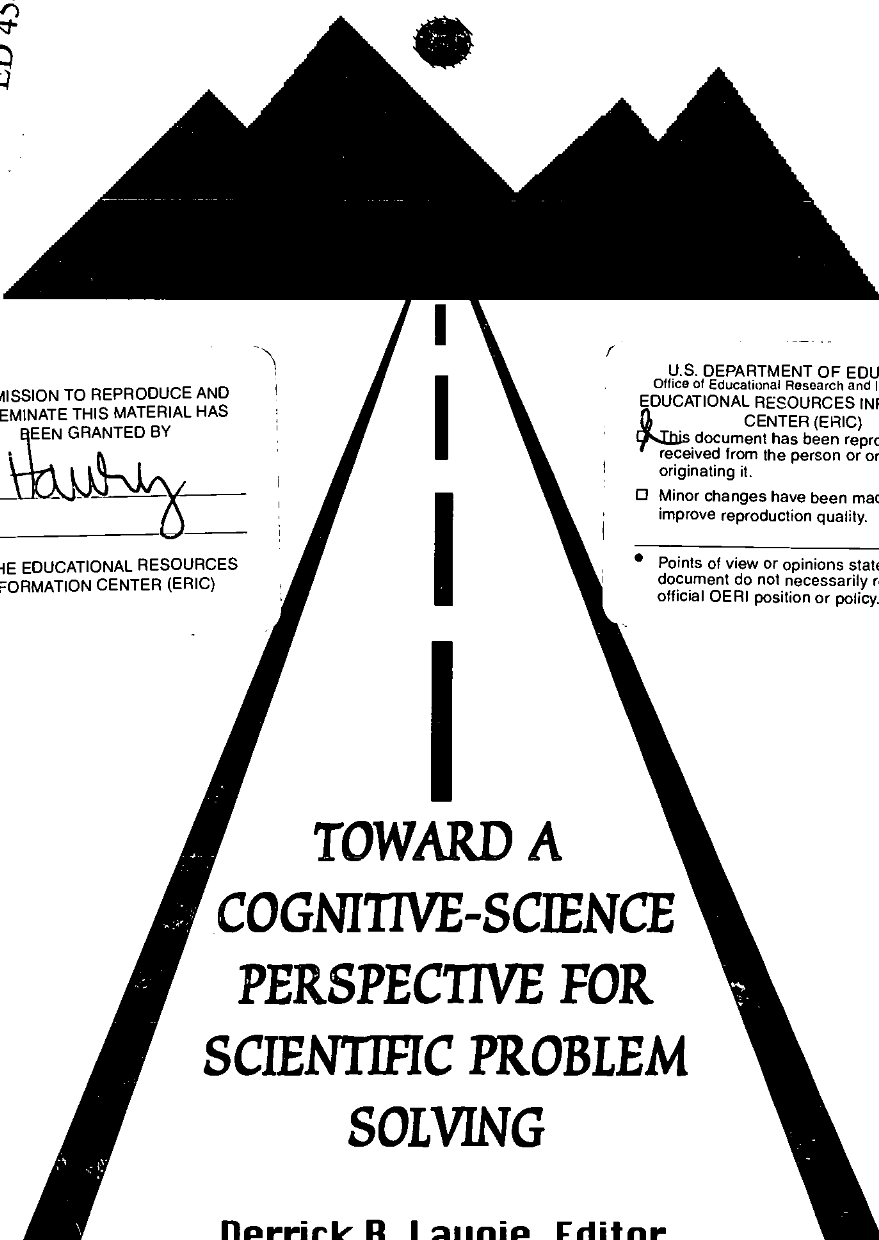
DOCUMENT RESUME

ED 454 060		SE 064 917
AUTHOR	Lavoie, Derrick R., Ed.	
TITLE	Toward a Cognitive-Science Perspective for Scientific Problem Solving. NARST Monograph, Number Six.	
INSTITUTION	National Association for Research in Science Teaching.	
PUB DATE	1995-00-00	
NOTE	342p.	
AVAILABLE FROM	NARST, 1929 Kenny Road, 200E, Columbus, OH 43210-1080.	
PUB TYPE	Books (010) -- Collected Works - General (020)	
EDRS PRICE	MF01/PC14 Plus Postage.	
DESCRIPTORS	*Cognitive Processes; Computer Assisted Instruction; Elementary Secondary Education; Higher Education; *Problem Solving; Science Education; Teaching Methods; Technology	

