weight of the liquid in the vessel plus the forces exerted vertically by the side walls--this is called a Type 4 solution parameter.

The above solution frameworks and solution parameters can be combined, since problem solving consists both of choosing a framework and selecting a solution parameter. This combination yields eight possible lines of reasoning for solving problems in the MILKO cluster. These lines of reasoning are shown in Figure 1, on page 104, and characterized briefly below. (The first number in parenthesis refers to the type of solution framework; the second number refers to the type of solution parameter.)

(1,1) This is unlikely to exist, because as soon as pressure is explicitly expressed as the sum of $\underline{h}_i$, ρ_i, and $\underline{g}$, the two states have to be considered separately.

(1,2) Since this line of reasoning involves general principles it is likely to appear as a starting point. It does not lead to the right solution. The subject thinks of the problem without referring to the two states and recalls the general definitions and the relation of pressure to density. Based on the fact that, before and after mixing the "stuff" above, the base is the same and equally distributed, he or she gives the wrong answer.



	(1)	(2)	(3)	(4)
System (1)	1,1	1,2	1,3	1,4
State (2)	2,1 *	2,2	2,3 *	2,4 *

O ≡ Unobserved lines of reasoning
* ≡ Successful lines of reasoning

Figure 1. Framework for Characterizing Lines of Reasoning

(1,3) The general definition of pressure is recalled by the subject. The pressure as total weight of fluid over area is the same in both cases. The answer is again wrong, as with any solution involving a system framework.

(1,4) Again, this one is unlikely to exist. If the subject recalls the role of side wall forces, he or she typically sees that they are different in the two cases, so that they must be worked out independently.

(2,1) The relationship, $P = \Sigma\rho_i g h_i$, is used to study the two cases separately. This line of reasoning can lead to a correct solution both in a calculation problem, as well as in a general one (without numbers). For example, the subject assumes that the average surface of each liquid will be treated as a parameter of the problem (i.e., assuming the subject does not work with a cylindrical vessel).

(2,2) This line of reasoning is difficult to distinguish from (2,1). The difference is that here the subject only recalls the general pressure relationship without working it explicitly, in which case he or she is using (2,1). Thus, (2,2) cannot give the correct answer and is likely to represent an intermediate rather than a final stage in the reasoning process.

(2,3) The definition of P = F/A is applied separately for the different states. It can lead to a correct solution only when applied carefully ~~in the calculation problems.~~ It usually comes together with the visualization of the central column.

(2,4) The subject recalls that the force exerted on the bottom is not due to the weight of the fluid alone, but also to the vertical forces exerted by the side walls. This line of reasoning can lead to a correct solution both for a general problem and for a calculation problem.

Six of these eight lines of reasoning are plausible means of attacking the MILKO problems. These six lines of reasoning are illustrated as sequences of problem-solving steps in Figures 2a, 2b, and 2c, below and on the following page.

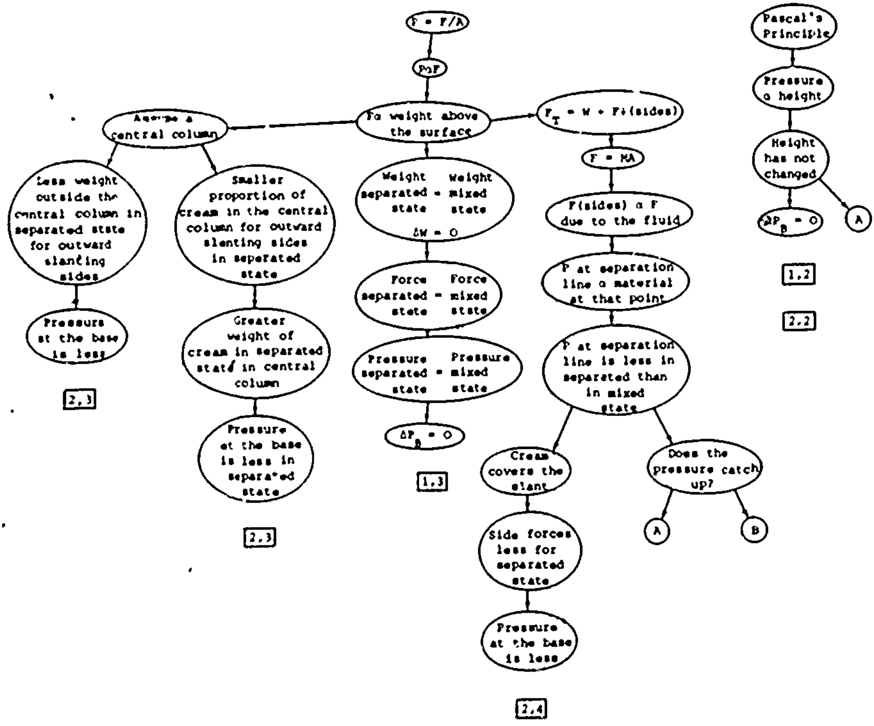


Figure 2a.

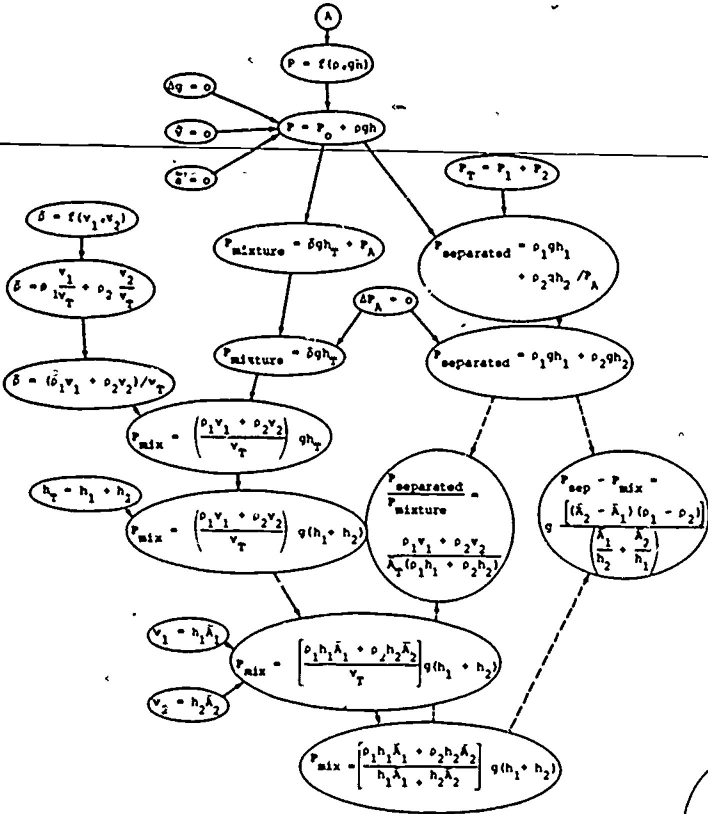


Figure 2b.

2,1

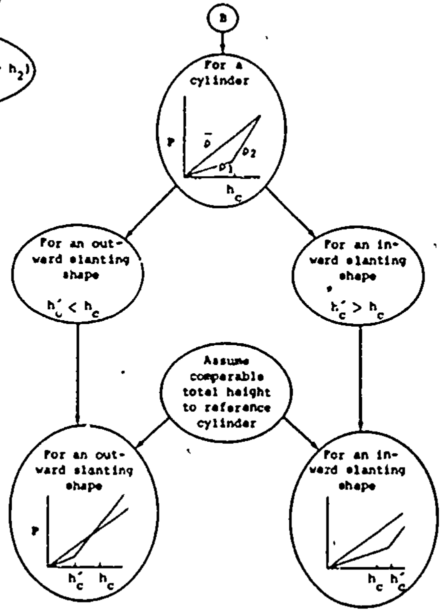


Figure 2c.

2,2

The MILKO Cluster

The first question asked is how subjects responded to the cluster of problems. Figure 3, which presents the average number of problems solved by subjects in each group, shows that the more expert subjects solved more problems in the cluster than the less expert subjects. The difference in average number of problems solved was greater between graduate students and undergraduates than between graduate students and professors. The data in Figure 3 tell only part of the story, however. In addition to knowing how many problems in each cluster were solved, the investigators are also interested in which problems were solved, the point in the sequence of problems at which a correct solution first occurred, and the line of reasoning used to achieve a correct solution.

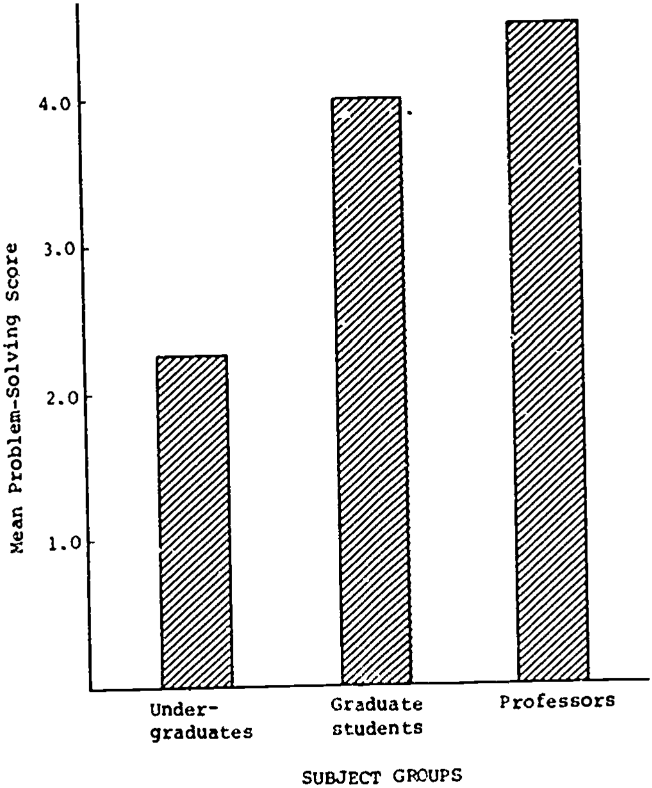


Figure 3. Mean Problem-Solving Score for Subjects in Each Group

Figure 4 shows the lines of reasoning employed by the undergraduate subjects. Figures 5 and 6, on page 109, provide similar information for graduate students and professors, respectively. According to Figure 4, four of the six plausible lines of reasoning were used by undergraduate subjects. Two of these lines of reasoning (2,1 and 2,3) resulted in

112

success and two others (1,3 and 2,2) did not. Figure 4 also shows that some lines of reasoning were begun for one problem in the cluster and then dropped in favor of a different line of reasoning in a subsequent problem.

In the case of some problems (ERLENO and CALCULO), a particular line of reasoning was abandoned within a problem and replaced by one of the alternatives. One undergraduate hit upon a successful line of reasoning in the ERLENO problem and another undergraduate adopted a successful line of reasoning in the CALCULO problem. Both of the successful lines of reasoning used by the undergraduates were based upon representation of the problem in terms of states rather than a system,

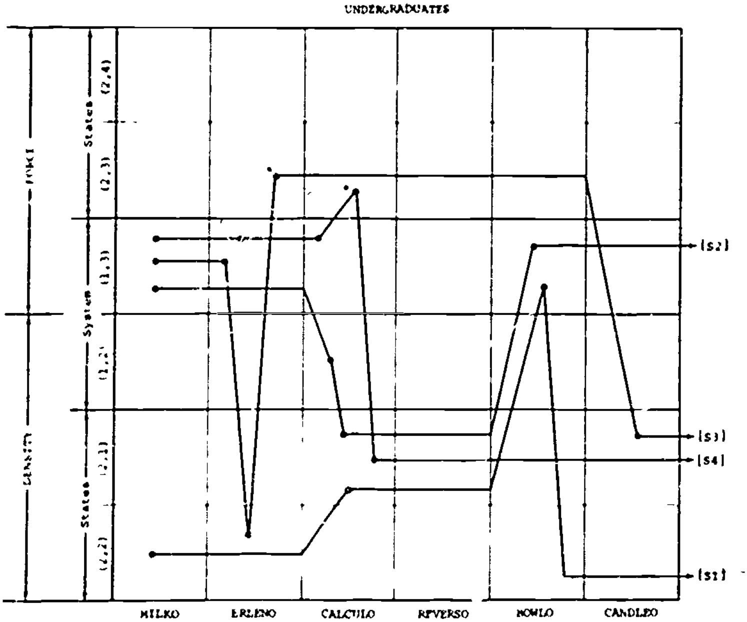


Figure 4. Lines of Reasoning Used by Undergraduates

and both successful lines of reasoning employed a density, rather than a force, solution parameter. Three of the undergraduates worked the original problem in the cluster the same way, namely, by adopting a line of reasoning (1,3) based upon P = F/A and a representation of the problem as a single system rather than as a series of states.

Figure 5 shows that all the graduate students were successful; they were also more similar to each other than undergraduates in the lines of reasoning they employed. Like three of the undergraduates, all of the graduate students began with line of reasoning (1,3). Unlike the undergraduates, the graduate students did not pursue this unsuccessful line of reasoning beyond the CALCULO problem. In the case of the professors, Figure 6 shows that all subjects were successful, although, like the

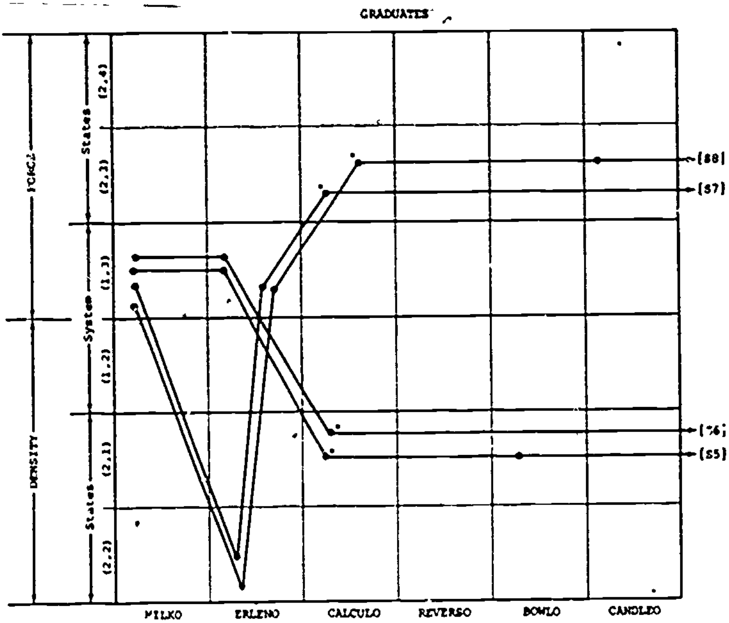


Figure 5. Lines of Reasoning Used by Graduate Students

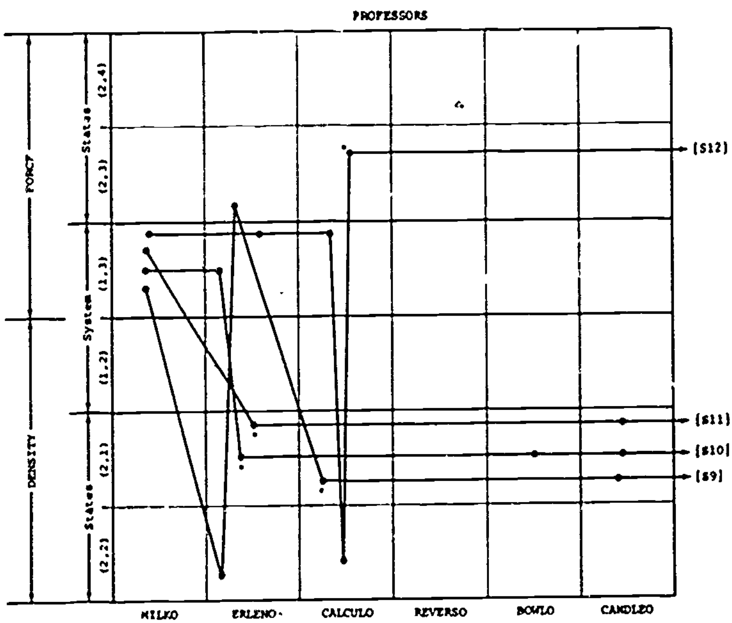


Figure 6. Lines of Reasoning Used by Professors

undergraduates and graduate students, they began with line of reasoning (1,3). Three out of four professors pursued the same line of reasoning (2,1) once they achieved a correct solution.

Taken together, Figures 4, 5, and 6 also show that some lines of reasoning were used more than others, and that two lines of reasoning (2,1 and 2,3) in particular led to a correct solution. Undergraduates considered successful lines of reasoning less often than graduate students and professors, and even when they considered them, such lines of reasoning were more likely to be abandoned by the undergraduate subjects..

The final aspect of the data that should be presented concerns the impact a line of reasoning had upon the behavior of subjects in each group. Excerpts from protocols given by subjects were chosen to illustrate the effect of various lines of reasoning on subjects' thinking based upon the acceptability of the result they generated. These excerpts are presented in Figures 7, 8, and 9, on pages 111 and 112. Undergraduates tended to accept, but not understand, the result provided by their calculations (see Figure 7). Graduate students were convinced, on the basis of their calculations, and were successful if they hit upon a line of reasoning that permitted a correct solution. This solution behavior is shown in Figure 8. Professors, on the other hand, were both able to identify the implications of a correct calculation and also generalize it to other situations. Thus, professors would often create the REVERSO example for themselves on the basis of their CALCULO response.

An interesting exception to the behavior of subjects in the expert group occurred for subject 12, who obtained the correct calculations in the CALCULO problem, but refused to believe them. This individual essentially adopted a parallel solution strategy in which he followed through (correctly) on the implications of his CALCULO response, i.e., he did subsequent problems correctly. As the more extended excerpt from his protocol in Figure 9 shows, however, he treated this aspect of the solution as hypothetical, maintaining to the end that the pressure under any of the conditions in the cluster would not change!

Implications For Science Teaching

Because a particular effort of instruction cannot cover everything, there is always a question of selection or coverage. Presumably, selection from among the things that might be taught is done to foster transfer to new content and application. The good news is that such selectivity reduces what must be taught and makes day-to-day activities of instruction and learning manageable. The bad news is that concepts or ideas may be left out that cannot readily be filled in on the basis of what is learned. Since little guidance is offered other than tradition in how to select material for textbooks, there is a risk of repeating past mistakes. This seems to be the case in the types of problems here.

CALCULO

. . . The pressure above will be less--or the pressure on the bottom will be less in the second one, because the total pressure above it . . . Now I have to explain why I got that answer, right?

Okay, the pressure would be different by . . . (under his breath) .73, that's equal to 23 . . . 32 times oh 9.81 is equal to, no, that's grams. We want kilograms, so that would be .032 times 9.81 is equal to . . . .032 times 9.81 is equal to .314 newtons per cm. squared. And the reason for that is because in this one . . . (pause) This one would have a . . . The separated state would have a less . . . less of a density--or less of a pressure on the bottom than the unseparated state, the mixed state, because . . . there's less weight total above the bottom in the separated state than there is in the mixed state. There this pressure would be less . . .

. . . SO WHAT DO YOU CONCLUDE FROM THAT WITH RESPECT TO THE OTHER PROBLEMS? DOES IT ALTER YOUR ANSWERS?

. . . No, it doesn't, because in a uniformly shaped container the pressure on the bottom . . . is going to be equal . . .

Figure 7. Sample Protocol from One Undergraduate

REVERSO

. . . ..But this procedure (the cluster) helped me to understand the original problem . . . I assumed that the shape was independent--but as I go through the problems, then I see that the shape is important, and we had to consider what shape had more pressure for the separated state and what shape had less pressure for the separated state. So this is a very good procedure . . .

Figure 8. Sample Protocol from One Graduate Student

MILKO

. . . The pressure near the base of the bottle does not change because the pressure is the ratio of the force with which the gravity pulls on the contents of the bottle and the surface of the base of the bottle. Because both of these remain constant, the pressure should not change . . .

ERLENO

. . . The problem reduces to the previous one, of which we already know the answer. And the answer is, that the pressure on the bottom does not change . . .

CALCULO

. . . The thing is--the problem is that because the masses are the same in both cases the pressures have got to be the same, there's no way around that. So, I'm essentially sticking to my guns as far as the previous two problems are concerned. Point being that I would rather distrust my calculations than conservation of mass . . .

REVERSO

. . . The easiest way to do this I should have thought, again on the assumption that the previous stuff is right, is going back to this stuff about doing it in columns, and design a vessel in such a way that the narrowest part would be at the bottom. And then, again on the assumption that this is right, that kind of separation would lead to an increase in pressure upon separation . . .

BOWLO

. . . It would increase upon separation, the reason why, I have already discussed in the previous problem . . . (REVERSO) . . . one would be led to believe that the shape of the vessel is the one thing that is important, in particular, it is important whether the top of the vessel-free surface is broader or narrower than the bottom. Since, in the other case, the top is narrower than the bottom, in a bowl the top is broader than the bottom, and therefore one would expect the pressure to increase . . .

CANDLEO

. . . Okay, my real answer is that, sticking to my original guns, the pressure wouldn't change.

Figure 9. Sample Protocol from One Professor

It makes perfectly good sense to choose as examples for instruction just those instances of concepts or principles that seem most readily apparent. The results of instruction designed around such "clear cases" may be, however, that students are not able to either fill in or generate the details necessary to make application in nonprototypical situations. Although "clear cases" imply wide generality, they may not "work" in particular situations. In the case of the MILKO problem presented here, a situation is encountered in which the general prototype "Pascal's Principle" holds only for homogeneous fluids. This is stated in textbook discussions. The fact that one might encounter nonhomogeneous fluids in oddly-shaped containers is simply not addressed.

The issue for teaching is, simply, that one may need to consider much more carefully than previously whether examples foster the kind of transfer they are believed to do. The difficulty is more than simply choosing different examples, since any prototype could, in theory, lead to the same sort of difficulty. Rather, what the data suggest is that one may need to deal more explicitly with exceptions and elaborations of the prototypes presented here. The goal of instruction is not only to enlarge the vocabulary of prototypes over which students have mastery, but also to provide sufficient detail so that each prototype can be used in the domain of situations to which its principles and concepts apply. If experts are more successful than novices in applying what they know, it is in part because their knowledge is more richly detailed and penumbral than that of novices (Bromley, 1980). The work presented in this paper suggests that a dependence upon "clear-case" thinking may occur in areas of subject matter covered in introductory college physics as well as at more advanced levels, where the prototypes themselves may be more difficult to learn, quantum mechanics, for instance.

Research such as that begun at the University of Minnesota may have its greatest impact upon the field of science teaching through the identification and study of "holes" in the cognitive fabric woven by experience in a discipline such as physics. By understanding better what such holes are like, the lines of reasoning that support them and the lines of reasoning that overcome them, attention can be focused upon what should be taught, as well as what should be learned, in the science curriculum.

References

Bromley, O. A. Physics. Science, 1980, 209, 112-121.

Chase, W. G., and Simon, H. A. Perception in chess. Cognitive Psychology, 1973, 4, 55-81.

Cohen, H. A. The art of snaring dragons. (A. I. Memo 338). Boston: Massachusetts Institute of Technology, Artificial Intelligence Laboratory, May, 1975.