ABSTRACT

This book attempts to review and extend current knowledge and perspectives on scientific problem solving by utilizing the framework of cognitive science. While the focus of the monograph is toward problem solving, its ultimate goal is the improvement of science teaching. Twelve chapters address problem solving in a variety of science content domains including biology, chemistry, physics, mathematics; process skills domains: hypothesizing and predicting; contexts: everyday and school settings; and educational foci: assessment, use of analogies, and applications of technology. With varying degrees of emphasis, each chapter addresses problem solving relative to research-based background, theoretical ideas and issues, practical applications for science teachers, and directions for further research. Knowledge representation, transfer, processing capacity, cognitive growth, distributed cognition, and context are topics that continue to remain challenging for problem solving researchers. With many of the chapters describing applications of computers, it is clear that technology will play an increasing role relative to the teaching, learning, evaluation, and research of scientific problem solving. Individual chapters contain references. (SAH)

NARST MONOGRAPH, Number Six, 1995



TOWARD A COGNITIVE-SCIENCE PERSPECTIVE FOR SCIENTIFIC PROBLEM SOLVING

Derrick R. Lavoie, Editor

A Monograph of the National Association for Research in Science Teaching

Number Six, 1995

TOWARD A COGNITIVE-SCIENCE PERSPECTIVE FOR SCIENTIFIC PROBLEM SOLVING

Derrick R. Lavoie, Editor
Department of Biology
University of Northern Iowa
Cedar Falls, Iowa 50614

Contents

Preface

Why Problem Solving?

This NARST monograph, *Toward A Cognitive-Science Perspective For Scientific Problem Solving*, was conceived in response to our Nation's need for a population of scientifically literate individuals who can think and solve problems. A focus on problem solving seems to have taken a back seat, not only in our classrooms, but in our respected science education research circles. This monograph is an attempt to renew science educators' focus on problem solving as one of the most important subjects of our research and teaching efforts. The term problem solving, while often evoking contrasting perceptions (as evident in this monograph), is inevitably viewed as the application of knowledge for the purpose of reaching solutions. It is the search for solutions which motivates and develops our problem-solving abilities in school or research as well as that which determines our survival both as an individual and as a species. Further, with our so called nation of scientifically illiterate, it becomes vital that we renew our efforts to improve problem solving at all levels.

Why Cognitive Science?

Cognitive science is an interdisciplinary discipline that blends research on neuro-cognition, language, computers, problem solving, learning theory, decision theory, and social-contextual interactions. Its goal is to understand how we communicate, conceptualize knowledge, and solve problems. Two useful orientations of cognitive science directed toward achieving this goal for science education are information processing and social-linguistics. Information-processing emphasizes the structure and mechanics of knowledge generation and application. Through methods such as protocol analysis, word categorization, and concept mapping, information-processing research has sought to identify the cognitive processes and structures involved with scientific problem solving and concept formation. Social-linguistic cognitive science emphasizes the situated-social context which is dependent on domain specific knowledge, language, and human interactions. Using interpretive methods of cognitive anthropology, research from this framework has sought out the details of the social processes of communicating, reasoning, and decision making during scientific problem solving in authentic every-day settings. Both orientations of cognitive-science have been useful for defining the "science" of science teaching.

Chapter Highlights

Toward A Cognitive-Science Perspective For Scientific Problem Solving is an attempt to review and extend current knowledge and perspectives on scientific problem solving by utilizing the framework of cognitive science as described above. While the focus of the monograph is toward problem solving, its ultimate goal is toward the improvement of science teaching. Twelve chapters address problem-solving from a variety of science content domains (e.g., biology, chemistry, physics, and mathematics), process-skill domains (e.g., hypothesizing and predicting), contexts (e.g., everyday and school settings), and educational foci (assessment, use of analogies, and applications of technology). With varying degrees of emphasis, each chapter addresses problem-solving relative to research-based background, theoretical ideas and issues, practical applications for science teachers, and directions for further research. Knowledge representation, transfer, processing capacity, cognitive growth, distributed cognition, and context are topics that continue to remain challenging for problem-solving researchers. With many of the chapters describing applications of computers, it is clear that technology will play an increasing role relative to the teaching, learning, evaluation, and research of scientific problem solving.

As with many edited works, the reader will find several consonant perspectives throughout the chapters (e.g., problem solving is a constructive process; prior knowledge is critical to problem-solving success; both procedural and declarative knowledge are necessary for learning science and solving science problems; the problem-solving process is as important as the problem-solving product). The reader will also find a few contrasting perspectives (e.g., authentic versus traditional settings; general problem-solving skills versus domain specific skills). This editor feels the contrasting viewpoints to be one of the strengths of the monograph. The following paragraphs briefly preview each chapter.

The introductory chapter by Wheatley establishes a foundation for the monograph by providing an overview of problem solving from a constructivist perspective which becomes a predominant theme throughout the following chapters. Wheatley addresses the much debated question: what is problem solving? He discusses several types of problem-solving heuristics and builds an argument that an exploratory mind set is the most important characteristic of successful problem solvers. He finally raises important research issues relative to problem-solving settings and the use of general heuristics.

The second chapter by Lavoie centers upon the cognitive-processing skills of hypothesizing and predicting (hypothetico-predictive reasoning) within a biology problem-solving context. Lavoie argues that hypothetico-predictive reasoning is crucial to the scientific process and scientific problem solving. He examines the information-processing details of how procedural and declarative knowledge, associated with hypothetico-predictive reasoning, is identified, organized, applied, and modified during problem solving. Based on this analysis, which includes explicit models of the hypothetico-predictive process, he provides recommendations for teaching strategies as well as questions for science-education research.

Hauslein and Smith focus the third chapter on the structure of knowledge and its relationship to scientific problem solving. They include a comprehensive review of the literature relative to knowledge structure and the techniques that are used to determine such structure. They enrich the information-processing paradigm with inclusion of the connectionist model of brain function. The chapter provides good background and raises important research issues for the science education researcher concerned with improving students' problem-solving ability as well as conceptual understanding of the subject matter.

The fourth chapter by Roth compares research on problem solving within traditional well-defined settings to problem-solving research within ill-defined (authentic) settings. He argues that while the information-processing paradigm has been useful to cognitive science for modeling problem solving in traditional settings it is inadequate for authentic settings. Using many examples from his own research with open-ended (inquiry) physics teaching, Roth demonstrates how the situated cognition (social-linguistic) orientation of cognitive science can be a viable alternative for understanding authentic problem solving.

In the fifth chapter, Doran, Helgeson, and Kumar deal with the content, process, and context components of assessment in scientific problem solving. They outline several difficulties with trying to assess problem-solving knowledge and provide several mechanisms for integrating problem-solving assessment with traditional planning and instruction. Using multiple examples, they discuss a variety of assessment techniques such as performance assessment and the use of embedded formats for determining structural knowledge, process skills, and logical thinking skills.

Linn and Clark's sixth chapter describes additional strategies for the assessment of students' authentic problem-solving. The first section chronicles an historical account of problem solving by describing how and when problem solving, assessment, and curriculum merged together.