Department of Physics, University of Saskatchewan, Saskatoon. Graduate School Examination, 1978.

Rutherford, F. J., et al. Project physics course. New York: Holt, Rinehart, and Winston, 1970.

Eyges, L. Physics of the Mossbauer Effect. American Journal of Physics, 1965, 33, 790-802.

Feynman, R. P., Leighton, R. B., and Sands, M. The Feyman lecture on physics. New York: John Wiley and Sons, 1960.

Freier, G. D. University physics: Experiment and theory. New York: Appleton-Century Crofts, 1965.

Greeno, J. G. Trends in the theory of knowledge for problem solving. In D. T. Tuma and F. Reif (eds.), Problem solving and education: Issues in teaching and research. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1979.

Halliday, D., and Resnick, R. Fundamentals of physics. New York: John Wiley and Sons, 1970.

Hayes-Roth, F. Learning by example. In A. Lesgold, J. Pellegrino, S. Fokkema, and R. Glaser (eds.), Cognitive psychology and instruction. New York: Plenum, 1978.

Heller, J. I., and Greeno, J. G. Information processing analyses of mathematical problem solving. Paper presented at National Science Foundation Seminar on Applied Problem Solving, Northwestern University, March, 1978.

Jackson, D. J. Classical electrodynamics. New York: John Wiley and Sons, 1975.

Larkin, J., McDermott, J., Simon, D. P., and Simon, H. A. Expert and novice performance in solving physics problems. Science, 1980, 208, 1335-1342.

Neches, R., and Hayes, J. R. Progress toward a taxonomy strategy transformation. In A. Lesgold, J. Pellegrino, S. Fok'ema, and R. Glaser (eds.), Cognitive psychology and instruction. New York: Plenum, 1978.

Purcell, E. M. Berkeley physics course: Electricity and magnetism, Vol. 2. New York: McGraw-Hill, 1965.

Sears, F. W., and Zemansky, M. W. University physics. Reading, Mass.: Addison-Wesley, Inc., 1970.

Simon, D. P., and Simon, H. A. Individual differences in solving problems. In R. Siegler (ed.), Children's thinking: What develops? Hillsdale, N. J.: Lawrence Erlbaum Associates, 1978.

Weidner, R. T., and Sells, R. L. Elementary modern physics. Boston: Allyn and Bacon, 1973.

CHAPTER 8

UNDERSTANDING AND PROBLEM SOLVING IN PHYSICS[1]

Jill H. Larkin, Carnegie-Mellon University

Introduction

This paper is about what cognitive psychologists are beginning to know about how people learn to use formal quantitative knowledge, such as knowledge of physics principles, to relate and describe real-world phenomena. Such application of formal knowledge is central to meaningful use of mathematics or science of almost any level. The work described here has its theoretical roots in modern cognitive psychology, particularly in the branch called information-processing psychology. The basic view is that human cognition consists of the sophisticated processing of immense amounts of information, both information coming into the system through sensory organs and information stored internally in the brain. A good description of this basic view of the human mind is given by Simon (1979).

Most researchers in this area of psychology are concerned with building precise models of relatively limited aspects of cognition. This aim has two consequences for the way research often proceeds. First, mental processes are traced in detail, often by using think-aloud protocols (transcriptions of an individual's comments as he or she thinks aloud while performing a task). These processing traces are very different from performance data (e.g., test scores). Furthermore, the traces are often not averaged to give aggregate results, but used directly as the input for formulating models of individuals. The second implication is that the models constructed are mathematically formulated

[1]The introductory paragraphs and the literature review in this paper were originally part of a paper titled "The Cognition of Learning Physics," American Journal of Physics, in press, 1980. The discussion of naive and scientific representations, and of expert-novice differences, was originally part of a paper titled "Understanding, Problem Representations, and Skill in Physics," delivered at the NIE-LRDC conference on Thinking and Learning Skills, to be published as part of the proceedings of that conference.

and, because they are complex, they are usually implemented on computers. In the last ten years, the computer has given researchers a tool for mathematical modeling that is both precise and capable of handling the complexities required for at least first-order models of human knowing and learning.

In the recent past, cognitive psychology has developed immensely in exciting new directions. Ten years ago, studies in problem solving were only beginning and were largely concerned with how people solve puzzles or play games. Since that time, a large amount of interest has developed in problem solving of the kind done by students in various science and mathematics courses. The reason is clear. As psychologists began to understand the mechanisms of problem solving in simple and constrained situations (puzzles and games), they became ready to extend their theories to richer situations. However, they were not ready to take on the full complexity of problem solving done by professionals, or by individuals, in their personal life.

Textbook problems form a large and readily available set of problems that have been designed explicitly to capture some important features of real-world problems, and to be manageable for use in the classroom and, therefore, also in the psychological laboratory. Furthermore, the domains of science and mathematics (in particular, physics) have had more appeal than less structured domains (sociology, English literature), because the former domains are tautly organized around a relatively few, readily identifiable principles, a structure which facilitates understanding the psychological mechanisms of problem solving.

The trend has been from puzzles to structured textbook problems and is now beginning to address, first, the process of learning to solve problems and, second, open-ended problems more closely resembling those that might be encountered in real professional situations. In general, information-processing models have stressed cognitive factors, in contrast to social or emotional factors; but as the ability to understand mental phenomena increases, it is becoming possible to model some of these factors as well (Fiske and Linville, 1980).

This paper first discusses a sequence of work addressing the role of understanding in solving physics problems. In particular, the "scientific" understanding used by scientifically experienced individuals seems very different in content from the "naive" understanding characteristic of less experienced people. This difference in understanding relates to important differences in the problem-solving approaches of experts and novices in science, differences that are beginning to show us what knowledge is essential for expertise.

The later sections of the paper provide a brief review of literature relating to the psychology of learning and knowing in scientific domains.

To make the discussion more concrete, consider the following problem in elementary mechanics:

A toboggan of mass m starts from rest at a height h above a valley down a perfectly slippery hill. As shown in Figure 1, the toboggan then moves up over a hill with a circular cross section of radius R and a top located a distance y above the valley floor. What is the height h if the toboggan just barely leaves the snow surface at the top of the circular hill?

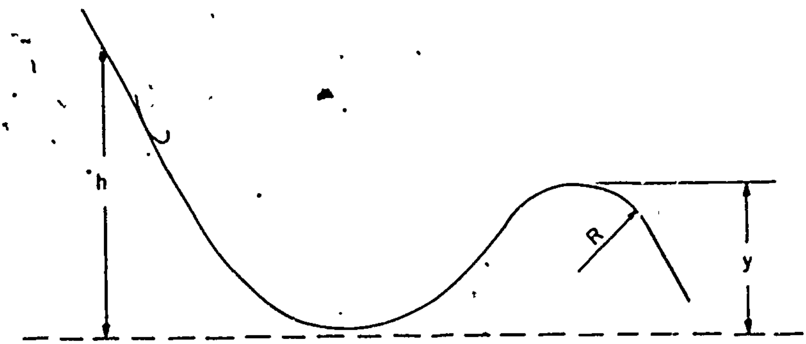


Figure 1. A Toboggan on a Slippery Hill

What happens when an individual reads and understands this problem? Any understanding involves processing the string of words in the problem to construct some internal representation consisting of entities and relations between them. These entities must be meaningful in that they are connected to many other entities in the memory of the understander. It is necessary to distinguish among several kinds of understanding based on the nature of the entities used in the internal representation of the problem. Most people probably construct for the preceding problem what could be called a naive representation, composed of entities related to familiar situations. Figure 2, on page 118, suggests elements in such a representation. The concept of moving is central. Three locations are mentioned, and for each there are attributes of the motion, such as starting from rest, moving fast, or the roller-coaster sensation of just skimming the top of a hill. A more adequate reflection of how people represent this problem would be much more complex and would certainly be different for different individuals.

The important point is that, although it is called naive, this every-day representation of the problem is not trivial, and constructing it depends on a lot of pre-existing knowledge about toboggans and other things (e.g., roller coasters). The importance of rich, intensely structured knowledge in understanding all prose is being explored intensively in the development of computer-implemented language

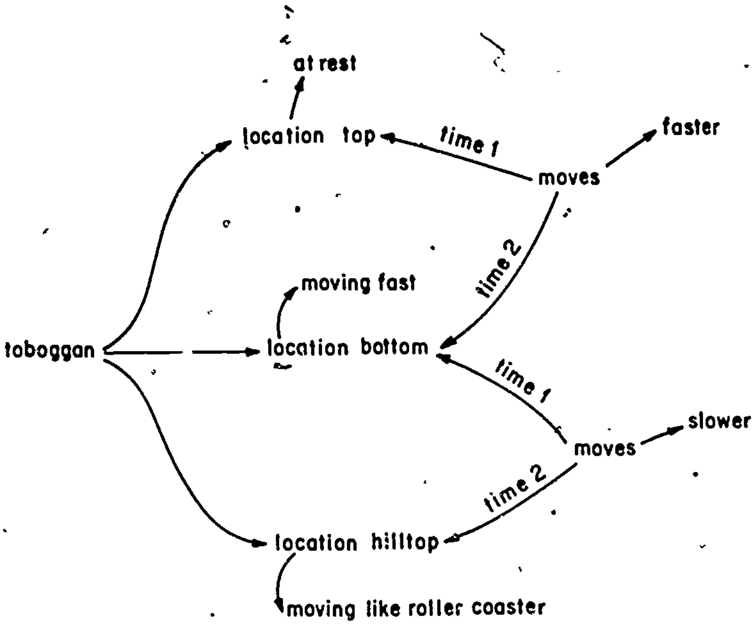


Figure 2. A Naive Representation for the Toboggan Problem

understanders (Shank and Abelson, 1977) and in studies of how people understand stories (Rumelhart, 1975). This paper is concerned less with this naive understanding than with the special kind of understanding that comes from the knowledge possessed by individuals trained in science.

Scientific Understanding

The kind of understanding described above involves simply building some internal representation of the important entities and relations in a text segment. As described above, this kind of understanding is far from trivial, and doing it effectively involves a great deal of knowledge about what kind of entities and relations are important. However, the kind of understanding discussed above is still not sufficient to solve problems.

In solving physics problems, the operators for producing new information involve physics principles, relations that allow the generation of new information from existing information. Solving a problem consists of applying a sequence of these principles until the desired information is generated. The difficulty is that the operators, as usually taught and used by expert solvers, often do not act directly on naive problem representations. For example, the toboggan problem

discussed earlier can be solved through some combination of the following main principles:

- The total force on a system is equal to its mass times its acceleration.
- The total energy is conserved for a system on which no non-conservative forces act.

But a naive representation for these problems, one constructed by a novice solver, (Figure 2) does not contain any forces or energies or systems. These entities are physics entities. They can be added to a problem representation only by a solver who has special knowledge of how to construct these entities from those appearing in the naive problem representation. The term scientific representation will be used for representations that include entities such as systems and forces that appear in some scientific discipline (physics) but not in the naive understanding of untrained individuals.

Figure 3 shows a hypothesized scientific representation for the toboggan problem. The representation for the atwood-machine problem is, in fact, very similar to representations built by PH632, the physics problem solver built by John McDermott and Jill Larkin and intended to replicate major steps of the work of an expert solver (McDermott and Larkin, 1978).

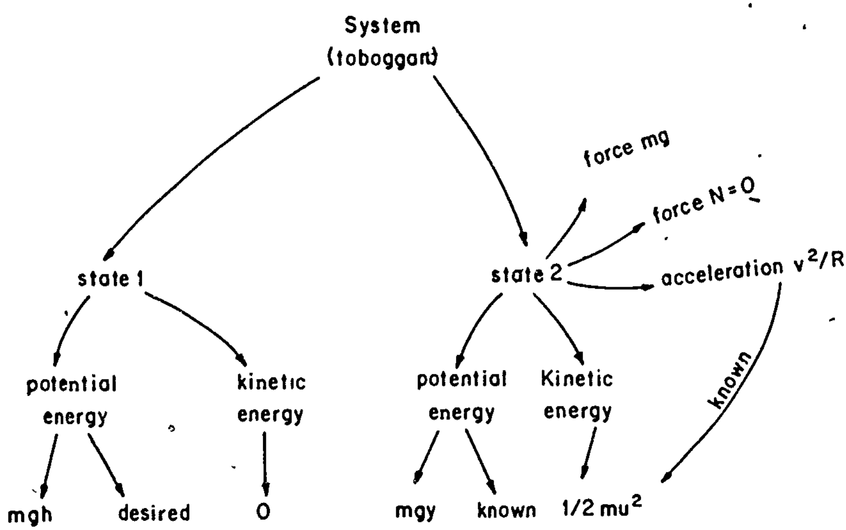


Figure 3. Scientific Representation for the Toboggan Problem

The scientific representation is more abstract than the naive representation, since many entities in the naive representation correspond to one scientific entity. For example, any object or collection of objects can be a system. Any contact (or other interaction) between systems can be characterized as a force. This abstraction is central,

even key to making unfamiliar problems familiar. Two problems with very different naive representations can have identical scientific representations. Thus, if one has the knowledge to solve one of them in its specific representation, the other problem becomes trivial once the scientific representation is constructed. For example, the scientific (energy) representation of a freely moving pendulum is identical to the scientific representation given in Figure 3 for the toboggan moving down the hill.

Abstraction, of course, comes in varying degrees. For example, the physics problem-solving program, ISAAC, developed by Novak, uses as its main problem representation something between the naive and scientific representations described in this paper (Novak, 1976, 1977; Novak and Araya, 1980). The entities in ISAAC's representations are idealized real-world objects, for example, perfectly rigid levers and noncompressible supports. ISAAC does not construct a more abstract physical representation involving entities like torques that do not correspond to real-world objects. Instead, its operators act directly on the collection of idealized objects to generate sets of equations.

Once a scientific problem representation (like the one in Figure 3) is constructed, it is relatively easy to imagine how operators reflecting knowledge of physics could act to produce equations needed to find desired values or expressions. For example, for the toboggan problem a full statement of the principle of conservation of energy is:

> If a system goes from one state to another and no work is done on it, then the energy of the initial state equals the energy of the final state.

This statement provides a template for an equation that can be written using expressions or values for the initial and final energies of the toboggan, entities that appear in the relevant scientific representation (Figure 3). The resulting equation, $\underline{mgh} = 1/2\underline{mv}^2 + \underline{mgy}$, provides a large part of the solution of the toboggan problem.

To distinguish the equations used to solve the problems from the scientific understanding of the problem discussed earlier, these equations will be called the mathematical representation of the problem. Mathematical representations are still more abstract than scientific representations. In other words, one entity in a mathematical representation, e.g., an equation in a particular form, can characterize scientific representations for several different problems. A striking example is an alternating-current circuit with capacitance and inductance. Compared with the pendulum and the toboggan problems considered earlier, this situation has a very different naive representation and a substantially different scientific representation, but it has an identical mathematical representation.

Expert-Novice Differences

This sequence of problem representations (naive, scientific, mathematical) has been described with the aim of stating and supporting with evidence the following hypothesis:

> The central difference between expert and novice solvers in a scientific domain is that novice solvers have much less ability to construct or use scientific representations.

The most easily interpreted evidence is that of Chi, Feltovich, and Glaser (1980). They find that, when asked to sort physics problems into categories, inexperienced solvers use categories corresponding to what is called here a naive representation of the problem. Specifically, their categories corresponded to objects mentioned in the problem, such as wheels or inclined planes. In contrast, experienced solvers sorted problems according to what is called here a scientific representation, with categories corresponding to force problems or energy problems.

A second source of evidence is the puzzling, but now repeatedly documented phenomenon that when generating equations expert solvers tend to work "forward" starting with equations that involve mostly known quantities; while novices work "backwards" starting with an equation involving the desired quantity (Larkin, McDermott, Simon and Simon, 1980; Larkin, 1980a). For example, the computer-implemented problem solver, ABLE, in its more ABLE and less ABLE forms, produces problem solutions with principles applied in an order characteristic of more and less experienced human solvers (Larkin, 1980a, 1981). Table 1 shows the order in which ABLE would apply principles to solve the toboggan problems.

Table 1. Order of Principles Applied in Typical Expert and Novice Solutions to the Toboggan Problem

NOVICE	EXPERT	NOVICE	EXPERT
$U_i = \underline{mg}\mathit{h}$	$F_g = \underline{mg}$	$K_f = \underline{\tfrac{1}{2}mv^2}$	$U_f = \underline{mgy}$
$E_i = U_i$	$F = F_g$	$a = \underline{v}^2/R$	$E_f = K_f + U_f$
$E_i = E_f$	$F = \underline{ma}$	$F = \underline{ma}$	$E_i = E_f$
$E_f = U_f + K_f$	$a = \underline{v}^2/R$	$F = F_g$	$E_i = U_i$
$U_f = \underline{mgy}$	$K_f = \underline{\tfrac{1}{2}mv^2}$	$F_g = \underline{mg}$	$U_i = \underline{mg}\mathit{h}$

In the novice solution, typically the first principle invoked would be one involving the desired quantity $\underline{h}$ (italicized in Table 1). Since this equation introduces an unknown quantity U_i (the initial potential energy of the toboggan), U_i becomes a desired quantity and cues the access of some principle involving it, here $E_i = U_i$, the total energy of a system equals the sum of its individual energies. The remainder of the principles in the novice column are produced by this search procedure. Underlining indicates a currently desired quantity. This strategy is a general means-ends strategy applied to a mathematical

representation of the problem. Each equation is compared with the desired final state of the problem (an equation involving $\underline{h}$ and only known quantities). Deviations from this desired final state are reduced by substituting for them.

The typical expert order for using principles is very different. Principles are invoked when they can be used to find a new quantity. For example, in Table 1 since $\underline{m}$ and the gravitational constant $\underline{g}$ are known, the principle $F_g = \underline{mg}$ can be used to find the gravitational force, i.e., express it in terms of known quantities. Once F_g is known, $F = F_g$ "finds" the value of the total force F. Underlining indicates newly found quantities. Similarly throughout the solution process, experienced solvers seem to work forwards in what is called elsewhere a knowledge development strategy (Larkin, McDermott, Simon and Simon, 1980; Larkin, 1980a; Simon and Simon, 1978).

These results have seemed puzzling because the backward strategy seems more intelligently directed toward the goal of finding the desired quantity. However, this is the case only if one imagines the problem being solved in its mathematical representation. Then, the only cue as to what would be a useful equation is whether or not that equation would contribute directly to finding the desired quantity. In that representation, the means-end strategy observed in novices is probably the best strategy available.

In contrast, if the problem is solved in the scientific representation (Figure 3), the first principles accessed are those that use entities in the naive representation to construct entities in the scientific representation; for example, $F_g = \underline{mg}$ and $U_i = \underline{mgh}$. These are relations that produce entities central to the scientific representation (forces and energies) from entities central in the naive representation (masses and heights). Then, major principles act to relate quantities that are known or desired in the scientific representation.

Detailed study of novice problem solutions (Larkin, 1980a, 1981) also reveal that they experience unusual difficulty with the principle: The total force on a system is equal to the sum of the individual forces on it. Indeed, in 22 solutions produced by 11 novice solvers for two moderately difficult problems in mechanics, this was the only principle consistently misapplied. It is suspected that this difficulty arises because, in the naive representation used by novice solvers, forces do not appear explicitly. Thus, they are easily omitted (overlooked) or misinterpreted.

Further evidence for the importance of a scientific representation in the problem-solving processes of experts comes from two detailed studies of experts solving more difficult problems. In one study, six subjects, graduate students or professors in physics, solved a single difficult problem in mechanics (Larkin, 1977). Protocols from these subjects were analyzed considering only the initial section, before any equation was written, presumably before the subject began working with a mathematical representation of the problem. With the exception of one subject (who became thoroughly confused and never solved the problem), each subject considered one or more possible scientific representations.

For example, one subject considered thinking about the forces in the problem or considered thinking about the virtual work in the problem. If a subject discarded a possible representation, he or she did so without ever writing an equation. If the subject used a selected scientific representation to write a corresponding mathematical equation, she or he was then always successful in using that representation, and the associated equation, to solve the problem correctly. Two subjects considered just one scientific representation. Within seconds after reading the problem these subjects made a statement like, "I know how to do this one, it's virtual work." These subjects thus recognized from the immediate naive problem representation (or perhaps even from the work strings) what scientific representation would be appropriate. But note that they identified a scientific representation (virtual work), and not a mathematical representation (an equation).

A second detailed study involved a single expert subject solving five moderately difficult mechanics problems (Larkin, 1980b; McDermott and Larkin, 1978). The aim was to develop a computer-implemented, problem-solving system that would implement steps analogous to those seen in the protocols of the expert subject. This particular subject was very articulate and complete in his descriptions. The protocols are clearly divided into sections that represent reasoning about the naive representation (blocks and strings); sections involving forces (scientific representations); and sections involving mathematics. Thus, in building a system to produce solutions analogous to this required the investigators to consider explicitly these different kinds of representations and to build in capabilities for constructing and using them.

A final set of data that fits with the hypothesis stated above comes from Clement and his colleagues (Clement, Lochhead, and Soloway, 1979). They find that a large number (47%) of entering engineering students do not solve the following problem correctly:

> Write an equation using the variables S and P to represent the following statement: "There are 6 times as many students as professors at this university. Use S for the number of students and P for the number of professors."

A hint, "Be careful: Some students put a number in a wrong place in the equation," had no significant effect, increasing the percentage of correct solutions by only 3% to 5%.

The student solvers apparently do not have difficulty with mathematical representations; over 90% of them can reliably solve simple equations. And many clearly indicate that their naive representation of the problem is correct, making statements like, "There's six times as many students, which means it's six students to one professor...."

What causes these errors could be translating directly from a correct naive representation, six students to one professor, to an incorrect mathematical representation, $6S = P$. What may be lacking is an intermediate (scientific) representation of how the number of students is related to the number of professors in the situation described. Clement and his colleagues address this issue by asking

students to use an active operation to solve the problem, specifically to write a BASIC computer program to find the number of professors from the number of students. Writing the program encourages construction of an intermediate (scientific) representation reflected by students statements; for example, "If you want to even out the number of students to the number of professors, you'd have to have six times as many professors."

In summary, problem solving in general, and particularly in physics, involves first understanding the problem. The model of understanding presented in this paper is that the solver (human or machine) constructs an internal representation of the problem. These representations depend crucially on the knowledge the solver already has. In understanding scientific problems, it has been argued that two different representations play crucial roles. First, there is the naive representation that would be constructed by individuals without special training in the science. Then, there is the scientific representation that includes entities known only to people with training in the science.

The kind of representation used is crucial to solving problems, because it determines the kind of operators that can act. If a solver works with a naive representation of a physics problem, a representation that contains no forces, he or she is exceedingly unlikely to consider initially using force principles and may be prone to errors in finding forces. On the other hand, if a scientific representation involving forces has been constructed, then the use of these principles becomes straightforward.

It has been argued, on the basis of a variety of evidence, that a central difference between expert and novice problem solvers is the extent to which they can use these specialized scientific representations. Relevant data include how experts and novices categorize problems, the order in which experts and novices access and use principles, the effect of teaching novices a scientific representation scheme, and detailed studies of the work of expert solvers on moderately difficult problems.

The remainder of this paper turns to the more general issue of the connection between the applied science of practical education and the basic science of modern cognitive psychology.

Science Instruction

The field of science instruction, like cognitive psychology, has come a long way in the last ten years. First, there is an increasing use of at least primitive theories of how learning relevant to science occurs. Examples of such theories include the learning hierarchies of Gagne (1970), the framework for meaningful learning provided by Ausubel (1968), and the rich theory of growth and change of Piaget (Inhelder and Piaget, 1958). A second advance is the development of instructional formats that can serve as vehicles for theory-based instruction.

Examples include programmed instruction (both printed and computer implemented,) the personalized system of instruction, or Keller plan (Green, 1971), as well as a few efforts to design instruction based closely on educational theory (Novak, 1979, 1980; Markle, 1969).

Thus, as cognitive psychology has been developing theories that are increasingly relevant to science instruction, the field of science instruction itself has become both more theory-based and better versed in instruction that is sufficiently systematic that it can be related tightly to theory. The time is approaching when theoretical advances in the basic science of cognitive psychology will provide profoundly useful input into science instruction; and, concurrently, the field of science instruction will provide an ideal applied test bed for these theories.

However, currently there is still an enormous gap between the detailed limited theory-based tests performed in cognitive psychology and the broad important tasks addressed by science educators. In the long run it is believed that this gap will be bridged with instruction that is tightly connected to theories, well tested in the laboratory, much as the design of a bridge is tightly connected to the well-tested theories of mechanics.

But tight connection between good theory and useful applications is probably at least five years in the future. For the present, the most useful work involves jumping this gap in the following ways. On the one hand, there are cognitive psychologists who study, with increasing care, important educational problems. They dissect out what seem to be the essential elements of educational problems and use these problems as the subject of their basic research. These individuals are very directly involved in building a cognitive theory relevant to instruction. On the other hand, there are science educators who work in a complementary way, becoming very well informed about cognitive psychology and using current theoretical advances as suggestions for how to design good instruction, although these suggestions are very far from any complete prescription. The following section summarizes some of this work.

Related Work

There is no way to be comprehensive in this review, and its inclusions and exclusions reflect the interests and biases of the investigators. It is hoped that the following comments will provide some readers with a summary that does justice to the fascinating and vital nature of this research and will give other readers the means to begin to explore it more seriously The area of problem solving and of learning to solve problems is one of particular interest in current cognitive psychology. Greeno (1976a, 1976b) works in geometry and children's word problems. Anderson (Anderson, et al., 1980) is producing a detailed model of learning from practice in geometry. Brown and Van Lehn (1980) have a fascinating model of what can go wrong in children's understanding of

addition and subtraction algorithms, and recently Resnick (1979) has been using these insights in developing instructional patterns to remedy these difficulties.

Artificial intelligence (AI) is a branch of computer science concerned with developing computer software that is "intelligent" in various senses. Many of the issues addressed there overlap with the interests of cognitive psychologists and probably also with those of science educators, although the technical nature of this literature makes it difficult to read. Some of the most promising work has been done in problem solving in physics by four separate individuals.

Novak (1977) is modeling the development of problem representation discussed earlier in this paper. His program, ISAAC, can read physics problems in natural English and construct abstracted problem representations consisting of standard wheels, levers, etc., which can then be used to generate equations to solve the problem. De Kleer (1977) in two different areas provides an exceedingly disciplined and thoughtful analysis of the kind of thinking that is required when an expert makes qualitative inferences about difficult situations. In other papers, he addresses an area of mechanics and the more complex world of electronics (de Kleer, 1975, 1977, 1979). Bundy (Bundy, et al., 1979) makes use of some of the ideas of de Kleer and also standard objects similar to those of Novak--ideas which he imbeds in a logical processor based on predicate calculus. Luger (1979) has matched some of the output of Bundy's system against the work of human solvers. Bundy (1975) has also worked in the domain of algebra, and Lewis (1979) has used this framework in trying to understand the kinds of errors made by human subjects of algebra.

There is also a considerable body of science educators who interact strongly with cognitive psychologists and utilize their theories extensively. Lochhead and his colleagues (Lochhead, 1978) are working to elucidate the natural knowledge of physics that students bring to a physics course. They, and also Champagne and Klopfer (Champagne, Klopfer, and Cahn, 1980), find that these natural theories persist throughout physics instruction and often strongly influence what individuals learn. McDermott, with her colleagues (Trowbridge and McDermott, 1980a, 1980b), does similar work on more detailed questions of exactly what aspects of physical concepts, like velocity and acceleration, are understood by beginning university students.

Reif at the University of California, Berkeley, in collaboration with several students, has done several pieces of intriguing work. With Eylon (Eylon and Reif, 1979), he addresses the issue of how organization in text affects what students retain from text and what uses they can make of that text. With St. John (Reif and St. John, 1979) he formulated explicit goals for the physics laboratory and designed a full set of instructional materials, incorporating such issues as estimation, planning of the laboratory, and relating of measurement to basic theoretical issues. With Larkin and Brackett (Reif, Larkin, and Brackett, 1976) he addressed how students can be taught a general skill for reading scientific text, so that without special help from the unassembled text they can reliably retain such important information as the meaning

of symbols introduced, typical magnitudes for these symbols, and scaling properties involved.

In later work, Larkin and Reif (1979) describe an experimental effort to teach novice problem solvers in physics some of the strategies informally observed in experts. Reif (1980) continued this work using both strategies observed in experts and strategies systematically designed to be as optimal as possible. In related work, Schoenfeld (1979) designed and taught a set of heuristics for problem solving in university-level mathematics.

Related to the area of AI is what is sometimes called "intelligent" computer-assisted instruction (ICAI). Workers in this area are using state-of-the-art, artificial intelligence techniques to construct sophisticated computer-based teaching systems. An example is the work of Goldstein (1976), in which a learner plays a game requiring logical inferencing and is aided by a coach that not only understands the game but also has a model of the cognitive and affective state of the learner. Understandably, much of this work is supported by computer companies who hope to be able to market inexpensive but sophisticated teaching devices in the near future. Examples are the work of Lewis (1980) at IBM on teaching checking techniques in algebra and that of Miller (1980) at Texas Instruments.

Developmental psychology is the study of how children grow and change. This area is of potential interest not only to those concerned with instruction of children, but for all who are trying to understand intellectual growth and change. Over several years, Siegler (Richards and Siegler, 1979) has studied children's developing understanding of real-world relations involving proportions. Klahr and others (Klahr and Siegler, 1977; Klahr and Robinson, 1980) have documented children's ability to solve simple problems.

Summary

The goal in this paper has been to give a flavor of the growing work in information-processing psychology that addresses the fundamental educational questions of how people meaningfully learn and know. This has been done first by describing current work in physics. This work begins to elucidate the central role of sophisticated understanding in skilled problem solving.

The issues discussed here illustrate a growing body of work in cognitive psychology that has the potential for impact on science instruction. Although we are still far from psychological theories that would tightly guide the design of instruction, the time is fully ripe for working toward such theories and for beginning to apply such theories as there are. As has been suggested in the review of relevant studies, this work can proceed from both sides of the gap—with

cognitive psychologists studying in detail key aspects of instructional problems and with instructional designers becoming creative interpreters and elaborators of the theories that now exist.

References

Anderson, J. R., Greeno, J. G., Kline, P. L., and Neves, D. M. Learning to plan in geometry. In J. Anderson (ed.), Cognitive skills and their acquisition. Hillsdale, N. J.: Lawrence Erlbaum, 1981.

Ausubel, D. P. Educational psychology: A cognitive view. New York: Holt, Rinehart, and Winston, 1968.

Brown, J. S. and VanLehn, K. Repair theory: A generative theory of 'Bugs.' Technical Report, Xerox Palo Alto Science Center, June 1980.

Bundy, A., Byrd, L., Luger, G., Mellish, C. and Palmer, M. Solving mechanics problems using metalevel inference. Proceedings of the Sixth International Joint Conference on Artificial Intelligence, August 1979, 2, 1017-1027.

Bundy, A. Analysing mathematical proofs, or reading between the lines. Research Report 2, University of Edinburgh, Department of Artificial Intelligence, May 1975.

Champagne, A., Klopfer L., and Cahn, L. Interactions of students' knowledge with their comprehension and design of science experiments. Technical Report, University of Pittsburgh Learning Research and Development Center, 1980.

Chi, M.T.H., Feltovich, P. J., and Glaser, R. Representation of physics knowledge by experts and novices. Technical Report 2, University of Pittsburgh, Learning Research and Development Center, 1980.

Clement, J., Lochhead, J., and Soloway, E. Translating between symbol systems: Isolating a common difficulty in solving algebra word problems. Technical Report, Cognitive Development Project, Department of Physics and Astronomy, The University of Massachusetts, Amherst, 1979.

de Kleer, J. Qualitative and quantitative knowledge in classical mechanics. Master's thesis, Massachusetts Institute of Techology, December, 1975.

de Kleer, J. Multiple representations of knowledge in a mechanics problem solver. International Joint Conference on artificial Intelligence, 1977, 5, 299-304.

de Kleer, J. Causal and technological reasoning in circuit recognition. Unpublished doctoral dissertation, Massachusetts Institute of Technology, September, 1979.

Eylon, B. and Reif, F. Effects of internal knowledge organization on task performance. Technical Report, Group in Science and Mathematics Education, University of California, Berkeley, 1979.

Fiske, S. T. and Linville, P. W. What does the schema concept buy us? Personality and Social Psychology Bulletin, 1981, in press.

Gagne, R. The conditions of learning. Second edition. New York: Holt, Rinehart, and Winston, 1970.

Goldstein, I. An athletic paradigm for intellectual education. Technical Report, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1976. Logo Memo 37, AI Memo 389.

Green, B. A. Teaching physics by the Keller plan at MIT. American Journal of Physics, 1971, 39, 764-770.

Greeno, J. G. Indefinite goals in well-structured problems. Psychological Review, 1976, 83, 479-491.

Greeno, J. G. Cognitive objectives of instruction: Theory of knowledge for solving problems and answering questions. In D. Klahr (ed.), Cognition and instruction. Hillsdale, N. J.: Lawrence Erlbaum, 1976.

Inhelder, B. and Piaget, J. The growth of logical thinking, from childhood to adolescence. New York: Basic Books, 1958.

Klahr, D. and Robinson, M. Formal assessment of problem-solving and planning processes in preschool children. Cognitive Psychology, 1981, 13(17), 113-148.

Klahr, D. and Siegler, R. S. The representation of children's knowledge. In H. Reese and L. P. Lipsitt (eds.), Advances in child development. New York: Academic Press, 1977.

Larkin, J. H. and Reif, F. Understanding and teaching problem-solving in physics. European Journal of Science Education, 1979, 1(2), 191-203.

Larkin, J. H., McDermott, J., Simon, D. P., and Simon, H. A. Models of competence in solving physics problems. Cognitive Science, 1980, 4, 317-345.

Larkin, J. H. Problem solving in physics. Technical Report, Group in Science and Mathematics Education, University of California, Berkeley, 1977.

Larkin, J. H. Enriching formal knowledge: A model for learning to solve problems in physics. In J. Anderson (ed.), Cognitive skills and their acquisition. Hillsdale, N. J.: Lawrence Erlbaum, 1981.

Larkin, J. H. Skilled problem solving in physics: A hierarchical planning approach. Journal of Structural Learning, 1980.

Larkin, J. H. The cognition of learning physics. American Journal of Physics, 1981, 1, in press.

Lewis, C. Organization of skill in algebra. Paper delivered at the workshop on cognitive processes in algebra, University of Pittsburgh, August, 1979.

Lewis, C. Skill in algebra. In J. Anderson (ed.), Cognitive skills and their acquisition. Hillsdale, N. J.: Lawrence Erlbaum, 1981.

Lochhead, J. Teaching students how to learn. In Proceedings of the Third International Conference on Improving University Teaching, University of Maryland Press, 1978.

Luger, G. F. A comparison of human and automatic problem solving in two areas of mechanics. Technical Report, Department of Artificial Intelligence, University of Edinburgh, 1979.

Markle, S. M. Good frames and bad: A grammar of frame writing. New York: John Wiley and Sons, 1969.

McDermott, J. and Larkin, J. H. Rerepresenting textbook physics problems, pages 156-164. The Canadian Society for Computational Studies of Intelligence, 1978. University of Toronto press.

Miller, M. L. Using plans to generate hints in an intelligent tutoring system for elementary graphics programming. Technical Report, Intelligent Interactive Systems Branch, Computer Science Laboratory, Central Research Laboratories, Texas Instruments Incorporated, 1980.

Novak, G. S. and Araya, A. A. Research on expert problem solving in physics. Technical Report NL-40, The University of Texas at Austin, Department of Computer Sciences, 1980.

Novak, G. S. Computer understanding of physics problems stated in natural language. Unpublished doctoral dissertation, University of Texas at Austin, March, 1976.

Novak, G. S. Representations of knowledge in a program for solving physics problems. International Joint Conference on Artificial Intelligence, 1977, 5, 286-291.

Novak, J. D. The reception learning paradigm. Journal of Research in Science Teaching, 1979, 16, 481-488.

Novak, J. D. Learning theory applied to the biology classroom. American Biology Teacher, 1980, 42, 280-285.

Reif, F. and St. John, M. Teaching physicists' thinking skills in the laboratory. American Journal of Physics, 1979, 47(11), 950-957.

Reif, F., Larkin, J. H. and Brackett, G. C. Teaching general learning and problem-solving skills. American Journal of Physics, 1976, 44, 212-217.

Reif, F. Teaching problem solving and other cognitive skills: A scientific approach. The Physics Teacher, 1981, 19(5), 310-316.

Resnick, L. B. Syntax and semantics in learning to subtract. Paper presented at the Wisconsin Conference on the Initial Learning of Addition and Subtraction Skills, Racine, Wisconsin, November 26-29, 1979.

Richards, D. D. and Siegler, R. S. The development of time, speed, and distance concepts. 1979 Meeting of the Society for Research in Child Development.

Rumelhart D. E. Notes on a schema for stories. In D. C. Bobrow and A. Collins (eds.), Representation and understanding. New York: Academic Press, 1975.

Schoenfeld, A. Can heuristics be taught? In J. Lochhead and J. Clement (eds.), Cognitive process instruction. Philadelphia, Pa.: The Franklin Institute, 1979.

Shank, R. and Abelson, R. Scripts, plans, goals, and understanding. Hillsdale, N. J.: Lawrence Erlbaum, 1977.

Simon, D. P. and Simon, H. A. Individual differences in solving physics problems. In R. Siegler (ed.), Children's thinking: What develops? Hillsdale, N. J.: Lawrence Erlbaum, 1978.

Simon, H. A. The information-storage system called 'human memory.' In H. Simon (ed.), Models of thought. New Haven, Conn.: Yale University Press, 1979.

Trowbridge, D. E. and McDermott, L. C. An investigation of student understanding of the concept of velocity in one dimension. American Journal of Physics, 1980, 48(12), 1020-1028.

Trowbridge, D. E. and McDermott, L. C. An investigation of student understanding of the concept of acceleration in one dimension. American Journal of Physics, 1981, 49(3), 242-253.

Chapter 9

INVESTIGATING SCIENCE CLASSROOMS

The third symposium of the conference was based on the recognition that education, at least in schools, is a social enterprise. It occurs within a particular set of social contexts: classrooms, schools, school districts, states, as well as a context that includes federal policies, teacher education institutions, and culture in general. All of these contexts impact on teacher-student interactions in the classroom.

Concern with what is "really going on" has again led to observational studies in intact classrooms. The new approaches to field studies can avoid the limitations of the action research of the 1930s and 1940s by maintaining consistency with a theoretic base. Several promising theoretic bases are being advocated as appropriate for such studies: ethnographic, sociological, psychological, and ecological. Each has advantages and disadvantages in application to studies of the complexity of classrooms. But recurrent questions, such as, "What is going on in this science (mathematics) classroom?" need to be addressed. Theory-based field studies, to be explored in the three chapters that follow, hold greater promise in resolving such questions than did the empirical studies of the past.

Bloom (1980) has suggested that educational researchers should investigate the "alterable variables," variables that can be changed through intervention to cause an increase in achievement. Time on task and cognitive entry are two such alterable variables. Both have been identified through careful studies in classrooms. Bloom argues that research to identify such alterable variables can bring about profound changes in schools.

Ethnographic studies in classrooms based on a "constructivist," rather than a "behaviorist," theoretical framework were advocated by Magoon (1977). The "constructed worlds" of both teachers and students, the formulation of descriptions and/or explanations of objects and events, were defended as essential objects of study for understanding educational processes. Such studies have been accomplished successfully using ethnographic methodologies (Cicourel, et al., 1973; Lortie, 1975; Rist, 1975); yet, ethnographic studies of science classrooms are rare.

Bronfenbrenner (1976) has called into question the prevalent approaches to educational research and has proposed a new perspective in method, theory, and substance. He refers to this perspective as the experimental ecology of education. Bronfenbrenner contrasts his

ecological model, which conceptualizes environments and relationships in terms of systems, with educational research that deals with person-environment and environment-environment relations by treating variables separately.

The three papers presented in the following chapters include both quantitative and qualitative approaches to investigating science and mathematics classrooms. Two of the three papers report studies undertaken from an ethnographic frame of reference.

Jane Stallings' major interests have been in studying teaching and learning in elementary school classrooms. Her studies have been concerned with the improvement of instruction. She finds that little research on teaching has been conducted in science classrooms. In Chapter 10, Stallings points out the need for studies that include attention to curricular content and the quality of teachers' planning and execution of their intentions.

In her studies of secondary mathematics instruction, Stallings found marked differences between instructional patterns in general mathematics classes and those in geometry classes. Using the "Classroom Snapshot," an objective observation system, and observations of teacher-student interactions, Stallings found that teachers and students ~~used~~ their time quite differently in these two kinds of classes. Stallings demonstrates that classroom research such as she had done can be used to help teachers see what they are doing in class and that such information can affect instruction in positive ways.

Jane Stallings recently formed a new organization, the Teaching and Learning Institute in Mountain View, California. Her undergraduate studies were in elementary education and the natural sciences. She received her doctorate in child development and elementary education from Stanford University. Among her publications are Learning to Look: A Handbook on Classroom Observation and Teaching Methods and "Allocated Academic Time Revisited or Beyond Time on Task."

One school of thought about how best to study the classroom supports unobtrusive observation as a major method of study. Perry Lanier's study of general mathematics and algebra classes relies on this procedure. His investigations are supplemented by psychometric data about all students and clinical interviews of a selected group of students.

Lanier raises an issue of practical, in contrast to theoretic, research. He argues that research in mathematics and science education has lost touch with the classroom and ignores the knowledge of practitioners. Although he does not define what he means by theoretic research, it becomes evident that nearly all research studies lacking classroom observations of one type or another fall into this category. His major plea is for balance in research approaches, feeling strongly that current research practices, especially in mathematics education, is

out of balance, that the theoretic is heavily overrepresented in research studies.

Perry Lanier is Professor of Teacher/Mathematics Education in the College of Education and the Institute for Research on Teaching at Michigan State University. His undergraduate work was in business administration and mathematics. He earned an M.A.T. in Mathematics Education, an M. Ed. in School Administration, and a doctorate in elementary education, the latter from the University of Oklahoma. Professor Lanier began his career in education in the elementary school, teaching in the middle grades, and later teaching science and mathematics in the upper grades. His major research interest is the teaching of mathematics at the junior high school level. Among his publications are Animated Mathematics and a chapter on mathematics in Modern Elementary Curricula. A report of the research methodologies he used in a comparison study of general mathematics and algebra classes is presented in Chapter 11.

In contrast to Lanier, Jack Easley presents a case for the participant-observer as a valuable approach to research in science and mathematics classrooms. The plethora of approaches to field or case studies in classrooms has led Easley to initiate the development of a cognitive theory of school science and mathematics. In Chapter 12, he sketches the major elements of such a theory, contrasting what he characterizes as the normal scientist's approach as "outsider" with what he terms a cognitive ethnography, the "insider" approach. Easley suggests that human cognition be considered within an ethnographic context in which the thinking, both immediate and reflective, of teacher and students is taken into account as part of the cognitive background that a participant-observer uses in observing and interpreting classroom events and their consequences.

Jack A. Easley is Professor of Teacher Education at the University of Illinois, Urbana-Champaign. Before completing his doctoral studies in Science Education at Harvard University, he taught physical sciences and mathematics courses in high schools, junior colleges, and at the university level. His B. S. degree was in Physics and his M. Ed. in Science Education. His continuing interest in cognition and instruction led him to spend a year as Visiting Scholar at the International Center for Genetic Epistemology at the University of Geneva. Among his numerous publications are (with Robert Stake) Case Studies in Science Education, and (with J. M. Gallagher) Knowledge Development, Volume 2, Piaget and Education.

Studies in science and mathematics classrooms represent a very small portion of the studies conducted in the last 20 years of research in science and mathematics education. The difficulties of conducting such studies are apparent in the three chapters concerned with this problem. Methodological problems continue to confront both research and research consumer in disentangling the classroom ecosystem from the observer and all that he or she brings to such an investigation.

Careful classroom studies can make unique contributions to the teaching and learning of science and mathematics. These contributions are of a different genre and add new components to understanding the subject. In another decade, ethnographic or ecological approaches to science and mathematics instruction should be contributing a much larger portion of research studies than they have in the past.

References

Bloom, B. S. The new direction in education research: Alterable variables. *Phi Delta Kappan*, 1980, *61*(6), 382-385.

Bronfenbrenner, U. The experimental ecology of education. *Teachers College Record*, 1976, *78*, 157-204.

Cicourel, A. V. *Cognitive sociology*. London: Macmillan, 1973.

Lortie, D. C. *School teacher: A sociological study*. Chicago: University of Chicago Press, 1975.

Magoon, A. J. Constructivist approaches in educational research. *Review of Educational Research*, 1977, *47*(4), 651-693.

Rist, R. C. Ethnographic techniques and the urban school. *Urban Education*, 1975, *10*, 86-108.

Chapter 10

CLASSROOM RESEARCH: IMPLICATIONS FOR MATHEMATICS AND SCIENCE INSTRUCTION

Jane Stallings, Teaching and Learning Institute

During the 1970s, most of the research on teaching focused on effective instruction in compensatory reading and mathematics classes. The theory was that if the basic skills of low-income students were improved, their chances to gain meaningful employment would also improve. Because such a large proportion of funds available to education were allocated to categorical aid programs, most of the research and evaluation funds were directed toward identifying effective instructional strategies for low-achieving students. Thus, it is important to keep in mind that reviewing the findings from the research on teaching in the 1970s may not generalize to high-achieving students.

Research on Teaching Basic Mathematics and Reading

One of the most potentially useful variables to emerge from the research on teaching during this era was student time on task. Many educators are now convinced that if student time on task is increased, an increase in student achievement will follow. This belief is based on a considerable amount of research that focused on the length of school days, actual scheduled class time, time allocated to academic subjects, and student engaged time.