The authors point out the shift from an emphasis on general ability and science information toward a focus on metacognition and knowledge integration. The second section discusses the relationship between problem solving, assessment, and the Computer as Learning Partner project. The final section considers applications of portfolio assessment. The authors use many illustrations and descriptive examples throughout their chapter.

Kumar and Helgeson center the seventh chapter on applications of computer technology, particularly multimedia technology, for the assessment of scientific problem solving. They discuss several advantages of multimedia such as its non-linear learning and teaching formats, individual adaptability, and capacity for immediate feedback. The authors describe the use of a multimedia system for improving students' problem-solving skills in chemistry and highlight computerized adaptive testing. The relationship between human information processing and computer processing is re-defined.

A cognitive-science constructivist teaching strategy referred to as "anchored instruction" is the focus of the eighth chapter contributed by the Cognition and Technology Group at Vanderbilt. Hypercard controlled videodiscs, rich in embedded information, are used as "anchors" for engaging students in authentic problem-solving environments. The chapter describes how anchored instructional environments have led to cognitive social environments that are stimulating for learning, teaching, and assessing scientific problem solving. Several collaborative projects are discussed.

The ninth chapter by Glynn, Duit, and Britton recognizes the importance of analogies as conceptual tools to aid problem solving and scientific understanding. The authors begin with a historical examples of how analogies have led to scientific discoveries and inventions. Throughout the chapter are found many other examples illustrating how analogies are used in scientific problem solving to help students relate prior knowledge to new knowledge, identify problems, and develop hypotheses. They define analogies, discuss analogy-based science misconceptions, and describe a constructivist model of analogical problem solving. They also show how this model can be used to train students to strategically generate, apply, and evaluate analogies during problem solving.

In the tenth chapter, Bodner and Domin examine how successful and unsuccessful problem solvers in chemistry disembed relevant information from a problem and then manipulate and construct (represent) such information into a form meaningful for solving the problem. Utilizing many examples from their own research base, the authors discuss the characteristics and implications of verbal/linguistic, symbolic, and methodological forms of problem representation for problem-solving success. They use schema theory to generate hypotheses to explain several representational differences and end with several questions for further research.

Willson's eleventh chapter is a comprehensive analysis of a variety of research methodologies for investigating scientific problem solving. While emphasizing cognitive-science problem solving, he meshes his chapter with an examination of Gestalt, psychometric, science content, and humanistic problem-solving methods. Often using a historical approach, he not only outlines the research methodologies of a particular domain, but also the knowledge claims of that domain. He ends with several criticisms of problem-solving research in science education and offers guidance for future research.

The last chapter by Peterson is devoted to synthesizing the previous chapters and further expanding the cognitive-science perspective of scientific problem solving. She identifies and analyzes three dominant themes of the monograph: constructivism; the cognitive structure of knowledge; and the contextual nature of problem solving. The latter part of the chapter discusses thought provoking lines of future inquiry including the developmental nature of problem solving and the relationship between neuro-cognitive process and problem-solving success. Peterson also points out several omissions in this monograph which include a need to develop a broader cognitive-science framework that reflects a neuro-cognitive foundation.

Acknowledgments

My editorial work on this monograph was inspired by NARST colleagues, many of whom are included here as chapter authors or reviewers. I would like to thank all the authors for their timely efforts to complete their drafts and for their important internal reviews of each others' work. I would also like to thank the external reviewers, Bill Holliday, Dorothy Gabel, Bill Robertson, Arthur White, and Larry Yore for valuable comments of selected monograph chapters. I especially thank Larry Yore, for his generous guidance as monograph coordinator. I also thank Rick Duschl, John Staver, Amy Stout, and the NARST Publications Committee for support throughout the project. Lastly, I thank my wife and family for their understanding and support of the many hours away at my computer.