Length of School Day

The length of a school day in elementary schools or the length of a class period in secondary schools defines the maximum amount of time available for instruction. Within that maximum amount of time is the actual amount of time that is allocated to reading and mathematics instruction. Several studies by Stallings, et al., at SRI International indicate that the length of the school day or the length of class period is not related to student achievement in reading and mathematics. First-grade classrooms in the National Follow Through Observation Study (Stallings, 1975) varied as much as one hour and 30 minutes in length of school day; secondary class periods for remedial reading varied from 40

to 55 minutes (Stallings, Needels, and Stayrook, 1979). Such differences are not associated with reading or mathematics gain. These findings indicate that student learning depends on how the time that is available is used.

Academic Time

Although length of school day was not a predictor of student gain in these studies, the amount of time spent in academic subjects has been found to be positively associated with gain in mathematics and reading. In the Follow Through Observation Study (1975), time spent in mathematics, reading, and academic verbal interactions was related to achievement. Time spent working with textbooks and workbooks (as opposed to time spent with puzzles, games, and toys) was related to achievement in reading and mathematics. Time spent in small groups (as opposed to one-to-one instruction) was also associated with student academic gain. Conversely, time spent in more exploratory activities was positively related to scores on a nonverbal problem-solving test and to a lower student absence rate. Similar relationships were also found in a study of California third grade Early Childhood Education classes (Stallings, et al., 1977). These findings indicate that what teachers teach and how time is spent make a difference in what students learn.

Research conducted at Far West Laboratory as part of the Beginning Teacher Evaluation Study (BTES) (Fisher, et al., 1977) found associations between academic learning time and student achievement. Powell and Dishaw (1980), reporting data from the BTES, showed that the time allocated to reading and mathematics for second graders ranged from 62 to 123 minutes per day. This time for fifth graders ranged from 49 to 105 minutes per day. The correlation of allocated learning time with achievement varied from one sub-test to another.

Time on Task

The amount of time students actually spend on task is, of course, related to student gain. Teachers can allocate time in their schedule for mathematics or reading, but if students socialize and misbehave, gain will not be achieved. Powell and Dishaw, in the BTES study, reported that the engaged time of second-grade students varied from 38 to 98 minutes, and that of fifth-grade students varied from 49 to 105 minutes. Student-engaged time was positively associated with student achievement in all tests and at both grade levels.

Stallings, et al., (1979) in a study of secondary classrooms found a negative relationship between student gain and interruptions such as the loud speaker or tardy students. Each time one of those disruptive events occurred, students' eyes left their work and it took time to get back on task. In recent reviews and studies, there has been a consistency in the relationship between time on task and student accuracy

rates (Berlinger, et al., 1976; Brophy, 1979; Good, 1980; Rosenshine, 1977).

Allocation of Time to Specific Activities

The body of knowledge emanating from the research on teaching in the 1970s suggests that teachers should allocate more time to academic subjects, keeping in mind ability levels, and students should be kept engaged in the tasks. Such a recommendation will confirm what most teachers and administrators already know. However, the recommendation is not very helpful unless more specific statements are made about how to engage students and how to use academic time. It is of interest to know what activities occur within a class period, and how the time for those activities is distributed. If a class period has 45 minutes, how long does it take the teacher to get the show on the road? What is the balance of silent reading, written assignments, recitation/discussion, and instruction? Does the distribution of class time across activities make a difference in student achievement? If so, does this difference vary among students with different achievement levels?

A two-phase study conducted in 87 secondary remedial classrooms found that the allocation of time to specific reading activities significantly affects student reading gain; further, this distribution of time affects students in specific reading levels in different ways. This study, Teaching of Basic Reading Skills in Secondary Schools (TBRSSS), was funded by the National Institute of Education (NIE) and carried out at SRI International by Stallings, Needels, and Stayrook (1979).

Focus of Instruction

To whom should teachers focus their instruction: individuals, small groups, the total group? During the last decade, considerable energy has been directed toward the development of individualized programs. Federal, state, and local funds have been spent to develop programmed reading, mathematics, and science books. All of these programmed materials were aimed at providing children with activities in which they could progress at their own rates. It was assumed that if students were working at their own pace through a series of sequential exercises, learning would occur. It did for some students, but not for others. In general, there has been a great disillusionment in individualized instruction. Some students learn best when new information is presented to a small group of students who are operating at a similar pace (Stallings, 1975; Stallings, Needels, and Stayrook, 1979). Learning occurs when students read aloud, hear others read aloud, and hear others ask questions and respond. Hearing and speaking as well as reading and writing help students integrate and retain information. Individualized programs based almost totally on workbooks do not allow for this type of group learning.

At a conference sponsored by the National Institute of Education in 1978, where the failure of individualized instruction was discussed with teachers, teachers reported that they felt relegated to being record keepers. Workbooks were relied on to provide instruction for students. It appears that students need the interaction of teachers. A teacher can develop concepts with a group and can change examples or illustrations to coincide with the group's background experience. If students do not understand, the teacher can find yet another example. Books or machines do not do that. Books or machines provide opportunities to practice and reinforce what teachers are teaching, but research suggests (Stallings, 1975) they are not sufficient to provide the instruction that students need.

Interactive Instruction

In the study of time allocated to specific activities (TBRSSS), it became clear that teachers who were interactive in their instructional style had students who achieved more in reading. An interactive style included providing oral instruction for new work, discussing and reviewing student work, providing drill and practice, providing acknowledgement for correct responses, and supportive corrective feedback for incorrect responses. The correlation between interactive instructional style and student learning was replicated in Phase II of the study.

In summarizing the research on active instruction, Brophy (1979) states: "Learning gains are most impressive in classrooms in which students receive a great deal of interaction with the teacher, especially in public lessons and recitations that are briskly paced but conducted at a difficulty level that allows consistent success."

A Comparison of Instructional Styles in Secondary, General, and Advanced Mathematics Classes

Mathematics instruction was studied in 11 San Francisco Bay urban and suburban high schools (Stallings and Robertson, 1979). The sample included 91 classrooms sorted into three types:

Type I included general mathematics and pre-algebra, required for high school graduation.

Type II included algebra I and geometry, required for entry into the California university system.

Type III included trigonometry, algebra II, and calculus, required for entry into some university science and mathematics programs.

Each of the 91 classrooms was observed for three class periods using a low-inference objective observation system. The allocation of time to activities, the focus of instruction, and the teachers' instructional

styles were recorded. Instructional style findings were based on approximately 900 recorded interactions.

The manner in which mathematics instruction differs in general mathematics and geometry classes can be studied with these data. The SRI observers informally reported that general mathematics classes appeared to be much less task-oriented, and the teachers seemed less involved in teaching the students. To see whether or not this was true, we compared the kinds of activities occurring in the Types I, II, and III classrooms.

Data from the observation of activities (see Table 1) indicate that Types II and III teachers offer more instruction (S5) and discussion (S6) than do the Type I teachers. They also spend less time doing classroom management tasks during the class period, such as grading papers or keeping records.

Table 1. Percentage of Occurrence of Activity Variables for Three Types of Mathematics Classes

VARIABLES		All Classes Type I $\bar{x}$	All Classes Type II $\bar{x}$	All Classes Type III $\bar{x}$
S1	Teacher class management, no students	23.222	20.260	15.175
S4	Making assignments	0.929	1.039	2.077
S5	Instruction	14.895	25.353	30.512
S6	Discussion/review	8.091	20.800	23.216
S8	Written assignments	33.755	14.807	11.122
S9	Test taking	3.865	2.147	2.054
S11	Social interaction	10.610	12.672	13.797
S12	Student uninvolved	11.022	6.252	3.471
S13	Discipline	0.[illegible]62	0.200	0.040

*Type I = General Mathematics, Pre-algebra
Type II = Algebra I, Geometry
Type III = Algebra II, Trigonometry, Calculus

Type I teachers have students doing written assignments (S8) three times as often as do Type III teachers. Students in Type I classes were observed to be not involved in any activity (S12) more than three times as often as Type III students. These data indicate that Type I students do not receive as much interactive instruction (S5, S6). They are given written assignments (S8) more often, while the teacher does class management tasks such as grading papers, making lesson plans, or keeping records (S1). In an SRI study of secondary reading classes, we found a negative relationship between this type of noninteractive instruction

and student achievement. On the other hand, in that same study we found a strong positive correlation, .63, between interactive instruction (discussion, S6) and student achievement. This type of instruction is seen in Type II and Type III classes.

A corroboration of this difference in teaching in Type I classes and in Types II and III classes was found when we compared the teaching behavior of 11 teachers in our sample who taught both a Type I class and a Type II or III class. Six of these teachers were observed in Type I and Type II classes, and five were observed in Type II and Type III classes. These individual teachers' classsroom processes were investigated to determine if differences existed between the teaching processes for the same teacher in the different types of mathematics classes.

Type I Versus Type II Classes

An analysis of teachers' profiles indicated that all six of the teachers who taught both Type I and Type II classes tended to assign more written work to the students in their Type I classes. Five of these teachers spent more class time doing management tasks and not interacting with their Type I students. While all of the teachers spent more time with their Type II students in questioning and feedback activities, more interactions regarding mathematics were observed in the Type II classes for all six teachers.

These data indicated that the same teacher structured her or his class differently for different levels of mathematics classes. In the Type I classes, students worked alone doing written assignments more and the teacher devoted more class time to classroom management activities, such as correcting papers and organizing materials. Because of this type of classroom structure, the same teachers showed less interactive instruction with general mathematics students than with algebra I or geometry students. Time spent in instructing was almost twice as much for the algebra I and geometry classes as for the Type I classes. Type I students evidenced more misbehavior and less time on task. The main difference seemed to be that teachers tended to be noninteractive with Type I students, and with Type II students they were interactive.

Type II Versus Type III Classes

For the five teachers who taught both Type II and Type III classes, no clear differences were found in the methods used in these classes. Teachers varied on whether an observed teaching process occurred more in the Type II or Type III classes. The teachers in this group interacted more with the entire class when instructing their Type III classes. In the Type II classes, the teachers tended to interact more with groups or individuals. These teachers were also observed conducting more class discussion in their Type III classes. This might indicate a trend for teachers to conduct more interactive discussion with students in their higher level classes and to involve the entire class in the discussion.

Discussion

These data reveal findings similar to the analysis involving all classes in the study. Clear distinctions can be found in the teaching processes for the same teacher in the Type I and Type II classes. However, less difference was shown in the manner in which the same teacher conducted Type II and Type III classes. The same teacher tended to interact less with students in the Type I classes than with those in Type II classes. Type I students were assigned more independent desk work while the teacher used class time to correct students' work or organize materials. Nine of the eleven teachers led more class discussion in their higher level class.

The implication of this finding is that students in general mathematics classes may not receive the teacher attention and instruction required to achieve well and continue in mathematics studies.

Differential Treatment of Men and Women in Mathematics Classes

A hypothesis proposed was that differential treatment of men and women students would be observed in mathematics classrooms and this will be related to enrollment in advanced mathematics classes. Observation data were used to study this hypothesis.

Observation

Twenty-two geometry classes were the focus of this analysis. Approximately 50% of the students in these classes were women. Therefore, the number of teacher interactions with men and women should be about equal, if equal treatment was being given. Table 2, on page 142, displays the average frequency of observed interactions of teachers with male and female students during a geometry class. These computations were based on the total of 22 geometry classes in the study.

Though few of the differences are statistically significant, the trend is rather clear. Men were spoken to more often than were women (F10, F11). Men asked more questions (F15, F16) and teachers asked men more questions (F18, F19). Women volunteered answers as often as did men (F34, F35) but the men were called on to respond more frequently than were women (F40, F41). Men received a little more individual instruction (F43, F44) and social interaction (F54, F55). Acknowledgement (F62, F63), praise (F65, F66), encouragement (F80, F81), and corrective feedback (F100, F101) were given slightly more frequently to men than to women.

The only variable that occurred more frequently with women was positive interactions (F104), i.e., teachers smiled or laughed more with women than they did with men.

Table 2. Mean Frequencies of Interaction Variables for Geometry Classes

VARIABLES	ALL CLASSES			
	FEMALE		MALE	
All verbal, student-teacher interaction/math	F10	38.69	F11	46.88
Student question/math	F15	5.60	F16	6.59
Teacher question/math	F18	3.72	F19	4.60
Volunteer responses to teacher/math	F34	1.23	F35	1.67
Teacher response/student question/math	F40	5.49	F41	6.32
Teacher instruction/math	F43	5.48	F44	6.56
Teacher-student social interaction	F54	0.97	F55	1.41
Teacher acknowledge student/math	F62	2.87	F63	3.59
Teacher praise student/math	F65	0.26	F66	0.34
Teacher encourage student/math	F80	0.23	F81	0.26
Teacher-student interaction/ relevance of math	F88	0.00	F89	0.07
All corrective feedback/math	F100	8.03	F101	10.13
Positive teacher-student interaction	F104	0.41	F105	0.40
Negative teacher-student interaction	F111	0.06	F112	0.14

Discussion

Overall, the data in Table 2 support the interpretation that men students were treated somewhat differently than were women students in geometry classes. This is in agreement with Bean's finding (1976) that showed teachers initiating more contacts with men students than with women students and with Dweck and Reppucci (1973) who suggest that mathematics teachers provide more feedback to men. Although we found such differences, they did not relate to the enrollment of women in advanced mathematics classes. On a questionnaire, women stated they were more likely to enroll in classes where the instruction provided was actively interactive regardless of whether the interaction was with men or women.

Research on Teaching in Science Classrooms

Are the instructional strategies that were found to be effective in reading and mathematics classes similar for science classes? Very little research on teaching has been conducted in science classrooms. We can hypothesize that interactive instruction would be effective, and that students staying engaged with their tasks would enhance their learning.

Whereas drill and practice of basic skills in reading and mathematics were helpful to the low achiever, an inquiry approach might be more appropriate for teaching science to low achievers. How should time be distributed across activities? How much time is spent for demonstrations or experiments and how much time for discussion, review, or written assignments? How are these variables related to student achievement in science? To answer such questions, we have developed a comprehensive observation system to be used in science classrooms.

The observation system is composed of two sections. The first, the Classroom Snapshot, is so named because it records the environment and the participants in the classroom as if they were being photographed at one instant. It records every person in the classroom in the activity in which they are engaged and shows with whom they are engaged. The distribution of adults and students among the activities that are occurring simultaneously are recorded as the observer places them on the grid, going clockwise around the room. Essentially, the Snapshot provides data to assess the activities occurring, the materials being used, grouping patterns, teacher and adult participation, and students in activities independent of adults.

As shown in Figure 1, the Snapshot, on page 144, the classroom activities are listed down the left side of the page, and the materials are listed across the top. The observer records the information in each appropriate circle, recording each unique grouping occurring in the classroom. A completed Snapshot documents the number and kind of groupings, the activity and materials of each group, and whether an adult is present.

The letters at the beginning of each row indicate the placement of each category of participants in the classroom: T = Teacher, A = Aide, O = Other adult, I = Independent student. If it is a team teaching arrangement, both teachers are shown in the activity with the student or students they are working with when the Snapshot is recorded. The 1, S, L, and E in the bubbles in the row relate to the number of students with whom the teacher is working: 1 = one student, S = 2-8 students, L = 9 to one less than the total group, E = Everyone.

The bubbles marked in Figure 1 indicate that the teacher is giving instruction to one student in a workbook. The other students are working alone in workbooks. An aide is doing some classroom management task. Five grids are completed each 45-minute period and each class period is observed three days in a row.

SCIENCE

CLASSROOM SNAPSHOT

MATERIAL → ACTIVITY ↓		01 Textbook	02 Workbook Worksheet	03 Test	04 Materials/ Models to Look at	05 Machine	06 Chalkboard	07 Materials to Use/ Manipulate	08 No Material
01 Reading Silently	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
02 Reading Aloud	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
03 Making Assignments	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
04 Instruction/ Explanation	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
05 Discussion/ Reviewing Assignments	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
06 Demonstration	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
07 Written Assignments	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
08 Taking Test Quiz	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
09 Laboratory Work	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
10 Social Interaction	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
11 Student Uninvolved	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
12 Being Disciplined	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
13 Classroom Management	T	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	A	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	O	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E
	I	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E	I S L E

Figure 1. The Snapshot, a Comprehensive Classroom Observation System

Each Snapshot is followed by a five-minute interaction observation that is focused on the teacher. Approximately 300 interactions are recorded per class period. The nature of the instruction can be assessed. It is theoretical? Are there demonstrations? Are the questions thought provoking? Is the corrective feedback guiding and supportive?

The advantage of objective observation systems such as this are several. First, observers are trained to a criterion during a seven-day session and are checked regularly for acceptable reliability (>.80). Thus the analyst can have confidence in the data. Second, these data are easy to use in testing relationships between classroom instructional strategies and student outcomes. Third, teachers are quite amenable to being observed with such a system since it is a straight count made by a person carefully trained to be objective. It is not someone's subjective opinion of what is occurring. The data can be used for staff development as well as for program evaluation.

The limitations of the objective observation are that the actual content of the curriculum cannot be described nor can the quality of the teacher's planning and execution be assessed. Some combination of quantitative and qualitative data would be useful to describe effective science teaching.

The system described above has been used in classes where physics, health, biology, and general science were being taught. The interaction patterns were similar to those reported for the mathematics classes. In the more advanced classes there were demonstrations and discussions. The instruction was interactive. In the general science and health classes, the students were working in workbooks and the teachers were monitoring, not instructing. Research in classrooms strongly suggests that students are more likely to learn in classrooms where the instruction is interactive. This sample is too small to make any statements about science instruction. A study is needed where a representative sample can be observed, criterion or normative test data can be gathered, and student attitudes toward science can be assessed. Effective instructional strategies need to be identified for each grade level and science subject.

Translating Research on Teaching to Teacher Training Programs

From observational research, teachers can be given specific feedback that shows clearly how they distribute time across activities and how they interact with students. Figure 2, on page 146, shows the type of profile that was presented to teachers of reading in secondary classrooms. Similar profiles could be generated for science teachers. The variables used in a quasiexperiment of Teaching Basic Reading Skills in Secondary Schools (Stallings, Needels, and Stayrook, 1979) have considerable face validity that makes the findings understandable to teachers. The fact that the findings were generated from classes similar to the ones in which the teachers were working lent credibility

TEACHER NAME: SAM JONES
LAKEWOOD HIGH SCHOOL
WASHINGTON UNIFIED SCHOOL DISTRICT

STANDARD DEVIATIONS FROM THE MEAN

MEAN FOR ALL CLASSES

SNAPSHOT VARIABLES	RECOMMENDATION	Less Frequent ← → More Frequent (<2 S.D. · 1 S.D. · 0 · 1 S.D. · 2 S.D.>)	ALL CLASSES	SAM'S CLASS PRE	SAM'S CLASS POST
Teacher class manage/no students	Less	O ← X	28.	46.	07.
Total silent reading	Less	O ← X	08.	09.	05.
Total reading aloud	More	X → O	09.	06.	29.
Total making assignments	OK	O ← X	06.	04.	00.
Total instruction	More	O → X	10.	18.	13
Total discussion	More	X → O	02.	04.	18
Total practice drill	More	XO	01.	00.	00.
Total written assignments	OK	X→O	27.	17.	21.
Total test taking	More	XO	02.	00.	00.
Total social interaction	Less	O ← X	04.	15.	07.
Total student uninvolved	Less	O ← X	05.	15.	01.
Total discipline	OK	XO	00.3	00.	00.
INTERACTION VARIABLES					
Teacher to individual student	More	X → O	89.	48.	113.
Teacher to group	More	X → O	13.	01.	38.
Teacher to class	OK	O ← X	44.	137.	37.
Teacher direct question, reading	More	X → O	41.	23.	50.
Student response, reading	More	X → O	42.	19.	43.
Teacher praise, support, reading	More	X → O	13	08.	34.
Teacher corrective feedback	More	X → O	13	19.	53.
All interactions/reading	More	X → O	208.	230.	304.
All interactions/behavior	OK	X → O	06	01.	11.
All interactions/positive	More	X → O	04	00.	04.
All interactions/negative	OK	XO	00.	00.	00.
Student comments, assignment	OK	X→O	09.	01.	03.
All interactions/class assignment	Less	O ← X	69.	144.	17.
Teacher manage class, no student	Less	O ← X	33.	73.	05.

X = Pre-training observation
O = Post-training observation
→ = Direction of change

Figure 2. Profile of Sam Jones' Pre- and Post-Training Observation

to the research. The variables used in the study were very specific and translating them into recommendations for teachers was not a difficult task. Each teacher received her or his own set of recommendations for behavior change based on three days of observation in a class of her or his choice.

For example, we observed Sam Jones' Period 3 class prior to a series of in-service workshops. He received the behavior profile shown in Figure 2. Sam Jones was spending 46% of the class time in management tasks (see pre-test score for the first variable.) This indicates that Sam was spending approximately one half of the class time not being involved with students, e.g., grading papers or keeping records. The mean for all teachers on this variable was 28%. After interpreting the study findings to Sam, we made the recommendations shown in the left column. Our recommendations were to provide more instruction, more discussion, more feedback, and to do less paper grading and record keeping during the class time.

A series of six workshops, which were very supportive and interactive, focused on how to manage classroom time, how to offer activities for different ability levels, and how to interact and be supportive to students. Following these workshops, Sam Jones was observed to be doing only 7% management activities (see post-test score). His style of instruction became much more interactive. The profile shows how time spent in activities changed from one point to another. Interestingly, Sam had been teaching for 35 years. He found the changes made his teaching more fun, and the students became more involved. Overall, the teachers in this quasiexperiment did change behaviors as recommended, and their students gained more in reading than did a comparison group of students in classrooms of teachers who did not attend the workshop.

Where to from Here?

For a few years following the announcement of Russia's Sputnik, policymakers and educators focused their attention on improving science and mathematics education. A number of experimental science programs were well funded until the end of the sixties. During the 1960s, the civil rights movement was making the general populace painfully aware that low-income children were not thriving in our schools. In response to the need for improvement in the life chances of this segment of society, vast sums of money were channeled into educational programs serving the economically and educationally handicapped. The primary focus of these programs was to improve basic reading, writing, and computation skills. In the early seventies, as funding became limited and the effectiveness of the programs was not clear, funds for exploratory mathematics and science programs were drastically reduced or eliminated.

Having witnessed in the 1970s the plunge of the college-bound students' test scores, many educators and policymakers are now feeling compelled to reinstate excellency in our high schools. In California, there is a new focus on the average and above-average student. The great effort to assist the low-achievement and handicapped students has sometimes led to neglect of the middle and upper segments of our student bodies. Higher standards for graduation are being considered for these students.

The state of North Carolina is trying to find a remedy for working with these gifted and talented students. They have recently opened a School of Science and Mathematics. It is a free residential public high school that offers a pre-collegiate curriculum to 150 bright, technology-oriented eleventh-grade students. The students, 30 of whom are black, include sons and daughters of farmers, construction workers, professors, and business people. They were selected out of 900 applicants and were identified as those who have the potential for making contributions to the fields of science and mathematics. Though the school has been attacked as being elitist, Governor James B. Hunt, Jr. (1980) has backed it because, "I'm concerned about the loss in productivity in American industry and the loss of our competitive edge in the

whole technological field. At the high school level, we simply are not doing the best job we can do."

Clearly, it is time to focus again on excellence in instruction for average and above-average students. It is important to keep in mind that all students in our public schools need to be taught so that they can enjoy and contribute to the fullest extent that they are able. This means that no single segment of society should receive the lion's share of the resources.

At this time in the history of education, we must consider for now and for the future what kinds of citzens the school system should be developing. What kinds of people are required to solve the problems of food, clothing, and shelter in a world of expanding population and dwindling resources. What kinds of technological skills are required to function productively in this society? What kinds of basic and human skills are required to make democracy work?

The pendulum of educational emphasis swings back and forth. It is important to find a middle position that will allow all students to thrive and fulfill their potential. Classroom teachers and teachers of teachers have a special trust. Research conducted in the 1970s indicates very clearly that what is happening in classrooms now makes a difference in how our children will grow and develop in the 1980s.

References

Bean, J. What's happening in mathematics and science classrooms, student teacher interactions. Paper presented at the meeting of the American Educational Research Association, San Francisco, California, 1976.

Berlinger, J., Jacobsen, K., McKenna, B. H., Stake, R. E., and Yashinsky, J. A study of testing practices in the Royal Oak (Michigan) public schools. Michigan: Royal Oak Michigan School District, 1976.

Brophy, J. Teacher praise: A functional analysis. Paper presented to the American Psychological Association, 1979.

Dweck, C., and Repucci, N. D. Learned helplessness and reinforcement responsibility in children. Journal of Personality and Social Psychology, 1973, 25, 109-116.

Fisher, F., et al. Beginning teacher evaluation study. Technical Note IV-16. Far West Laboratory, San Francisco, California, 1977.

Good, T. L. Classroom expectations: Teacher pupils interactions. The social psychology of school learning. New York: Academic Press, 1980.

Hunt, J. B. Boarding school for brains. Time Magazine, 1980, 116(17), 87.

Powell, M. and Dishaw, M. A realistic picture of reading instructional time. Reading Research Quarterly, 1981, in press.

Rosenshine, B. Instruction principles in direct instruction. Champaign: University of Illinois, 1977.

Stallings, J. Implementation and child effects of teaching practices in follow through classrooms. Monographs of the Society for Research in Child Development, 1975, 40(7-8), 50-93.

Stallings, J., Cory, R., Fairweather, J., and Needels, M. Early childhood education classroom evaluation. SRI International, Menlo Park, California, November, 1977.

Stallings, J., Needels, M., and Stayrook, N. How to change the process of teaching basic reading skills in secondary schools, Phase II and Phase III. SRI International, Menlo Park, California, 1979.

Stallings, J., and Robertson, A. Factors influencing women's decisions to enroll in advanced mathematics courses. SRI International, Menlo Park, California, May, 1979.

Chapter 11

MATHEMATICS CLASSROOM INQUIRY: THE NEED, A METHOD, AND THE PROMISE

Perry E. Lanier, Michigan State University

Overview

In 1978 my colleagues and I began participant observation studies in 12 junior high school mathematics classrooms. The classrooms of primary interest were ninth-grade general mathematics classes, but for comparative/contrastive purposes we also looked in algebra classes, remedial classes, a combined seventh-, eighth-, ninth-grade class, four eighth-grade classes taught by the same teacher, and a ninth-grade social studies class that included a number of students from an observed general mathematics class.

Our studies were driven by the consensus among secondary mathematics teachers and chairpersons or supervisors that ninth-grade general mathematics classes are unrewarding experiences for both teachers and students. Furthermore, the consequences significantly affected students' curricular decisions throughout high school and constrained their opportunities upon graduation. That is, if they elected to enter college, their range of choice among "majors" was severely limited given their weak mathematics preparation. Similarly, if they elected to enter the job market, their weak mathematical abilities eliminated them from competition for the more desirable positions. Our research objectives were to accurately and narratively portray life in general mathematics classrooms with an eye toward improvement, but, to date, the focus has been on capturing "what is" in contrast to "what can be."

In ascertaining what life for teachers and students is like in these classrooms, we used field research methodology. Basically, this is an adaptation and extension of ethnography, the method of the anthropologist, for purposes of studying educational settings. Field notes from classroom observations were the primary data source, but these were augmented by teacher interview data, student artifacts such as tests, and limited use of videotaping. The participant observers (data gatherers) were teachers or mathematics educators trained to conduct field research and educational anthropologists. As participant observers, we attempted not to intervene in the flow of instructional activity. We departed from this stance only when it was expedient to interact with a student or students who asked for assistance or otherwise initiated a dialogue with us. Our task was to be as unobtrusive as

possible; we made every effort to study the classroom in its naturalness and wholeness.

These investigations were conducted under the auspices of the Institute for Research on Teaching (IRT) at Michigan State University. Funding and support for the studies came from the National Institute of Education (NIE), M.S.U.'s College of Education, and the National Science Foundation. The NSF study was designed not only to observe in classrooms, but also to gather psychometric data on all students and clinical interview data on selected students in those classrooms.

This paper is based on this series of investigations in junior high mathematics classrooms. Major points of the paper are: a discussion of the need for mathematics classroom studies, including an argument for who should/could be conducting such studies; a consideration of the criteria prospective classroom researchers should employ in selecting "study sites"; an explication of the nature of field research in educational settings; and a presentation of some preliminary findings along with the implications for improving practice and for further research.

The Need for Practical Investigation in Mathematics Classrooms

In 1970 Schwab stated, "My own incomplete investigations convince me that we have not the faintest reliable knowledge of...what actually goes on in science classrooms."

Five years ago (1975) the National Advisory Committee on Mathematics Education (NACOME) Conference Board of the Mathematical Sciences reported, "The question, what goes on in the ordinary classroom in the United States? is surely an important one, but attempting to survey the status of mathematical education at 'benchmark 1975,' one is immediately confronted by the fact that a major gap in existing data occurs here. Appallingly little is known about teaching in any large fraction of U. S. classrooms."

A review of the 580 entries appearing in the tenth annual listing of research on mathematics education by Suydam and Weaver (1980) showed that 25 studies, slightly more than 4%, were conducted to address questions of classroom practice.

In one sense, this last piece of information is encouraging: That there are nearly 600 people studying some aspect of mathematics education in a given year is commendable. Yet one wonders about the apparent imbalance when the need for practical/action research has been noted by scholars, teachers, and study groups for at least five or ten years. Only 25 of the 580 studies were directed toward investigating the quality and nature of life in mathematics classrooms. The remainder can be categorized as being primarily concerned with the theoretic.

Since the focus of this paper is the investigation of science classrooms, a discussion concerning the plausible explanations of why so few studies can be classified as practical, and likewise so many as theoretic, will be foregone except for the following observation: The first volume of the new NCTM Professional Reference Series, Research in Mathematics Education (Johnson, 1980), contains two sections, "The Research Process" and "The Research Problems." In the "Process" section there is one chapter titled, "Types of Research." Within that chapter under the heading of "Case Study" is the following single statement (note parentheses), which possibly reflects the perceived importance, or lack thereof, of holistic classroom investigation:

> The case study has recently come to include some aspects of instruction (to a small group or even an entire class) with follow-up interrogation or observation of selected individuals (p. 23).

Clearly, the mathematics education research community by its actions and writings is not unanimously convinced the classroom is a promising arena for investigation. The field of science education, as represented by mathematics education, is in need of classroom research. This need is supported by research experience in the study of general mathematics classrooms. Such research can provide unique insights that will lead to the improvement of science teaching and science teacher education, as well as advance the science education research field.

Why Classroom Research?

The intent here is to portray a need for balance between research of the theoretic and research of the practical in the field of science education. Begle and Gibb (1980) cast the idea this way in a discussion of "New Directions" for research in mathematics education: "Research has developed competing explanations for partial views of learning...but generalizations from these partial theories are limited...they are insufficient for the planning and realization of classroom practices. It is necessary to consider not only the student or the curriculum but also the general context of learning and the teacher's role in effecting learning" (p. 15).

Davis (1977) attributed the following to David Hawkins: "One of the most important--and most neglected--aspects of science (is) the analysis and explication of practitioner knowledge." Davis continued: "Practitioners do know something, and a major stage in science occurs when theorists study practitioner knowledge and try to relate it to theoretical modes of thought. This stage has been by-passed by education. Teaching is studied with the implicit assumption that teachers do not know what they are doing and theorists must come in and tell them."

To point up further the need for practical research and convey more clearly what is and has been missing in the array of research activity in education, Davis asserted that:

Educational research in the United States has become administration-centered. It deals with the kinds of decisions that administrators make, the kinds of data that administrators have, the selection among options that administrators have available, and the kinds of outcomes about which administrators can be aware. This is very different from the data, decisions, and options that confront a teacher..... In educational research we have lost the ability to think in terms of individual cases. There has been a harmful symbiosis of administrative knowledge types, statistical methodology, and seemingly scientific generalizations. What is best for most people may not be best for you. A good teacher tries to make these individual adjustments, but research methodology tries to cast them aside (p. 31).

The argument Davis makes for the need in completing the cycle of educational research activity rests primarily on the notion that without knowing what the practitioner knows we have only a partial picture. However, there is an implicit point that seems to me to be of equal or greater significance. It is the notion of the practitioner as a user of research. That is, being in possession of knowledge about practice may indeed make our science more complete, but for what purpose?

Is not one of the goals of research to inform and improve practice? Do we not know or suspect that teachers are influenced by other teachers more than any other variable? Does it not seem reasonable that research on classroom practices might appear more relevant to teachers, and be more likely to be perused and used by them, than the results of theoretic research?

Tom (1980), in an argument for a conception that portrays teaching as a "moral craft" rather than an "applied science" suggested both a need for and a use of research knowledge from the classroom. He observed:

Despite the obvious differences in pedagogical knowledge and skill between the experienced teacher and the typical novice, the craftsperson teacher rarely attempts to pass systematically this accumulated wisdom to the next generation...teacher training programs contain little such codified knowledge and skills, and many professors...deny that such craft culture is valuable. Even experienced teachers often deny that their skills and knowledge... could be of value to other teachers.... In other words, all teachers must discover, "what works for them individually"--of matching strategies and ideas to one's unique classroom of youngsters. The result of not receiving craft culture in preservice training--except perhaps in student teaching--and of believing that all teachers must develop a personal teaching style is the conception of teaching as an individualistic enterprise that must be learned by trial and error (p. 320).

Certainly, each of us who has either taught or closely observed others teaching or learning to teach is familiar with the preponderance of learning by trial and error. If one outcome of classroom research was the reduction of an overdependence on trial and error, it seems

certain that practice would be improved. Further, it is conceivable that perusal and use of practical research would begin the extinction of antiresearch attitudes common among teachers and would subsequently generate an appreciation of theoretic research as well. Such a state is as as desirable for teachers as for researchers, for it is the theoretic that provides the <u>new</u> idea to be adapted by the practitioner.

What do we want to find out from classroom research? What do we want to become smarter about? Why conduct classroom research? Schwab (1970) contends:

> What is wanted is a totally new and extensive pattern of <u>empirical</u> study of classroom action and reaction: a study, not as a basis for theoretical concerns about the nature of teaching or learning process, but as a basis for beginning to know what we are doing, what we are not doing, and to what effect; what changes are needed, which needed changes can be instituted with what costs or economics, and how they can be effected with minimum tearing of the remaining fabric of educational effort (p. 313).

Clearly, Schwab is asking that we establish "what is" as objectively as possible, then follow that assessment with evaluative judgments of "what should and could be." His point on "minimum tearing of the remaining fabric of educational effort" brings to mind the change strategy of Postman and Weingartner in their book, <u>The Soft Revolution</u> (1971). Could Schwab have been saying to us that, had we considered these things, the curriculum reform movements of the post-Sputnik era would have been implemented differently with different consequences? For example, had we thought in terms of a "soft revolution" versus a "revolution in school mathematics," what would we have needed to know? What would we have done differently? This information of conditions and their implications for change can only be validly obtained from serious and systematic investigation of classrooms. Furthermore, the implementation of any innovation is perilously endangered without such information.

Bridgham (1977) alludes to the concern for knowing existing conditions in making an appeal for the research on teaching mathematics to consider the thought of teachers. He says:

> We have not yet found secure ways of arranging communication between the designers of curriculum materials and classroom teachers which acknowledges the strength of each, provides for effective collaboration in materials design, and which doesn't require that all the adapting and conciliating be done by one of the partners. One of the reasons this problem remains is that we lack good descriptions of teachers' curricular thoughts. What is it that teachers think about? How do they decide what to do, how to do it, what to hope for? What do they take into account? (p. 78)

In summary, the need for classroom research appears to be twofold. First, a knowledge picture of teaching is incomplete without the classroom consideration. Second, the knowledge derived and communicated from classroom research is likely to have explicit and useful things to say to the practitioner (a phenomenon that may subsequently endear the teacher to research). Inherent in the twofold need for practical knowledge and its use are the questions that should be addressed in the context of the classroom: What's happening in the prelude to, during the flow of, and following instruction? What are the teacher's thoughts and actions relative to the classroom experience? What are the learner's thoughts and actions relative to the experience? How is the content of instruction enacted?

Who Should Engage in Science Classroom Research?

A second important factor of the general problem of looking in science classrooms is the consideration, "Who should do the looking?" Should a portion of the 160 authors who reported the outcomes of theoretic inquiries be redirected toward practical studies? Perhaps! Should a significant number (there were 408 reported in 1979) of doctoral dissertation studies be focused on the practical? I think so! Clearly, these two modifications in the pattern of current activity would create an improved balance between the theoretic and the practical. There is yet another source of talent that can and should be tapped to engage in classroom research. This source is the set of professors and supervisors who work with teachers--preservice and inservice--in science education activities.

At Michigan State University, there are at least a half-dozen professors in mathematics education, another six in social science, and even more in the natural sciences whose primary professional activity is working with teachers. These people are capable, hard-working professionals who engage in teaching, an occasional development project (teacher education curriculum or school curriculum) and writing (expository articles, texts for teachers or school learners), but they seldom engage in research studies. Their primary concern is with practice. Supervisors of school science, mathematics, or social studies are another professional group concerned with practice.

This pool of professionals concerned with the practical might provide a resource for conducting classroom investigations. They are on the scene, interested in the problem, and are knowledgeable about classrooms. Perhaps the primary obstacles to implementing field studies by these professionals are the acquisition of field research knowledge and skills, and negotiating with superiors for time to conduct such investigations. Each of these obstacles is significant but worthy of consideration, given Davis' argument for the need to complete the educational research cycle. If 10% of these science teacher educators conducted and reported one classroom study per year to the annual <u>JRME</u> report and its counterpart in natural and social science education, there would be a more optimal balance between practical and theoretic inquiry.

There are at least two good reasons why this suggestion is plausible. First, these are people whose professional experience typically includes several years in the classroom prior to their teacher education work in higher education; therefore, they are in a unique position relative to issues of theory and practice. Additionally, there is no other set of professionals who, on the one hand, has such a unique opportunity for linking research and practice, while, on the other hand, has the unique responsibility to do so.

The matter of responsibility is the second reason for the suggestion to recruit science teacher educators for classroom research. Though science educators are not antiresearch, they frequently communicate two things that may contribute to the development of antiresearch attitudes in teachers. First, by not being engaged in research, the science educator is indirectly diminishing its significance as a useful activity in the domain of science teaching. Second, and potentially more damaging, is the inclination to have answers to, but not questions about, problems of practice. This stance may tend to convey teaching as being an activity that is based on dogma rather than science, which seems to be one consequence of the by-passed stage pointed out by Davis.

Looking to science teacher educators to augment the present level of research activity in science classrooms is based on the opportunity to link research and practice and to extinguish a practice that appears to be detrimental to a healthy research attitude in the profession.

Are All Mathematics Classrooms Equally Worthy of Investigation?

Given the need for research in mathematics classrooms, one is confronted with the question of classroom selection, i.e., a research site. Intuitively, it seems obvious that every potential research site is not as good as every other potential site. But that intuition generates the question of criteria of selection. To address this question, the evolution of the General Mathematics Project in the Institute for Research in Teaching will be related. How ninth-grade general mathematics came to be the focus of the inquiry will be explained. This will be followed by a discussion of guidelines for choosing classroom research sites. The purpose of relating the evolution of the General Mathematics Project is twofold. First, the selection of a research site is a complex matter that requires serious thought prior to a decision. Second, the problems that had to be resolved provide an example of how science teacher educators can begin to investigate systematically their own work with teachers in classrooms.

General Mathematics, a Strategic Research Site

Ninth-grade general mathematics as a research site emerged from three distinct events. The first of these events was the release, distribution, and subsequent deliberations of the 1975 NACOME report. Though the report contained six chapters, two of them, "Patterns of Instruction" and "Teacher Education" were especially upsetting for many professionals in mathematics education. It was in the "Patterns of Instruction" chapter that the jarring observation was made that appallingly little is known about what goes on in mathematics classrooms across the nation. The impact of the observation is intensified when one reflects on the level of effort exerted during the sixties on curriculum development and teacher development, though limited primarily to secondary education, and the level of effort, almost nil, exerted on ascertaining what happens to students in classrooms with these curricular materials and teachers. Having some achievement data on students in these classrooms provided little solace to such an obvious gap in our knowledge base.

The significance of the gap was given additional impact by the reports of teachers in the "Teacher Education" chapter that their most significant problems were those of dealing with motivation, laboratory learning, slow learners, learning styles of students, etc. "Lowest on the list are content topics" (p. 92). For many in mathematics education, these problems were ones that could be finessed by suggesting or implying that they could be dealt with by focusing on interesting and neat mathematical content. However, any distillation or interpretation of what was known and unknown suggested the need for classroom inquiry--but, which classrooms, and by what means? In retrospect, the NACOME report was a most influential precursory event in the evolution of the General Mathematics Project.

The second event leading to general mathematics as a research site was the creation, in 1976, by the National Institute of Education (NIE) of the Institute for Research on Teaching (IRT) with Michigan State University's College of Education. Retrospectively, this event had as much influence on how to look as it did on where to look. The IRT focus was to be on the study of teacher thought, but teacher thought in terms of learner, curriculum, and setting. Reading was the curricular area specified, with the option of considering other subject matter areas. In 1977 the IRT sponsored an invitational conference to consider research on teaching mathematics, which resulted in a proceedings publication by that title (Davis, 1977).

In addition to the mathematics conference, the Institute also held a conference on field research methodology in 1977. Subsequently, the Institute recruited a field researcher, and the College of Education began offering a field research seminar sequence designed to train personnel to conduct research in educational settings. As a result of the two conferences, I was prompted to take the sequence of seminars on field research. A practicum component in this training led me to observe in a general mathematics classroom taught by my colleague in mathematics education, Bruce Mitchell.

Mitchell's presence in the general mathematics classroom is the third key event in the evolution of the General Mathematics Project. In a revision of the undergraduate secondary mathematics methods class, Mitchell had negotiated with a local school district to teach a geometry class on a daily, year-long basis. He wanted the students to have a weekly field experience in his classroom as part of the methods course. Over this year (1976-77) he discovered that the regular teachers, in their informal exchanges in the hall, lounge, etc., frequently expressed concerns about their respective general mathematics classes. For two reasons, his own enlightment and his methods students' enlightment, he arranged to teach a ninth-grade general mathematics class in the same high school during 1977-78. This was the site selected for satisfying the practicum requirement of my field research training.

During the course of the year, two inquiries were prompted by my observations in Mitchell's class and furthered the emergence of the General Mathematics Project. The first of these was a question of clarification to mathematics supervisor, Charles Zoet. In a presentation at the annual University of Michigan mathematics education conference, Zoet asserted that he and his Livonia, Michigan, secondary teachers were not reaching half their students. His response revealed that these students were similar in many ways to those in Mitchell's general mathematics class. Following this, informal telephone polls of several mathematics supervisors across Michigan were made with the finding that at least half the students did take general mathematics. Furthermore, it was not only a class that generated disquietness and concern among the teachers, but it was equally disquieting and disliked by students.

Hence, general mathematics was clearly problematic for students, teachers, and supervisors. It was a problem that could only be significantly addressed by study of the problem where it existed--in the thought, actions, and consequences of and for the teachers and students in the general mathematics classroom. It was indeed a setting worthy of study, a strategic research site.

Guidelines for Selecting Classroom Research Sites

Shulman (1978) in an invited address to the AERA special interest group for Research in Mathematics Education argued that the <u>strategic research site</u>, as a concept, was a useful guide to educational researchers. His argument attempted to discern those features or qualities that appear to distinguish strategic research sites from other potential loci for empirical investigation. Shulman proposed that striking discontinuity, aberration, anomaly, or error can serve as a strategic research site for studies of human functioning in general and mathematics education in particular.

The discontinuity of relative satisfaction by mathematics teachers and supervisors tended to appear as consideration was given to their assessments of the lowest algebra class versus the general mathematics class. They were not always satisfied and happy with the algebra

classes, but there was a noticeable shift from positive to negative when considering general mathematics. Bruce Mitchell stated, "I just can't be me in the general mathematics class," a discontinuity in terms of his normal and expected style of teaching. Similarly, there was a discontinuity in student demeanor and attitudes as expressed by Mitchell, "Last year you'd walk down the hall and hear kids say in a positive manner, 'I'm in Mitchell's geometry.' You sure don't hear anyone saying anything about being in Mitchell's general math class."

An additional guideline in selecting a classroom research site is that of complexity, i.e., it should be rich enough to warrant being looked at from several perspectives, rather than being a relatively barren, single-issue phenomenon. The general mathematics class, for instance, represents multiple problems: learner problems--computation, reasoning, or reading deficiencies; curricular problems--scope of content; context problems--no one likes to be there and the consequences thereof; and teacher problems--how to motivate, what to expect.

In summary, the selection of a classroom research site should reflect the inquirer's personal interest, a setting with multiple problems that need and can be addressed, and some apparent aberration or anomaly. Finally, the research site should be focused clearly on a practical problem.

Participant Observation--A Method for Classroom Inquiry

The primary method used in the General Mathematics Project was that of participant observation. Since it was virtually impossible for an adult to come across as an adolescent student in any naturalistic sense, the respective participant observers needed to establish themselves as a natural part of the scene. As members of the scene, their role was to particpate as observers.

Why become part of the landscape? First, "belonging in some sense" is indigenous to the method of participant observation. That is, the method requires interaction with members in a form that differs from the interaction likely to occur between members and a visiting fireperson, for example. McCall and Simmons (1969) implicitly affirmed the necessity for belonging in their statement of what participant observation is:

> (participant observation) is most sensibly regarded...as the ... of methods and techniques that is characteristically employed in studies of social situations or complex social organizations of all sorts. These are studies that involve repeated, genuine social interaction on the scene with the subjects themselves as a part of the data-gathering process (p. 3).

Second, or in a corollary sense, belonging is imperative if the researcher is to understand, or produce knowledge about, a phenomenon. In other words, any meaning, the purpose of a descriptive study, derived from a phenomenon has validity if, and only if, the participant observer is a bona fide member of the scene under investigation.