Problem Solving from a Constructivist Perspective

Grayson H. Wheatley

Florida State University
219 Carothers Hall
Tallahassee, FL 32306-3032

Introduction

When we trace problem solving in school science and mathematics historically, we find that problem solving has often been thought of as, solving highly structured word problems appearing in texts. For example, there are uniform motion problems, mixture problems, and related rate problems. Often, the word problems were developed by the textbook authors not so much to develop problem solving but to provide practice for prescribed computational procedures. A student can usually know what method to use by identifying the method illustrated in the preceding lesson; a look back at worked-out examples is usually sufficient to determine the method to be used. Many of today's textbooks labeled as problem solving exercises are actually designed to practice demonstrated methods.

When viewed from a constructivist perspective, problem solving takes on quite a different meaning (von Glasersfeld, 1987, Segal, 1986), Wheatley, 1991). Constructivists see the individual as trying to make sense of their experiences. Thus much of cognition is problem solving while little of what typically occurs in school classrooms could be considered problem solving because the learner is rarely allowed to make decisions. In a classroom where instructional practices are compatible with constructivism, students are presented with tasks **before** being shown any solution procedures. The intention is that students will construct their own meaningful methods. Where problem solving is the explicit goal, non routine problems are usually selected. For example, Trowell and Wheatley (1994) describe a university mathematics problem-solving class where the following task was presented:

> Fraction of Singles - The fraction of men in a population who are married is 2/3. The fraction of women in the population who are married is 3/5. What fraction of the population is single?

This task proved to be problematic for the students of this class and a variety of solution methods were presented and discussed. This task was non routine for the class since previously none of the students had seen or devised a solution method for such a problem and the question could not be answered by performing computations directly on the numbers given.

ɔter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for ific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

What is problem solving?

The term problem solving has many meanings. Even in this book you will find that 'problem solving' can be interpreted in several ways. Polya posited four steps in problem solving; understand, plan, execute, look back. While these so called "steps" have been widely used by teachers and researchers alike, problem solving may not be quite so linear. Mason and Burton (1987) have challenged the Polya four-step approach to problem solving suggesting that there is much cycling back through the steps. Our study of individuals engaged in problem solving suggests that effective problem solving is much less organized than the Polya four steps would suggest (Trowell and Wheatley, 1994; Wheatley and Wheatley, 1982). In this chapter, *problem solving is what you do when you don't know what to do* - if you know how to do the task it isn't problem solving. Thus many of the tasks often used in science and mathematics classrooms as "problems" are in effect exercises. It can be argued that as teachers we can only suggest tasks - individual make their own problem to be solved. What is a problem for one person may be a trivial routine exercise for another. As Hundeide (1986) states "We tend to overlook the fact that problems only exist in relation to a background of expectations that are usually taken for granted" (p. 310). A student is engaged in problem solving when she experiences a situation which causes a perturbation - when the situation is unclear and no known methods seem to apply.

Polya's first step, understanding, is rarely accomplished before a plan is devised. Few scientists would state that they understand a problem before a solution has been obtained. In problem solving there is usually interpretation followed by a period of exploration in which patterns and relationships are constructed and mental images formed. Often, obtaining a solution to a stated question reveals fresh nuances which may suggest that the problem was not initially well understood. Sfard (1994) provides convincing evidence that mathematicians do not feel they understand a problem until they have a holistic image of the relationships. In her interviews, the mathematicians described carrying out procedures and solving problems and not understanding until much later, if at all.