Diesing (1971), in a discussion of the participant observer method as used by philosophers, explained why and how the participant observer must belong:

> The participant observer approaches scientific method from the inside; he attempts to take the point of view of the scientists who are using a particular method. He does this by becoming, as far as possible, a member of a scientific community, sharing its activities and discussions, familiarizing himself with the literature, problems, and personalities that are discussed, helping with the daily work in whatever way he can. Taking an inside point of view consists of taking one's concepts, distinctions, problems, logic, values, from the scientific community rather than imposing externally derived concepts and distinctions on it. One learns concepts and distinctions not just by asking people or reading an article but by participating in innumerable activities. In this way one is able gradually to note the distinctions that are made habitually, almost unconsciously, the procedures that are carried out routinely, the goals, assumptions, and modes in inference that are taken for granted in activity as well as in speech. In the terminology of communication theorists, one learns by communicating rather than by metacommunicating.
>
> The participant observer tests the adequacy of his account by seeing whether its various parts are acceptable and intelligible to the people he is working with, though not necessarily identical to their own verbal formulations. He does this not by asking their approval of an article--which tests mainly friendship and politeness--but informal discussion continued over a period of time. Or, expressed somewhat differently, he tests the adequacy of his understanding by acting on it and seeing where his actions are unintelligible or puzzling to others....
>
> The participant observer takes the inside approach not only to understand scientific work but also to evaluate it (p. 291).

Another consideration of a participant observer study, along with the necessity to belong, is an appropriate research paradigm. Mishler (1979), in a paper on "Meaning in Context," asserts that where the "study of situated meaning is central," as it implicitly is in an observer study, an approach other than the search for general or universal laws that describe relationships, of the form $Y = f(x)$, between two variables is necessary. To illustrate his point, Mishler identified and described three approaches--phenomenology, sociolinguistics, and ethnomethodology. The common ground of these three centers is on the dissolvement of subject and context. That is, meaning is always within a context, and contexts incorporate meanings.

Since the general mathematics classroom is, by reputation and consensus, exceedingly complex, the phenomenological approach seemed most attuned to our proposed observation study. As Carini (1975) stated, "The function of observing in phenomenological inquiry is to constitute the multiple meanings of the phenomenon...." The task is therefore not to determine the single meaning of an event, but to reveal the multiplicity of meanings. Thus, phenomenology was the accepted approach for the earlier observation study of general mathematics, since it was likely that many explanations of the phenomenon would be plausible. The focus of sociolinguistics on the meaning of language in context or ethnomethodology's central topic, the apparent invariance of rules, were of less interest.

Given these theoretic underpinnings of participant observation study, what are the domains where this method is most applicable? Diesing (1971) provided a response to this question in his summary descriptive paragraph of the method:

> The participant-observer method was first developed by anthropologists, though it is also frequently used by sociologists, social psychologists, political scientists, and organization theorists. Its primary subject matter is a single, self-maintaining social system. The system may be a small community with its own culture, or a larger society with its culture, or a small and relatively isolated neighborhood, or a gang, clique, voluntary organization, or family, or a formal organization or institution, or a person (clinical method), or a historical period. In each case the emphasis is on the individuality or uniqueness of the system, its wholeness or boundness, and the ways it maintains its individuality. The primary objective is to describe the individual in its individuality, as system or rules, goals, values, techniques, defense of boundary-maintaining procedures, and decisions procedures. In one important variant, the primary interest is in recurring processes within or around such individual systems (p. 5).

Diesing's description makes the method extremely well fitted for attaining the objectives of the General Mathematics Project--identification and characterization of the manner in which the group identity, classroom organization and process, peer culture, and teacher processes interact to influence mathematics learning both cognitively and affectively; and a focus on the contrasting perspectives of teachers and students on the meanings, events, and purposes of these classes.

As a method, participant observation seems particularly well suited to the investigation of mathematics and science classrooms, if the researcher's overarching questions emanate from a practical problem. As an example, the two sets of questions of concern in the NSF study, "The Ecology of Failure in Ninth-Grade General Mathemat[illegible] An Ethnographic, Experimental, and Psychometric Inquiry," were:

1. Who are the students who become the mathematically disadvantaged ninth-grade population? What are their mathematical abilities, attitudes toward mathematics, learning histories, and learning expectations? How do they wind up in general mathematics?

2. What effects do the two primary instructional environments--general mathematics or first-year algebra classes--have on the cognitive and affective mathematical development of early adolescents? How are those settings experienced by both teachers and learners, whose interactions define those learning environments and their consequences?

If these are representative of the questions that typically characterize a practical problem of the classroom, then the form of the answers to such questions becomes a concern. Clearly, the answers will not be in the form of statistical generalizations. They are rather more likely to be in the form of retrospective generalizations. According to Stenhouse (1978) retrospective generalizations are "organization(s) of experience in retrospect...attempts to map the range of experience rather than to perceive within that range the operations of laws in the scientific sense."

Though our analysis of the data gathered to answer the above questions is incomplete, the preliminary form of the answers appears to be that of retrospective generalization. The final section of this paper includes examples of these preliminary conclusions.

In summary, participant observation, in the anthropological sense, is an appropriate method to use when investigating classrooms. The questions asked and the answers suggested by analysis of the data will be in the form of retrospective generalization rather than predictive generalization, the expected form of answers to questions in a quantitative study.

Of What Use Are Findings from Classroom Research?

The following preliminary findings of the ninth-grade general mathematics investigation are presented as an example of outcomes of classroom research where the primary method was participant observation. Perusal of these preliminary findings will be illuminating in terms of their usefulness to teachers, policymakers, and researchers.

- Tracking, the policy commonly used for placing students within ninth-grade math--whether into algebra (college track) or general mathematics (vocational track), is usually highly correlated with the students' records in mathematics classes during junior high school. A survey (Belli, 1980) of mathematics chairpersons in Michigan secondary schools revealed that, while the most commonly used criterion for tracking students was teacher recommendation, the wishes of counselors, parents, and even students themselves played an important part in determining their final placement into a math track. Yet, regardless of how

students entered their respective tracks, there is a strong tendency for them to remain there as long as they stay in school. Tracking decisions that not only affect early adolescents' present learning environment but also their senior high school or college options seem premature at best and inequitable at worst.

• Teachers instruct general math classes differently than they instruct algebra classes. Further, these differences appear to be critical factors, since they include aspects of teaching that are recognized as clearly related to student learning, e.g., the general math students receive less direct instruction, less goal clarity, less assistance with seatwork, less encouragement, less opportunity for discussion (Weisbeck field notes, 1979). Teachers expect less of their general math students than of their algebra students on such factors as completion and submission of homework as well as achievement outcomes (Buschman field notes, 1980).

• Most secondary mathematics teachers find it easier to think about, to plan for, and to teach advanced classes than to do so for general math classes. Part of this imbalance stems from their difficulty in comprehending that students in general math can have serious problems in learning basic content. Their intuition tells them that the problem is something controllable by the affected person rather than a problem calling for their professional assistance. Teachers may come to hold these convictions because (1) they have enjoyed a history of successes and relative ease in learning the subject, and (2) their teacher education programs only prepared them to teach the advanced subjects, leaving them unequipped to teach students experiencing grave difficulty in learning basic mathematical skills. Hence, we need to know more about whether or not, and in what ways, teachers come to shift their primary emphasis from one of teaching subject matter to students to one of teaching students particular subject matter.

• Most teachers assigned to teach classes with a high percentage of youngsters identified as having low promise for successful achievement in mathematics (e.g., general math classes) have unusual difficulty in teaching these classes and often feel only marginally or not at all successful. These same teachers often enjoy success--by their own assessment as well as by reputation--as teachers of algebra I or II and geometry. The following reference to general math students was taken from observer field notes (Criss, 1979): "These are really nice kids, it's just too bad I have to teach them mathematics."

• The low incidence of success and high incidence of frustration and failure encountered by both teachers and learners in general mathematics classes have created unique instructional settings that are notorious for their unpleasantness, i.e., teachers and students in general math classes have a strong aversion to being there. The class is not a place where teachers or students anticipate teaching and learning mathematics, rather it is something to endure; a place where students anticipate doing a lot of problems and at year's end concluding their formal study of mathematics. During a taped interview (Lanier, 1980) one general mathematics student responded when asked if she was planning to take a math class in high school by tearfully stating she

would never take another math class because she hated mathematics so much. She did not see why everyone thought mathematics was so important to being successful.

• Though we know that most general mathematics students have a diversity of learning problems at a crtitical level, we do not know the precise nature of these problems. Consequently, corrective instructional activities for a particular student have little chance of success, and the probability of a class strategy working is even lower.

• Well established is the troublesome and problematic nature of teaching and learning general mathematics at the ninth-grade level. Not well established are practices that alleviate the problem. Hence, new inquiries should focus on the study of teacher thought and action as it relates to changing existing patterns of instruction and identifying more successful approaches to helping youngsters who find learning mathematics difficult.

• Overall, students placed in general mathematics classes appeared to be different, in certain important ways, from students placed in algebra classes. Interestingly, at the time of placement, group differences in areas of computation, achievement, test scores, or measures of math anxiety were not pervasive; more striking were differences in:

a. The amount of variation in types of learning difficulties experienced. Students' difficulties appear highly diverse and more individualistic in a general math class.

b. The students' approaches to memory and reasoning tasks required by the math content. General math students tend to have more concept fragmentation or misconceptions of the content, which may be related to a strong propensity to rely on rote memorization of rules and facts that later are insufficient to carry them through increasingly complex mathematics experiences.

c. The students' social interactions with the teacher and other students. Teachers speak of this difference in terms of lack of maturity, the lack of social graces, etc., all of which frequently place almost total responsibility for classroom behavior on the teacher.

d. The students' study habits. General math students usually spend less time on task during class time and are erratic in terms of completing out-of-class assignments. One algebra/general math teacher states, "I tell the eighth-grade teachers not to recommend anyone for algebra who has poor study habits" (Lanier field notes, 1979).

The study of general mathematics classrooms has resulted in preliminary findings that are of a different kind than those that result from theoretic studies. They have been presented to support the argument that classroom observational studies are needed to provide a complete understanding of mathematics teaching. This kind of study is essential for developing ways to improve the teaching of science and mathematics.

The previously cited statement of Schwab (1970) expresses the point most appropriately:

> What is wanted is a totally new and extensive pattern of empirical study of classroom action and reaction; a study not as basis for theoretical concern about the nature of teaching or learning process, but as a basis for beginning to know what we are doing, what we are not doing, and to what effect; what changes are needed, which needed changes can be instituted with what costs or economics, and how they can be effected with minimum tearing of the remaining fabric of educational effort (p. 313).

The findings reported here provide information about what is and is not going on in general mathematics classrooms, and to what effect. Teachers reading these findings would become aware of some changes they could and should make. In short, these findings hold promise for improving the practice of teaching mathematics in particular, and of the practice of teaching in general.

References

Begle, E., and Gibb, E. Why do research? In R. Shumway (ed.), Research in mathematics education. National Council of Teachers of Mathematics, Professional Reference Series. Reston, Virginia: NCTM, 1980, 3-9.

Belli, G. A survey of ninth-grade mathematics. Unpublished manuscript, 1980.

Buschman, J. General mathematics project field notes. East Lansing: Institute for Research on Teaching, Michigan State University, 1980.

Bridgham, R. Reaction to R. Davis' "Curricular implications for research on teaching mathematics." In Proceedings of the Research on Teaching Mathematics Conference. East Lansing: IRT, Michigan State University, 1977.

Carini, P. F. Observation and description: An alternative methodology for the investigation of human phenomena. North Dakota study group on evaluation monograph. Grand Forks: University of North Dakota Press, 1975.

Criss, J. General mathematics project field notes. East Lansing: IRT, Michigan State University, 1979.

Davis, R. Curricular implications for research on teaching mathematics. In Proceedings of the Research on Teaching Mathematics Conference. East Lansing: IRT, Michigan State University, 1977.

Diesing, P. Patterns of discovery in the social sciences. Chicago: Aldine, 1971.

Johnson, D. Types of research. In R. Shumway (ed.), Research in mathematics education. National Council of Teachers of Mathematics, Professional Reference Series. Reston, Virginia: NCTM, 1980, 20-29.

Lanier, P. General mathematics project field notes. East Lansing: IRT, Michigan State University, 1979.

Lanier, P. General mathematics project clinical interviews. East Lansing: IRT, Michigan State University, 1980.

McCall, G. J., and Simmons, J. L. (eds.). Issues in participant observation. Reading: Addison Wesley, 1969.

Mishler, E. G. Meaning in context: Is there any other kind? Harvard Educational Review, 1979, 49, 1-19.

NACOME (National Advisory Committee on Mathematical Education). Overview and Analysis of School Mathematics, Grades K-12. Washington, D. C.: Conference Board of Mathematical Sciences, 1975.

Postman N., and Weingartner, C. The soft revolution. New York: Delacorte, 1971.

Schwab, J. The practical: Language for curriculum. Washington, D. C.: National Education Association, 1970.

Schwab, J. Science, curriculum, and liberal education: Selected Essays. I. Westbury and N. Wilkof (eds.). Chicago: University of Chicago Press, 1978.

Shulman, L. The teaching of mathematics: Strategic sites for research on teaching. Paper presented at the annual meeting of the American Educational Research Association, Toronto, 1978.

Stenhouse, L. Case study and case records: Toward a contemporary history of education. British Educational Research Journal, 1978, 4(2), 21-39.

Suydam, M., and Weaver, J. Research on mathematics education reported in 1979. Journal for Research in Mathematics Education, 1980, 11(4), 241-311.

Tom, A. Teaching as a moral craft: A metaphor for teaching and teacher education. Curriculum Inquiry, 1980, 10(3), 317-323.

Weisbeck, S. C. General mathematics project field notes. East Lansing: IRT, Michigan State University, 1979.

Zoet, C. Personal communication, February, 1978.

Chapter 12

THE NEED FOR A COGNITIVE ETHNOGRAPHY OF SCHOOL SCIENCE AND MATHEMATICS

Jack Easley, University of Illinois

Bronfenbrenner (1976), and doubtless others, called for an ecological approach to the study of educational problems. This need is clear in understanding problems of equity, violence, talent development, and other instances where we recognize that societal problems are limiting the effectiveness of the schools. Part of my own graduate training was in ecology, and I have long been impressed by the accomplishments ecologists have made in understanding large, complex ecosystems where the physiology of the individual component organisms is also very complex.

However, I have to recognize that, in education, we are dealing with thinking individuals, from the occupants of the classroom to superintendents and officers of state and federal agencies. Whether we agree with these individuals or not, their thoughts have more influence on the system than disembodied goals or objectives of the organization. From a Piagetian framework, all these people can be said to assimilate aspects of their environment to the cognitive systems they each have, and their responses are seen as generated by the interaction of these selected aspects of their environment and their cognitive systems.

Just as the ecologist has to know something about the metabolic and assimilatory systems of typical organisms in each ecological niche, so we need to know much more than we now do about the cognitive systems of pupils, teachers, and other significant groups involved in schools. However, we also need to relate this knowledge to large scale interaction patterns of the system.

Teacher Cognition in Primary School Mathematics: An Ethnography

My recent experience with Bob Stake and many others in a study called Case Studies in Science Education (1978) has convinced me that

"the teacher is the key" to any significant changes in the school system that society might require.[1] Since 1978, I have been concentrating on finding out as much as I could about teacher cognition--that is, what teachers understand their problems to be while they are functioning as teachers, how they respond to their own perceptions of various classroom situations, and what they accept as solutions. For the most part, I have chosen to work with teachers in primary school mathematics classes, since I perceive that a major problem develops in most primary math classes that creates inequities later on. Pupils develop a dependency by having someone show them how to do a particular problem before they try it. I believe this dependency limits real success, in both science and mathematics education in secondary and tertiary institutions, to those who manage to preserve their independence of standard instruction.

Trying to prescribe behavior to reduce such dependency, however, is not my goal. Assuming that thinking individuals tend to convert prescriptions for their behavior by selective assimilation of these prescriptions to their own cognitive systems, which we now understand poorly, requires something more promising than conventional approaches. One such approach is to reflect the system back to its participants in such a way that they can assimilate the consequences of their own behavior to cognitive systems relating to their own goals, ideals, social norms, etc. One such reflection is what is called an ethnography. An ethonography portrays a style of life from the point of view of the participants in that style.

If teachers are to change their own perceptions by reading an ethnography, they do not need an external, material ethnography but an internal, cognitive one. What I have in mind is something like Gay and Cole's The New Mathematics and an Old Culture (1967), which is a cognitive ethnography of school mathematics that is said to have produced a big change in the curriculum of school mathematics in Liberia (Erickson, personal communication). I think teachers need a much more detailed study than that of Gay and Cole, one covering more facets of the curriculum. Perhaps the chief obstacle to such research is that it is more difficult for us to study our own educational system than it is for us to study the conceptions of mathematics in another culture.

Before describing various features of the approach I am taking, I would like to point out how my research methodology, or theory of method, compares with that in natural sciences and with the dominant view in educational research, as I see them. In Figure 1, the box drawn with large dashes in the upper left-hand corner represents a contemporary, post-Kuhnian, view among philosophers of science of the structure of most natural sciences. A qualitative theory or model (small dashes) of those things that are being studied directs the design of experiments used to extend and refine quantitative definitions and

[1]Traditionally, the child was seen as the key; now many argue that administrators are the key to change.

laws as well as the qualitative theory. The elements contained in the large solid-line box constitute a Kuhnian paradigm or research program (Lakatos, 1978), which is operated and evaluated by the thinking of scientists who are "outside," where they can, if they feel the need, consider competing theories. This system is influenced by other methodological principles and criteria and by cultural and evolutionary constraints.

The dashed-line box in the lower right-hand corner depicts the process I think most quantitative researchers involved in the study of science and mathematics education believe they are following. It seems they believe that experiments and correlations, obtained with little or no specific guidance from theory, can give rise to a quantitative theory of pupil learning. Researchers then interpret that theory for teachers who make decisions about, for example, how to promote thinking or learning. Whatever might be said about this as a caricature, the only thinking assumed is on the outside as far as the conduct of the research is concerned. The thinking that might be promoted in the pupils is strictly a product of teachers who are following policies implied in the findings or the theories developed by researchers and are, therefore, not thinking about the issues themselves. It is this view of research that leads researchers to focus on the pupil or, alternatively, the instructional policies of the institution of the school or district as the key to improvement.

The kind of research I am trying to develop is represented in the large dashed-line box in the upper right-hand corner of Figure 1. Here,

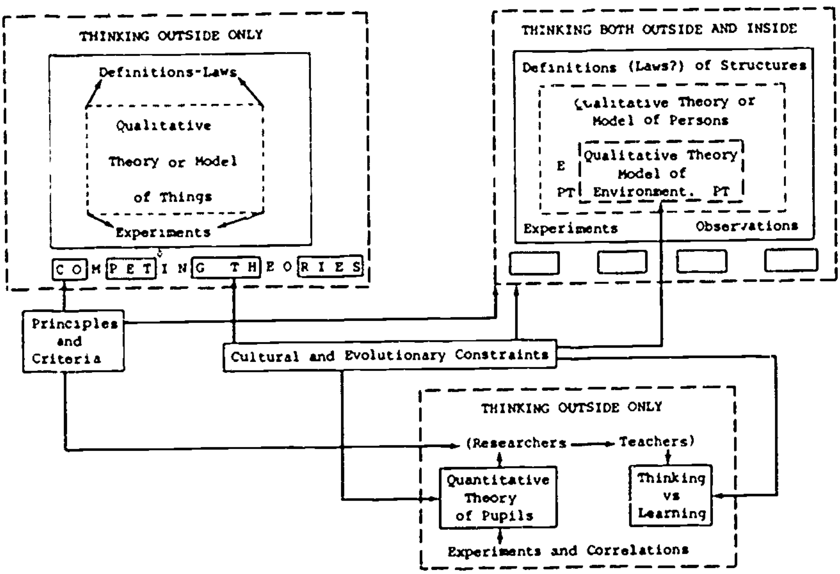


Figure 1. Three Research Metholodogies

we can see some similarities with, and some differences from, the other two forms of research. I propose to call this type of research "cognitive ethnography." The qualitative theory of pupils and teachers in an environment, represented by the letters P, T, and E in the box drawn with small dashes, must contain within it a qualitative theory of how the pupils and the teachers see their environment. How pupils and teachers think about their world determines, in very large measure, how they will interpret suggestions for changes in their world. In short, we must assume that thinking goes on inside the system being studied and interacts with it, as well as on the outside where the researcher's thoughts also interact with the developing qualitative theory. Just as the researchers develop experiments and observations from the theory and need to compare the theory with competing theories, represented by the small empty boxes outside the large solid-line box, so teachers and pupils inside the system have theories about the system that guide their experience of it and determine what they do as well as what they might say to the researcher about the system.

Another point the diagram represents is that the cultural and evolutionary constraints under which researchers operate also affect pupils and teachers. However, the principles and criteria of conducting research may be assumed to affect the researchers alone and not the pupils and teachers. For example, as a result of research studies and policy decisions in society at large, changes have occurred in many schools in the way both teachers and pupils regard low achievement in science and mathematics among minorities and girls. Basically, the researcher needs to assume that his or her thoughts while present in a classroom affect what he or she does and may indirectly change the way the teacher and the pupils think. Their thoughts and behavior, in turn, will affect the researcher's thoughts. In short, the researcher is part of the classroom ecological system. The researcher is inside, in contrast to the natural scientist who assumes he or she is outside of the system under study.

Participant Observation

The ethnographer's method for learning how the members of a social system think while functioning in it is called participant observation. By participating as a member of the social system, living in it as a fellow human being, the researcher can observe the ways of thinking of its usual inhabitants. But they may be experienced by the observer as much or more than they are observed as social processes. So the solid-lined box (upper right) is not excluding the observer who inhabits all of the boxes within--moving in and out as he or she changes roles. The apparent subjectivity of the participant observer is only overcome by methodological processes, chiefly his or her determination to look both within and without, for counter-examples to the qualitative theory already formulated. The observer should be surprised from time to time, as he or she detects a flaw in the efforts of the observed to conform to her or his expectation, or in her or his own experiences that violate theoretical expectations. An example should make this complex relationship clear.

In Figure 2, the box on the left represents the school. Children are coming into the school already under the influence of culture at home. They come oriented in somewhat different directions. Some of them can be recognized immediately by teachers as children who must be "saved," i.e., who must be brought back into the mainstream of the school--a set of cultural norms that includes the ideology of work. The teacher also perceives pressures of society, legal and governmental policies, equity policies for targeted groups--females, minorities, the handicapped, and the poor--that cause the teacher to perceive in a special way children in those categories who are moving away from the work ideology in mathematics. One of the by-products of that perception seems to be greater freedom for untargeted white, male, middle-class children. At least, those few who show a superior self-confidence in mathematics can get by without being socialized into the ideology of work.

Cognitive systems are implied for both teachers and pupils. For example, there are now strong legal forces teachers cannot afford to ignore if the targeted groups escape this socialization and should happen to perform poorly on tests. However, white male math mavericks, encouraged by their early freedom, can acquire self-confidence in attacking new problems. Such self-confidence serves them better in advanced scientific and mathematical study than the dependency learned by targeted groups who wait to be shown how to do each problem type.