Personal heuristics are constructed by the individual as they reflect on their experience, not as the result of direct instruction and practice. What to a university instructor may be an exercise can be a problem for students. On the other hand, I do not rule out "exercises" as problems. I have interviewed students for whom 83 - 38 was a problem. That is, they had constructed no procedure for determining an answer and thus had to problem solve to complete the task. The definition of problem solving used in this chapter does not exclude any task as being a problem for a student. I do not have in mind as problems just puzzle-type tasks or any specially contrived task such as the Tower of Hanoi. In contrast, much of what passes as problem solving in school science and mathematics is using a taught procedure on a similar word problem. In problem solving as defined in this chapter, "understanding" develops throughout the problem solving process. Since we are social beings, any personal formulation of a problem is clearly a function of our interaction with others. The very act of asking questions reflects a stance, an orientation toward life. In the broadest sense, we cannot *give* students problems. Problem solving instruction in schools is often the solving of well defined questions based on certain information provided, frequently with the method specified. In reality, we can only suggest tasks - students make the problem they solve. For some, the "problem" might be writing something on their paper to get partial credit and avoid looking stupid. Other students may interpret the "task" differently and choose to solve *their* problem, while others will make a problem from the task and become totally engrossed in making sense of the situation, not just obtaining an answer. They may solve their problem in more than one way in order to verify their solution or pose new 'what if' questions which they explore and solve. Thus we cannot isolate problem solving from the broader context of knowing.

I have come to believe we should opt for open ended, or what I like to call ambiguous tasks. For example, consider the question, "How much paper will be required to wrap five identical cubes as a single present with no overlaps?" A few moments considering this situation will reveal that there are a variety of interpretations of this task. How are the blocks to be arranged? Will the paper be pulled taut or adhere to the sides of cubes? What is meant by "how much paper", area? size of sheet? As students explore this task over several class periods presenting their interpretations, problems, solutions and answers, significant learning opportunities are created. There are significant learning opportunities just in the act of interpreting the problem. As individuals in a social setting attempt to negotiate their interpretations of a problem each can learn.

Heuristics

Problem solving heuristics are ways of coming to know a problem - actions which have the potential to lead to problem resolution. Heuristics differ from rules in that there is no guarantee that a particular heuristic will lead to an "answer." There is the danger that attempting to teach specific heuristics directly will result in the heuristics being viewed as new rules. Three types of heuristics can be described; exploratory, general, and domain specific. Since heuristics are personal, it is possible that what is a general heuristic for one person may be domain specific for another. Yet it may be helpful to think in terms of the three categories posed.

Exploratory heuristics

If a task is taken as a problem by an individual, he or she may have no idea how to begin. The most fundamental of all heuristics is DO SOMETHING. What one actually does will depend on many factors including their previously constructed schemes, intentions, and prior experiences with related problems. Doing something may involve mentally visualizing relationships, making a sketch, trying a number, testing a possible solution, using a special case or any of a variety of other moves. Constructing units often proves helpful. Many students are not confident problem solvers because they believe good problem solvers know what to do and will write down the solution in an orderly set of steps. Since they don't know what "the first step" is, they do nothing. Once they come to believe there is no one first step and that problem solving involves exploration, they are on their way to becoming effective problem solvers.

General heuristics

Much attention has been focused on general problem solving heuristics (Mason and Burton, 1991; Polya, 1962; Wheatley and Wheatley, 1982). Most school mathematics textbooks list a set of problem-solving heuristics to be taught. While there are differing views on how heuristics should be "taught," some researchers and many teachers have reported enhanced problem solving as a result of attention to heuristics.

Polya's four steps have often been interpreted as general heuristics; one first attempts to understand the problem, devises a plan, carries out the plan and looks back. More often Polya's steps have been view as characterizing phases in problem solving rather than heuristics. Lists of general heuristics usually include

Draw a diagram,
Make a list or chart,
Look for a pattern,
Guess and test,
Try a simpler case, and
Work backwards.

There is little doubt that knowledge of options in problem solving is useful. The more a person is consciously aware of possible directions to take, the better they are able to deal with the uncertainty of problem solving. However, it does not follow that these heuristics should be the focus of instruction *prior* to the assignment of tasks. It is likely that heuristics can become part of a student's integrated knowledge if they evolve out of their problem solving activity rather than being "taught" directly (cf. Lavoie, this Monograph). As students engage in problem solving cooperatively and present their solution methods to the class, opportunities exist for the teacher to call attention and even name particular heuristics being used by the students. Thus "Look for a Pattern" becomes the name for “What Marta was doing” rather than a method specified by the teacher to be used on a particular set of exercises.