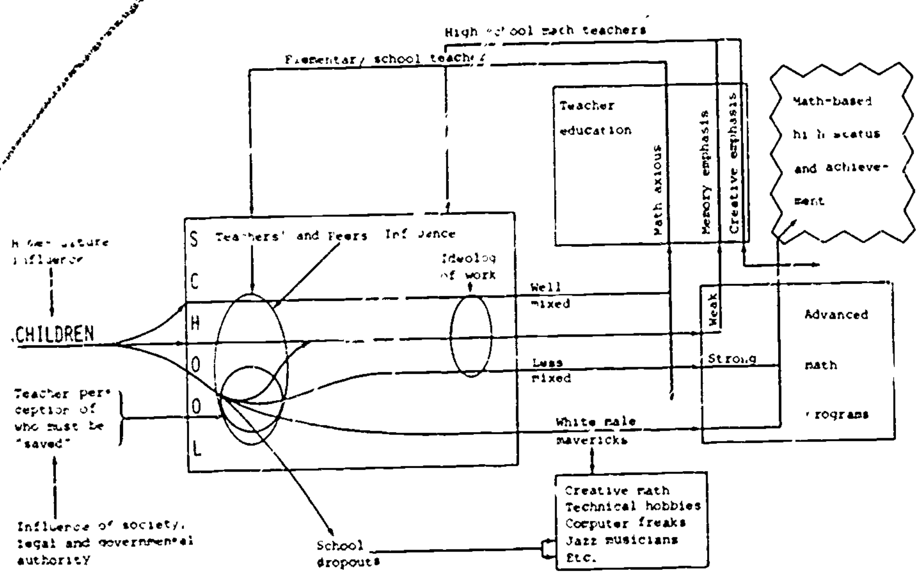


Figure 2. The "Life Cycle" of the Ideology of Work in School Mathematics

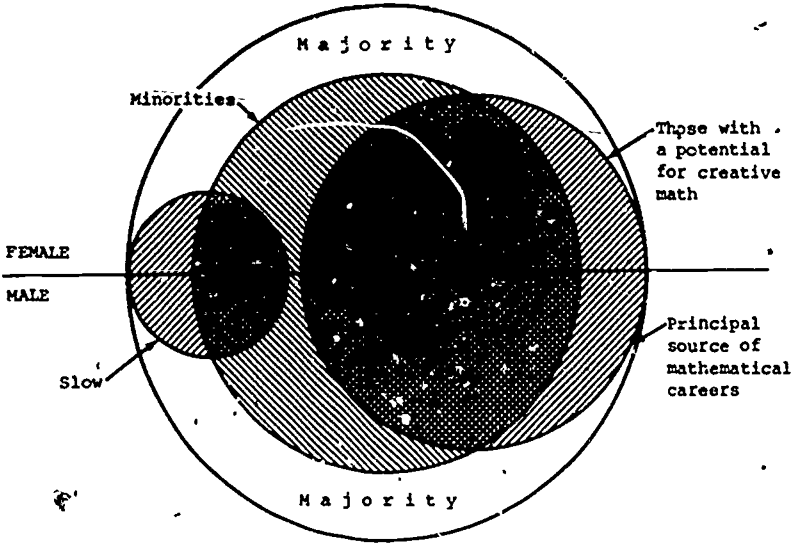


Figure 3. School Populations Targeted for Minimum Competency

The relationships of these targeted populations are shown in Figure 3. So the efforts of the teachers to bring minorities and girls up to some minimum standard through the idealogy of work may have little or no effect in changing the underrepresentation of targeted groups in math-based high status achievement occupations, such as the sciences, engineering, medicine, etc., which accord the highest status to creative problem solving.

Teacher education, represented in Figure 2 at the top center, selects future elementary teachers who tend to be math-anxious people and recycles their ideology of work. Some future high school math teachers, at least in this country, have a similar orientation towards work, while others may come from the white male math maverick source.

This is the kind of qualitative theory that needs to be developed and kept in mind, if one is going to do research that produces some kind of an effect on equity in mathematics teaching. Keeping such a theory in mind, of course, is not enough. It must be subjected to critical scrutiny, e.g., by a careful search for counter-examples. Once we locate a minority and/or female pupil who exhibits a freedom to explore nonstandard algorithms without being pressured into using only the standard ones, environmental conditions need to be closely examined to find out why the conjectured sociocognitive mechanisms are not working. Such study of a clear counter-example can lead us to a much more accurate definition of the mechanisms involved. For example, they may involve an absence of math anxiety on the part of that pupil's teacher, or the presence of a support group or resource person for the teacher.

Keeping such a broad ecological view in mind, we need to see in more detail what the thinking of the teacher and the child are at the classroom level (Figure 4). Children typically are involved in such things as counting. Here, teachers are drawing on the children's counting schemes, which have been learned at home from parents.

The teacher is thinking about the development of pupils' skill following an ideology of meaningful learning as the application of prior understanding. As the teacher introduces counting by 2s, 10s, 5s, etc., algorithms for adding (counting the total and counting on) and counting back in subtraction, we begin to see the typical American primary school mathematics curriculum emerging. The chief ideology operating here may be expressed by teachers as "teaching processes." They hear the term as jargon, but the meaning they attach is procedure-- a procedure that is rehearsed over and over again. Such procedures are considered to be meaningful, becaue they are based on prior skills and a presumed understanding of counting.

By the second grade, however, a conceptual failure clearly emerges as pupils need to use the decimal system in "borrowing" and "carrying"-- now called renaming and regrouping. Conceptual problems are evidenced in many ways, such as in children's failure to shift the tag sequence in counting money when the 1-to-1 correspondence rule should change to 1 to

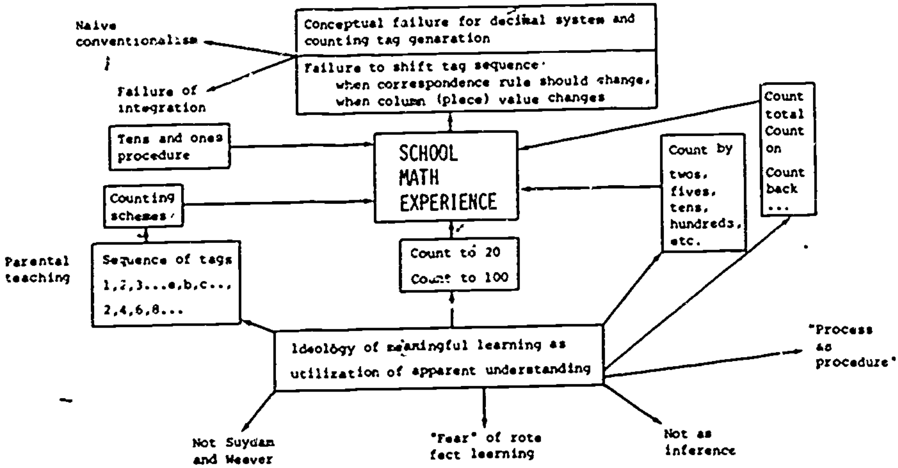


Figure 4. Conceptual Aspects of Primary School Mathematics

10, 1 to 5, or whatever. Teachers have different ways of explaining this; for example, as a memory problem. The interference of counting in conceptualizing multiplication can be recognized as what has been variously called preconceptions (Ausubel, 1963), alternative frameworks (Driver, 1973), or the "critical barrier phenomena" (Hawkins, 1979). By the fourth grade, we begin to see, emerging from the children's failure to understand the decimal system, a kind of naive conventionalism well described in Stanley Erlwanger's dissertation (1974). Because of a failure of conceptual integration, the only kind of integration that many children achieve is a kind of conventionalism. That is, they view mathematics as a massive collection of arbitrary, conventional rules and arbitrary, conventional "problems," to certain of which certain rules apply.

After locating female and minority math mavericks who are given as much freedom in school as some majority boys and making the adjustment of the descriptions that such counter-examples require, it is time to inform teachers of the theory, to see what happens when the mirror is held up to them.

First, the theory is explored with another group of teachers (for safety) who are not critical to further data. Of course, they deny immediately that they treat any majority male children differently from the others. The suggestion is made that children treat majority males differently according to their culture. Teachers may not realize that two children who behave quite differently are both trying to avoid doing their math for the same reason--boredom. The question is, how do the teachers feel about themselves? Are they intimidated by having children move ahead of the class? After differential response patterns are revealed to teachers, they may change their behavior with respect to such children.

How can the theory be subjected to a better test before it is revealed? Perhaps the researcher can examine her or his own reaction to a majority male math maverick to determine how teachers must feel about him. Perhaps there is another person--a research assistant--available who could be interviewed. The important idea that guides the investigator is to attain a sufficiently sympathetic reconstruction of the experience of the teacher, so that many teachers in many places will read it and respond with, "Yes, that's what it's like. Some little middle-class boys are just so clever and so confident, I can't come down on them hard because I don't really know what's best for them. I have a responsibility to the whole class and can't let everybody get away with the freedom they seem to thrive on." If the problem is brought to the level of conscious deliberation, it can be solved by most teachers with some sympathetic help from outside. But some solutions are likely to be better than others by external standards, such as the long-run improvement of minority representation in high levels of achievement in mathematics.

Teacher Accountability or Supportive Intervention?

Strengthening teacher accountability for pupil test scores is not a promising means for achieving fundamental change in social relations in the classroom nor for achieving a higher level of sophistication on the part of teachers in recognizing mathematical insights by children. A supportive intervention, on the other hand, can work, not so much by producing direct improvements in practice, as by deepening teachers' insights into what their practices are doing to children's attitudes about mathematics, and what some options are for changing their practices, in a tentative, piecemeal fashion. While a policy of accountability has the advantage of ease of dissemination, supportive intervention taps the sympathy and concern of workers in schools. Reaching the audience requires, as cognitive ethnographies, case studies that are so interestingly written that they have the potential for becoming best sellers, like Bel Kaufmann's Up the Down Staircase. A facile pen is still mightier than a thousand behavioral objectives.

A few dozen books that have recorded the experiences of teachers in popular novel form have already paved the way for much deeper insights that could come from analytical cognitive ethnography in the classroom. Perhaps some of the people Oscar Lewis (1966) wrote about have read his books and gained impetus for change from within the culture of poverty. Books written with both artistic talent and penetrating ethnological and cognitive insights about particular aspects of the life of classrooms could have a positive influence on the poverty of most mathematics teaching. Such books should report the dreams frustrated, hopes sometimes realized, and new ways of life discovered, not only for the cynics and critics of the schools, but for teachers and a concerned public. At present, there is little appreciation of the complex social and intellectual problems a first-grade teacher must solve in helping children to learn at a rate that satisfies them and in a direction that avoids conceptual traps.

Only a glimpse can now be provided into the solutions that are emerging from this work with place-value problems teachers experience in teaching primary grade mathematics, a curriculum permeated with two threads that cannot be woven together for most children--counting and place value. Counting emerges in all societies from parental teaching of sequences of tags. In Western societies, number names and the alphabet are taught at home along with nursery rhymes and songs. Place value is a thing of school. In oriental societies, however, grouping by fives and tens, hundreds, thousands, etc., seems to be more deeply embedded in the culture itself. The school turns counting into algorithms: count the parts, count the total, count on from the first number, while keeping track of how many are counted on (Figure 4). Count the whole, then count backwards while keeping track of how many are counted back. Count by 2s, 3s, 4s, 5s, 10s, 100s, etc. At first, it was difficult to consider seriously any essential conflict between counting and place value.

One of the advantages of the cognitive ethnographic approach over the usual psychometric and statistical approaches is that the kind of

qualitative theory that guides this research, the conjectured way of thinking, feeling, and seeing one's working environment, can be checked directly with the teacners and pupils who are the subjects of the research. However, the checking that is possible here is not only for additional validation, but for utility as well. If teachers or students find it personally useful to think about the social mechanisms in which they participate, or the cognitive mechanisms of their own minds, then there is a broader practical advantage to the research than theoretical inquiries destined for implementation through teacher education programs.

The Culture of Schools and the Culture of Mathematics Educators

If a closer look at what the teacners are thinking is taken, one discovers that, indeed, they are not interpreting meaningful learning in the way that Suydam and Weaver (1972) had interpreted it, but rather as a reaction against rote learning of number facts. Many teachers find support for math procedures that will guarantee success if rehearsed, in talk of "process" and "meaningfulness" by specialists who do not see process as procedure. Now, with this brief picture of the kind of thinking that is operating in most primary classrooms, let's move back and look at the broader ecological picture in which this system is embedded, where it acquires its terminology and prestige. Figure 5 represents key elements of two cultures. Above the dashed-line is the "culture of the schools," and below it the university-based "culture of mathematics educators."

As teachers make traditional kinds of presentations, dozens of such systems have to be provided to the children, who have to be informed and confirmed as to whether they succeeded or they failed. Their repeated successes lead eventually to global successes, and repeated failures lead to global failures. The culture of the school is protected by slogans (SL) and rationalizes the status quo. These small boxes represent piecemeal packages, materials, textbooks, and worksheets that can be introduced.

The university mathematics education ideology that operates generally in universities, regional laboratories, and professional organizations, also has its slogans and its packages. The university rationalizes success and failure competitively, i.e., more in terms of frequency distributions than absolute work accomplished. It has inspired numerous complete curriculum packages for the schools and theory-based programs and policies.

A certain tension exists between these two cultures. The university culture wants implementation of its programs. This is represented by the leftmost upward arrow. The teachers' union counterattacks saying that teachers have to take charge of their own in-service training and participate more in the training of teachers. More and more, we are finding this kind of counterpressure.

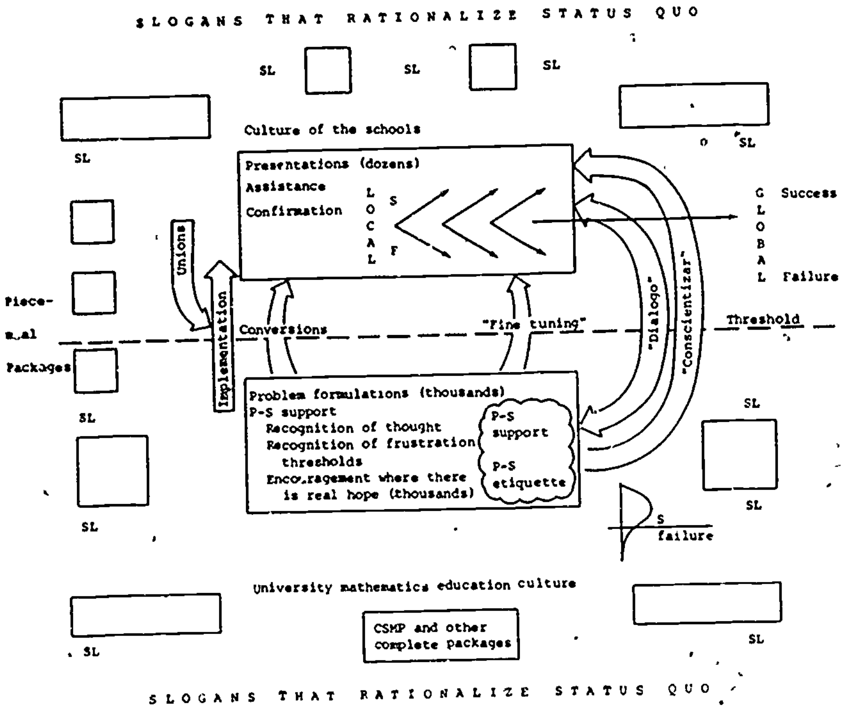


Figure 5. Interactions of Two Mathematics Education Cultures

Since this writer belongs to the university culture, and the research group is embedded in this culture, the first problem is how to build a bridge across to the school culture. What kinds of intercultural communication can be introduced effectively? Resource people are participant observers who serve as two-way bridges. It is not enough for them simply to be available in the schools with all the resources of the university culture at their disposal. Often, these are not the right resources. Helping teachers solve their problems demands the system represented in the lower of the two large boxes in Figure 5. Teachers must become capable of formulating math problems on the spot, thousands of them, in contrast to the dozens of processes teachers usually have teach, i.e., there are two orders of magnitude more things for a teacher to do.

If teachers are to learn to give problem-solving (P-S) support to all children, no matter how independent they may be, they need to recognize thought, to recognize frustration thresholds in children working on problems, to encourage them where there is need, and so forth. This kind of problem-solving support demands a problem-solving

classroom etiquette. This, in turn, usually requires a cultural change within the classroom so that problem solving is respected (Taylor, 1980). For example, it must be clear to everyone that no one interferes with or gives the answer to a child who is trying to work out a problem. The notion of process as procedure must also change.

Teachers who ask for help in teaching the so-called "renaming" algorithms can be introduced to ways of avoiding counting dependence, e.g., decomposition and recomposition based on what may be called subitized number facts from domino patterns, Cuisenaire rods, Diene's multibased blocks, and other concrete materials (Figure 6). Use of these materials makes it possible to see two 3s in the 6 pattern and two 2s in the 4 pattern. It seems clear from research on children's counting schemes (Gelman and Galistell, 1978; Stake, 1980) that these schemes are deep-seated and relatively inflexible, with strong social support at home and in schools.

However, multiplication requires a triple application of cardinal numbers to a collection of objects (three 2s equal 6), violating the several correspondence schemas used in counting. Visual patterns and decompositions seem to involve other schemas with no contradiction, so there is better overall payoff in using them for addition and subtraction. Place value, although usually taught before multiplication, involves the 1s and 10s times tables at the outset, and, therefore, the conception of a triple application of numbers. After this was recognized, ways were found that avoided counting altogether when one was teaching children who were having problems in learning place value. There is now a need to relate this cognitive mechanism to the child-based ideology of teachers that supports counting algorithms. When this relationship is pointed out to teachers, it seems to make sense to them.

Returning to the dissemination goals of the university-based "culture" in Figure 5, there seems to be little hope of getting a significant number of direct conversions, but one can begin to do some "fine tuning." This is usually acceptable to most teachers. One can start where the teacher is, introduce some fine tuning to make things work smoother for her or him; then gradually engage the teacher in a dialogue. When people engage in dialogue across two different cultures, particularly a sophisticated advanced culture with a more grass-roots culture that is tied to homes and traditions, learning is involved on both sides. By such dialogue, one can raise the consciousness of teachers. If teachers try they can raise the consciousness of university people. This writer has been engaged in such dialogues for about a year and a half, and some remarkable changes have occurred. These dialogues almost look like conversions but would not be such technically. These two terms, dialogue and consciousness raising, are taken from Paulo Freire's (1970) books about the implementation problems of agricultural technology.

As soon as children know the facts exhibited in the addition table of Figure 6, teachers can introduce inference schemes for plus or minus 1, and plus or minus 2, with which children can generate and fill in all the other cells. These plus or minus 1, plus or minus 2 inference

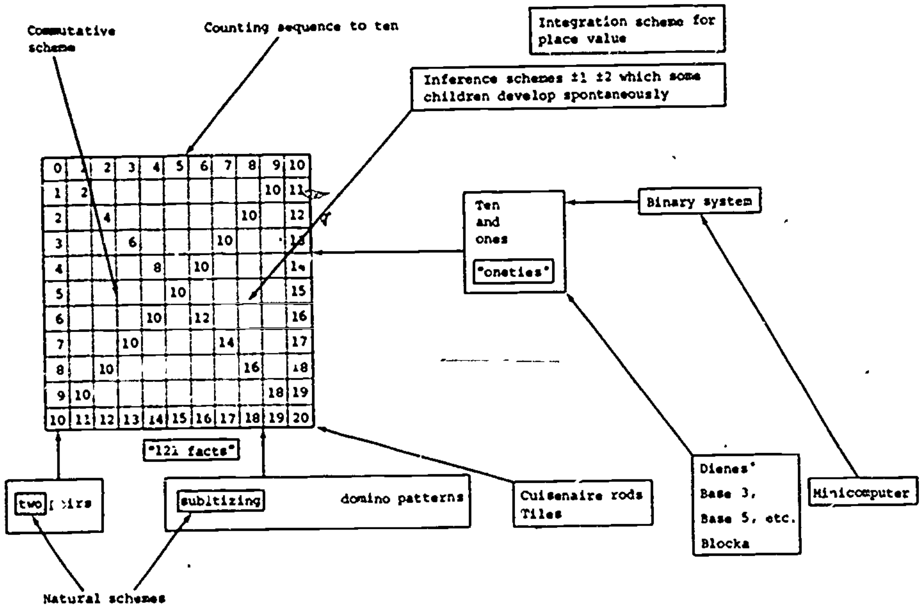


Figure 6. Schema for Integrating Place Value with One-Digit Addition Facts

schemes are really inferences, not procedures. One also has some choice as to which one to use, which encourages more creative problem solving.

Teachers seem especially gratified to discover that starting with diagonals and edges, the remainder of the square table can be filled by these four inference schemes that may seem, at first glance, to be somewhat hit or miss. So a new kind of system emerges out of helping teachers with their problems--not something brought in as a replacement curriculum. Other kinds of materials could be employed in developing the 10s and 1s scheme, for the object is to stimulate creative thought in both cultures, not to settle down on one system. After all, educators in the United States <u>are</u> committed to local control.

In this kind of intervention strategy, the idea is to respect the teacher's own starting point, conceptually and ideologically, and to try to find ways of working within that. The problem of a researcher is to find ways to describe this process.

When a social cognitive mechanism has been defined that is responsible for certain adverse (or beneficial) kinds of mathematics learning situations, the subjects themselves can also begin to learn to control these mechanisms. One result may be a declining frequency of the occurrence of the mechanism. This kind of theory building is quite different

from trait theory. A teacher who is math anxious in a given context may not be at all math anxious in another. A child who is counting dependent in one situation may not be in another. Moreover, the controlling situational variables are qualities perceived by the person, not absolute, objective qualities. This is both a challenge to empathetic theory development (a theory of understanding others) and a way to get rid of a lot of "noise" or "error" in the relational models. Policies based on variables accounting for 50% of the variance ignore a lot of deviance.

Researchers who are looking for policies that will select people for certain situations (e.g., teachers and pupils for class assignments, for promotions, etc.) according to personal traits may be disappointed in a theory that selects teachers for ways of thinking or for ideas that may be helpful or harmful in the educational development of children. Personnel decisions are now so much in the public domain, and so competitive in terms of public criteria of qualification, that no place is left for a decision in terms of whether or not the candidate for placement is interested in certain ideas about classroom work. Objective qualifications take over public administration, replacing much of the function of context designs.

Consequently, it is important that research into the social and cognitive mechanisms of mathematics education be addressed to teachers and educational administrators. The latter need research reports that are so readable and enlightening regarding the possibilities for improving contexts of teaching mathematics that they will see clearly the need for staff who promote better context through sympathetic listening, reflecting back the ideas they hear, injecting a bit of encouragement here and there and a new idea when it fits the problem raised. Some of the relations of other research traditions with this view of research proposed for mathematics education are further developed in Easley (1977, 1980).

Ethnographic studies in mathematics classrooms yield different and important knowledge about instructional processes. Examples of the kind of knowledge obtained and suggestions as to how such knowledge may be used have been briefly explored. The examples cited have been taken from research in mathematics classrooms. Similar studies in which the interactions of students, teachers, and particular subject matters, such as that of the natural sciences, are needed. Such studies provide insights into instruction that differs in kind from the knowledge gained from quantitative studies. The methodology proposed here requires the development of a theory of method incorporating thinking from outside, as a spectator, and thinking from inside--coming to an understanding of others, learning and adapting with others.

References

Ausubel, D. P. The psychology of meaningful verbal learning. New York: Grune and Stratton, 1963.

Bronfenbrenner, U. The experimental ecology of education. Teachers College Record 1976, 78, 157-204.

Driver, R. The representation of conceptual frameworks in young adolescent science students. Ph. D. dissertation, University of Illinois at Urbana-Champaign, 1973.

Easley, J. A., Jr. On clinical studies in mathematics education. Mathematics Education Information Report, ERIC Science, Mathematics and Environmental Education Clearinghouse, Columbus, Ohio, 1977.

Easley, J. A., Jr. Alternative research metaphors and the social context of mathematics teaching and learning. For the Learning of Mathematics, 1980, 1, 32-40 (Montreal: Concordia University).

Erlwanger, S. H. Case studies of children's conceptions of mathematics. Ph. D. dissertation, University of Illinois at Urbana-Champaign, 1974.

Erickson, F. Personal communication. At Michigan State University.

Freire, P. Pedagogy of the oppressed. Ramos, M., tr. New York: Continuum Press, 1970.

Gay, J., and Cole, M. The new mathematics and an old culture. New York: Holt, Rinehart and Winston, 1967.

Gelman, R., and Gallistel, C. R. The child's understanding of number. Cambridge, Mass.: Harvard University Press, 1978.

Hawkins, D. Critical barrier phenomena in science teaching. Outlook, 1979.

Kaufman, B. Up the down staircase. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Kuhn, T. S. The function of measurement in modern physical science. In H. Woolfe (ed.), Quantification: A history of the meaning of measurement in the natural and social sciences. Indianapolis: Bobbs-Merrill, 1961.

Lakatos, I. The methodology of scientific research programmes. In J. Worral and G. Carrie (eds.), Philosophical papers, volume 1. Cambridge, England: Cambridge University Press, 1978.

Lewis, O. Children of Sanchez. New York: Random House, 1966.

Suydam, M. M., and Weaver, J. F. Meaningful instruction in mathematics education. Columbus, Ohio: ERIC Information Analysis Center for Science, Mathematics, and Environmental Education, 1972.

Stake, B. Clinical studies of counting problems with primary school children. Ed. D. dissertation, University of Illinois at Urbana-Champaign, 1980.

Stake, R. E., Easley, J. A., Jr., and collaborators. Case studies in science education. Washington, D. C.: Government Printing Office (Stock Nos. 038-00-00377-1, -00376-3, and -00383-6), 1978.

Taylor, H. On classroom etiquette. In J. Easley, et al, Pedagogical dialogs in primary school math. Bureau of Educational Research, University of Illinois at Urbana-Champaign, 1980.

Chapter 13

IMPROVING EDUCATION IN THE SCIENCES THROUGH RESEARCH

James T. Robinson, Center for Educational Research and Evaluation

The purposes of the conference reported here and, therefore, of this publication were twofold. First, three fields of nontraditional research in science and mathematics education were reported. Each of these fields--the analysis of texts and curriculum materials, investigations of science understandings of individuals, and analysis of classrooms--is directly related to classroom instruction. Furthermore, each is just beginning to be applied to problems in science and mathematics education. Second, a variety of approaches is being applied to problems within the three fields of research. The preceding chapters reflect this variety but do not, of course, exhaust it.

Research in science and mathematics education has entered an exciting stage of development. New questions related to the subject matter of science and mathematics; questions that differentiate experts from novices, 11-year-olds from 13-year-olds; questions concerned with the context of their asking and the context in which answers are sought; and questions attending to the mechanisms of knowledge acquisition are being asked, and old ways of seeking their resolution are being questioned. This ferment has led to new knowledge of the complexities of text structure and comprehension, of the complexities of developing expertness in problem solving, and of the intricacies of the cultures of science and mathematics classrooms--and how little we know about these cultures.

The opportunities to expand significantly the understanding of knowledge acquisition in the sciences and mathematics over the next 15 to 20 years are most promising, more so than at any other time in this century. The research reported here is a beginning. When such research takes its place with traditional approaches, as well as currently unimagined approaches that will be developed in the future, the prospects are bright that research will contribute to and, indeed, influence and lead to improvements in the teaching and learning of mathematics and science.

The Significance of New Research Directions

Breaking new ground in research methodologies provides opportunities to examine extant practices. The previous chapters are intended to stimulate this kind of reexamination. They are also intended to provide accounts of alternative methodologies that can themselves be criticized. What are their limitations, what insights do they provide that are new contributions to knowledge, to practice? For this writer, several issues were raised that warrant discussion by members of the science and mathematics education communities. Others undoubtedly will be raised if the deliberations this publication is designed to stimulate do occur. The issues proposed below rest on the asumption that the goals of research in science and mathematics education are to produce knowledge that will, in the short or long term, contribute to the improvement of learning mathematics and science.

1. At what level of generality(ies) can research best contribute, given the current state of knowledge? Level of generality, in this context, refers to the extent of human activities subsumed within a problem. For example, problem solving, learning, creativity, and similar constructs have been, in many circumstances, investigated as unitary entities independent of subject matter and learner. A great deal of attention has been given to developing theories of learning—theories to apply to all learning of all subject matters and all age groups.

Asking much more specific questions such as, "How does an expert solve this physics problem?" may provide more knowledge relevant to science and mathematics education than has the more general research question. Perhaps there is a need to ask specific questions about learning specific subject matters of specific individuals before attacking the more general problem. If so, what subject matters and age groups will yield the most useful information? What different methodological approaches could be used in such research?

2. What kinds of expertise need to be used in research in mathematics and science education? Several papers presented in this work illustrate the value of cooperative research involving scientists, psychologists, and science educators in investigating certain educational problems. The culture or ecology of a science or mathematics classroom cannot be understood without giving attention to the content of teacher and student discourse and to the interactions of teachers and students with text and other learning materials. There is a need to know how students internalize such interactions. Do these processes differ in students of different ages or with different subject matter?

The conduct and analysis of interview data may require sophisticated knowledge of the subject matter and of interviewing techniques and their implicit risks, as well as the appropriate analytic techniques for handling such data. Would research be more fruitful if doctoral students in science and mathematics education worked with students in psychology and in a natural science discipline? Can faculty research be

structured in such ways that cooperative, interdepartmental research among graduate students can be facilitated?

3. Should "the obvious" be subject to research? This issue was most effectively raised for this writer by participant comments that the differences found in general math and algebra classes are "old hat" and well known to anyone who has taught in a secondary school. Research can be useful for this kind of situation if it contributes to a clearer definition of a problem that has become an accepted situation.

Instruction of the "other cultures," in Snow's idiom, is commonplace in mathematics and science education. It is commonplace in that it is well known to those who teach nonmajor, service, and other courses for "general" or citizen education in both high school and college. The existence of the situation and its ubiquitous presence in departmental and other curricular agendas has not made this a research issue. A great deal of curriculum development activity has been focused on the "nonacademically oriented" student (with its unwarranted assumption), but little research has centered on the culture of courses for nonmajors as compared to that of courses for majors. Ethnographic or other types of classroom studies conducted with differing theoretical frameworks may yield the data needed to produce more insightful research problems than are currently available. There is a special need to conduct research studies that will transform the reality of "what is" into a problem in need of study.

Both Armbruster and Deese raised questions about the style used in writing text for science instruction. Both suggested that many more student comprehension problems result from an inductive style of science text as compared to a deductive style. This appears to be an area in which there is a need for additional research.

Related to the question of style is the matter of maintaining student interest. In an effort to be "interesting," some textbook writers may abandon a logically coherent writing style. Armbruster and Deese suggested that such efforts compound the comprehension problems of inexperienced students. What are the trade-offs among interest, boredom, and a logically coherent, deductive style of textbook writing? Much detailed research must be undertaken to provide a complete understanding of these issues.

4. Considering the many complexities of analysis of the printed page, what features are most important for research in science and mathematics education? What types of analysis are most important for the improvement of science instruction? Walker, Armbruster, and Deese made particular decisions about what text features to use. What other aspects of text also may be important?

Walker pointed out the importance of field studies, suggesting that normal science classrooms are where research on use of texts by teachers and students must occur if results are to be valid. Collaboration of science teachers and researchers in classroom contexts where print

materials are put to different uses would be one approach with potential. Science and mathematics education programs could well include a stronger emphasis on the teacher as a collaborator in classroom research.

5. Both Armbruster and Deese suggested that there is a need for specific classroom instruction in how to read expository materials of the type usually included in science textbooks. Science teachers often complain that their students cannot read science texts adequately. The complaint generally implies that someone should have taught them how to read such materials. Yet, research studies are pointing up the unique problems of comprehending expository materials encountered by many students. This seems especially true for poor readers. They have less knowledge and must depend to a greater degree on information explicitly stated in the surface structure of the text. This issue involves those members of the student population who are poor readers and nonmajors in science or mathematics. What about other groups--poor readers who are majors, or able readers who are nonmajors?

Publishers, in response to pressures from teachers to simplify text and to reduce the "reading level," are turning many science texts into "media events." If the argument that learning from expository text ought to be a goal of science and mathematics education were accepted, what changes would need to be made in text construction, teacher education, and classroom practices?

6. Popularizers of science such as Conant (1951) and Bronowski (1953) portrayed science as "commonsense." They argued that science, but not each particular science, could be understood at a general level by anyone who could read and think. In approaching science from a philosophic and historic perspective, they attempted to provide the nonscientist a bridge to understanding science.

The task of the school, as currently practiced, is more complicated than providing general understanding. Science courses must enable students to learn technical details of particular disciplines and, occasionally, of interrelationships of the technical aspects of science to technology and society. The work of Johnson, Larkin, and Lochhead demonstrates clearly that science is not commonsense. When the understandings of individuals are investigated--individuals who have taken science courses--one is struck by the degree to which basic science concepts such as motion contradict, in such fundamental ways, the scientific explanations of experts in physics.

The results of research reported here and elsewhere can have immediate impact on science and mathematics educators and their students. For example, McCloskey, Caramazza, and Green (1980) suggested that those teaching science at any level should take into account the commonsense explanations of their students--preconceptions and misconceptions, from the scientist's perspective. They interpret their research and that of others as demonstrating that students' belief systems must be addressed, or science instruction may only serve to

provide students with new terminology for expressing erroneous beliefs. To make the point, they used an interview transcript of a student's prediction of the path of a ball after it is shot out of a curved (semi-circular) tube. The student used the terms momentum and angular momentum, in his explanation of the curved path he predicted the ball would follow.

Achievement tests and examinations used to determine students' understanding of the particulars of a science in both teaching and research situations have been shown by the research reported here to be clearly inappropriate procedures for answering fundamental questions about the knowledge people have of the natural world. Their use in determining grades and even more important, competence, warrants more discussion than can be accommodated in this publication.

Summary

The six issues raised here are intended to provide some "grist" for the intellectual "mills" of graduate research seminars in science and mathematics education. The papers presented in the earlier chapters gave examples of alternative methodologies that are offering new directions for research in science and mathematics education. This chapter ends with a list of references that extends the discussions of research methodologies and approaches. These papers reflect the deliberations about methodological issues that are taking place within the broader educational research community. The references were selected for their relevance to the issues raised by the conference papers and the references they cite.

Two significant and substantive papers with particular relevance to the underlying issues of competing research methodologies were presented at the annual meeting of the National Association for Research in Science Teaching, April 1981. These two papers will be especially valuable as resources in deliberations about the conduct of research in science and mathematics education:

Roberts (1981) presented an analysis of the metaphysical roots of the differences between quantitative and qualitative research. He examined the requirements for any research report and concluded with an argument for broadening the concept of legitimate research in science education. This thoughtful, scholarly paper is an important contribution to the development of research in science education.

An evaluation of naturalistic and conventional inquiry was presented by Welch (1981). He compared the advantages and limits of the two research approaches. The problems of judging the validity of data and the overall quality of naturalistic research were raised and discussed..

The conference was a stimulating and provocative experience for those in attendance. This conference report was prepared to provide a basis for evoking similar experiences for faculties and graduate students concerned with conducting research that will lead to improvement in the teaching and learning of science and mathematics.

References

Bloom, B. S. The new direction in educational research: Alterable variables. Phi Delta Kappan, 1980. 61(6), 382-385.

Bronfenbrenner, U. The experimental ecology of education. Educational Researcher, 1976, 5(9), 5-15.

Bronowski, J. The common sense of science. New York: Vintage Books, 1953.

Brown, A. L., Campioni, J. C., and Day, J. D. Learning to learn: On training students to learn from texts. Educational Researcher, 1981, 10(2), 14-21.

Conant, J. B. Science and common sense. New Haven: Yale University Press, 1951.

Cooley, W. W. Explanatory observational studies. Educational Researcher, 1978, 7(9), 9-15.

Elbaz, F. The teacher's "practical knowledge": Report of a case study. Curriculum Inquiry, 1981, 11, 43-71.

Gehlbach, R. D. Individual differences: Implications for instructional theory, research, and innovation. Educational Researcher, 1979, 8(4), 8-14.

Glaser, R. Trends and research questions in psychological research on learning and schooling. Educational Researcher, 1979, 8(10), 6-13.

Glass, G. V. Policy for the unpredictable (uncertainty research and policy). Educational Researcher, 1979, 8(9), 12-14.

Light, R. J. Capitalizing on variation: How conflicting research findings can be helpful for policy. Educational Researcher, 1979, 8(9), 7-11.

Leonard, W. H., and Lowery, L. F. Was there really an experiment? Educational Researcher, 1979, 8(6), 4-7.

Magoon, A. J. Constructivist approaches in educational research Review of Educational Research, 1977, 47(4), 651-693.

McCloskey, M., Caramazza, A., and Green, B. Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. Science, 1980, 210, 1139-1141.

Mishler, E. G. Meaning in context: Is there any other kind? Harvard Educational Review, 1979, 49(1), 1-19.

Reif, F. Toward an applied science of education: Some key questions and directions. Instructional Science, 1978, 7, 1-14.

Rist R. C. Blitzkrieg ethnography: On the transformation of a method into a movement. Educational Researcher, 1980, 9(2), 8-10.

Roberts, D. A. The place of qualitative research in science education. Invited address to the 54th Annual Meeting of the National Association for Research in Science Teaching, Grossinger's in the Catskills, April 5-8, 1981.

Schubert, W. H. Recalibrating educational research: Toward a focus on practice. Educational Researcher, 1980, 9(1), 17-24.

Shaver, J. P. The productivity of educational research and the applied-basic research distinction. Educational Researcher 1979, 8(1), 3-9.

Strike, K. A. An epistemology of practical research. Educational Researcher, 1979, 8(1), 10-16.

Tanner, D., and Tanner, L. N. Emancipation from research: The reconceptualist prescription. Educational Researcher, 1979, 8(6), 8-12.

Turnbull, W. W. Setting agendas for social research. Educational Researcher, 1979, 8(8), 9-12.

Varenhorst, B. B. In search of the illusive view of man. Educational Researcher, 1978, 7(9), 3-8.

Welch, W. Naturalistic inquiry and conventional inquiry: An evaluation. Invited address to the 54th Annual Meeting of the National Association for Research in Science Teaching, Grossinger's in the Catskills, April 5-8, 1981.

Research in Science Education: New Questions, New Directions

Participants

Harold M. Anderson
School of Education
University of Colorado
Boulder, CO 80309

Ronald D. Anderson
School of Education
University of Colorado
Boulder, CO 80309

Bonnie B. Armbruster
Center for the Study of Reading
51 Gerty Drive
University of Illinois
Urbana-Champaign, IL 61803

Carl Berger
School of Education
University of Michigan
Ann Arbor, MI 48109

Alfred Boemer
William B. Travis High School
1211 East Oltorf Street
Austin, TX 78704

Nicholas A. Branca
Department of Mathematical Sciences
San Diego State University
San Diego, CA 92182

R. Dean Brown
Department of Science Education
Colorado State University
Fort Collins, CO 80521

Ernest Burkman
College of Education
Florida State University
Tallahassee, FL 32307

David P. Butts
Department of Science Education
University of Georgia
Athens, GA 30602

Rodger W. Bybee
Department of Education
Carleton College
Northfield, MN 55057

William P. Callahan
Department of Special Education
Education Center, Room 150A
University of Northern Iowa
Cedar Falls, IA 50613

Jack L. Carter
Department of Biology
The Colorado College
Colorado Springs, CO 80907

Greg Christianson
2040 Lambs Canyon Road
Salt Lake City, UT 84108

George M. Clark
BSCS

Eunice Combs
BSCS

Malcolm Correll
320 20th Street
Boulder, CO 80302

James E. Davis
SSEC, Inc.
855 Broadway
Boulder, CO 80302

James Deese
Department of Psychology
Gilmer Hall
University of Virginia
Charlottesville, VA 22901

Jack A. Easley, Jr.
Professor of Teacher Education
Bureau of Educational Research
188 Education Building
University of Illinois
Urbana, IL 61801

Arthur W. Elias
Director for Professional Services
BioSciences Information Service
2100 Arch Street
Philadelphia, PA 19103

James D. Ellis
Visiting Assistant Professor of Science Education
Science Education Center
Education Building 340
University of Texas at Austin
Austin, TX 78712

Gaalen Erickson
Faculty of Education
University of British Columbia
Vancouver, B. C.
CANADA V6T 1W5

Fred N. Finley
226 Teacher Education Building
University of Wisconsin
225 N. Mills Street
Madison, WI 53705

Abraham Flexer
2790 Dartmouth
Boulder, CO 80303

Roberta Flexer
2790 Dartmouth
Boulder, CO 80303

James Gaskell
Faculty of Education
The University of British Columbia
2075 Wesbrook Mall
Vancouver, B. C.
CANADA V6T 1W5

Walter L. Gillespie
Acting Assistant Director for Science Education
National Science Foundation
1800 G Street, N. W.
Washington, DC 20550

Stephen P. Godomsky, Jr.
Division of Education
University of Maine at Farmington
104 Main St.
Farmington, ME 04938

Jeanne Dwight Gottschalk
BSCS

Richard E. Haney
Professor of Science Education
Department of Curriculum and Instruction
University of Wisconsin-Milwaukee
Milwaukee, WI 53211

Norris Harms
School of Education
University of Colorado
Boulder, CO 80309

Faith Hickman
BSCS

Paul DeHart Hurd
549 Hilbar Lane
Palo Alto, CA 94303

J. W. George Ivany
Dean of Education
Simon Fraser University
Burnaby, B. C.
CANADA V5A 1S6

Alvin A. Johnson
563 Leisure World
Mesa, AZ 85206

Paul E. Johnson
205 Elliott Hall
University of Minnesota
75 East River Road
Minneapolis, MN 55455

Jane Butler Kahle
Department of Biology
Purdue University
West Lafayette, IN 47907

Robert Karplus
Lawrence Hall of Science
University of California
Berkeley, CA 94720

W. Diane Keller
BSCS

Eileen Kintsch
435 College Avenue
Boulder, CO 80302

Leopold E. Klopfer
Learning Research and
Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Oliver P. Kolstoe
Professor of Special Education
University of Northern Colorado
Greeley, CO 80639

John J. Koran, Jr.
Rt. 3, Box 308
Alachua, FL 32615

Perry Lanier
Institute for Research in Teaching
Michigan State University
201 Erickson
East Lansing, MI 48824

Alma E. Lantz
Denver Research Institute
University of Denver
University Park
Denver, CO 80208

Jill H. Larkin
Assistant Professor
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Addison E. Lee
Professor Emeritus
Science Education Center
University of Texas
Austin, TX 78712

Ivo E. Lindauer
Department of Biological Sciences
University of Northern Colorado
Greeley, CO 80639

Joseph I. Lipson
National Science Foundation, SEDR
1800 G Street, N. W. (W. 638)
Washington, DC 20550

Arthur Livermore
Office of Science Education
American Association for the
Advancement of Science
1776 Massachusetts Avenue, N. W.
Washington, DC 20036

Jack Lochhead
Department of Physics and
Astronomy
University of Massachusetts
Amherst, MA 01003

Vincent N. Lunetta
Director, Iowa-UPSTEP
The University of Iowa
Iowa City, IA 52242

William V. Mayer
BSCS

Mary C. McConnell
2975 N. W. Imperial Terrace
Portland, OR 97210

Alan J. McCormack
Science and Mathematics
Teaching Center
University of Wyoming
University Station, Box 3992
Laramie, WY 82071

Lillian C. McDermott
University of Washington
Seattle, WA 98195

Joseph D. McInerney
BSCS

Jon D. Miller
Associate Dean for Research
Northern Illinois University
DeKalb, IL 60115

Irving Morrissett
SSEC, Inc.
Educational Resources Center
855 Broadway
Boulder, CO 80302

James L. Neujahr
City College of New York, A209
Convent Avenue at 138th Street
New York, NY 10031

John Nicholas
Canberra College of
Advanced Education
P. O. Box 1
Belconnen, A.C.T. 2616
AUSTRALIA

T. C. O'Brien
Teachers Center Project
Southern Illinois University
Edwardsville, IL 62026

Kenneth Olson
Science Education Center
University of Northern Colorado
Greeley, CO 80631

Roger G. Olstad
College of Education
University of Washington
115 Miller, DQ-12
Seattle, WA 98195

John J. Patrick
Social Studies Development Center
Indiana University
513 North Park Avenue
Bloomington, IN 47401

Rita Payton
School of Nursing
University of Northern Colorado
Greeley, CO 80639

Jerald J. Pelofsky
3101 So. 51 Street
Kansas City, KS 66106

Rita W. Peterson
National Science Foundation
1800 G Street, N. W.
Washington, DC 20550

Harold Pratt
Jefferson County Public Schools
1209 Quail Street
Lakewood, CO 80215

William C. Ritz
Natural Sciences Department
California State University at
Long Beach
Long Beach, CA 90840

Douglas A. Roberts
Ontario Institute for Studies in
Education
Department of Curriculum
252 Bloor Street West
Toronto, Ontario
CANADA M5S 1V6

James T. Robinson
CERE/BSCS

Norris M. Ross
CERE/BSCS

Ronald D. Simpson
North Carolina State University
326 Poe Hall
Raleigh, NC 27650

Mary Lee Smith
Research Associate
School of Education
University of Colorado
Boulder, CO 80309

Jane Stallings
The Teaching and Learning
Institute
409 Poppy Place
Mt. View, CA 94043

David R. Stronck
Department of Social and Natural Sciences
Faculty of Education
University of Victoria
P. O. Box 1700
Victoria, B. C.
CANADA V8W 2Y2

Constance P. Tate
Baltimore City School System
Division of Instruction and Curriculum Management/Office of Science
Oliver and Eden Streets
Baltimore, MD 21213

Richard R. Tolman
CERE/BSCS

Les W. Trowbridge
Science Education Center
University of Northern Colorado
Greeley, CO 80631

Richard L. Turner
Dean, School of Education
University of Colorado
Boulder, CO 80309

James R. Wafles
School of Education
University of Colorado
Boulder, CO 80309

Decker F. Walker
School of Education
Stanford University
Stanford, CA 94305

Jo Williams
CERE/BSCS

Dael L. Wolfle
Graduate School of Public Affairs
University of Washington
266 Smith Hall
Seattle, WA 98195

Robert E. Yager
Department of Science Education
University of Iowa
Iowa City, IA 52242

198