Domain Specific Heuristics

Certain heuristics are domain specific. For example, in solving a geometry problem, it may be necessary to draw in auxiliary lines. In a physics problem, it may be useful to symbolize the force vectors. Some domains of inquiry naturally suggest certain heuristics which are not useful in other settings. In debugging a computer program one may refer to or construct a flow diagram. Or they may parse the lines of code into modules. Each domain of knowledge may have heuristics which work well because of the type and nature of the knowledge organization. Situated cognition theorists (Brown, Collins, and Duguid (1989) place great importance on such domain specific heuristics. Lave (1988) goes so far as to question the psychological construct of general transfer. She presents evidence that mathematics competency is integrally dependent on the setting in which the individual acts. For example, grocery shoppers were much more accurate in determining best buys than with similar arithmetic tasks on a computation test in a school-like setting. It should be noted that Lave’s work and the other studies cited dealt with computational arithmetic.

Problem solving in science and mathematics involves more abstract and varied tasks requiring the construction of high level abstractions and thus different cognitive mechanism may be operating. It is important for teachers to do a careful analysis of tasks and determine the essential constructions students must make and then formulate tasks which provide potential opportunities for those constructions to be made.

In large part because of the particular philosophy of science which dominates school science and mathematics, attention has been focused on problem solving rather than problem posing. Problem solving has been framed as a machine-like (computer metaphor) activity performed by ‘subjects.’ But all cognitive activity is carried out by a person who has goals, intentions, expectations, and unique personal experiences. To ignore the individual in problem solving is to dehumanize teaching.

Personal Heuristics and Problem Orientations

Wheatley and Wheatley (1982) report on the problem solving activity of grade six pupils. In their study, students were observed as they engaged in solving non routine problems. While students had an opportunity to learn six particular heuristics, wide individual differences were noted in the strategy use of students. Interviews were conducted with 102 students as they attempted five non-routine problems. Not only were there heuristic preferences but particular heuristics tended to be favored on each of the problems. The strategies observed being used most often were, guess and test, make a list, look for a pattern, draw a diagram, simplify, and write an equation. When all five interview problems were considered, the order of heuristic use was guess and test, make a list, look for a pattern, and draw a diagram. Students showed definite strategy preferences.

Trowell (1994) studied the problem solving of prospective secondary mathematics teachers enrolled in a course on mathematics problem solving. She reported strong individual preferences among the students for particular heuristics. In her study, draw a diagram was widely used as was guess and test and look for a pattern. Some students adopted a heuristic of setting up a Cartesian coordinate system and used it extensively while other students never used this heuristic. Draw a diagram was widely used by all students. Guess and test was used by some students but rarely by others. The frequency of use of any heuristic varied widely among the students. These studies suggest that individuals solve problems in idiosyncratic ways; each person constructs schemes which guide and influence their problem solving activity.

Other Theoretical Perspectives

Problem solving has often been studied using a "fine grained" analysis of a few individuals as they are engaged in problem solving for the purpose of then designing a computer program to solve problems based on the insights gained from these detailed analyses (Larkin, 1983, Simon and Barenfeld, 1969). Much has been learned about mental activity of individuals. However, one limitation of these studies lies in the "mind as machine" metaphor underlying the research. The mind is considered to be like a computer with input, output and storage. Such a view fails to consider the intentions, beliefs, confidence and motivations of students. If we are to progress in understanding the process of becoming a problem solver we must construct explanations of students' problem solving activity taking into consideration the setting, emotions, and intentions. Throughout the research activity, we must not over simplify the thought processes but instead keep in mind the extreme complexity of mental activity.

Theoretical perspectives on problem-solving research can often be like living in Abbott's Flatland. Abbott (1950) described people who lived in a two-dimensional world. Some were shaped like triangles, some squares and other multi-sided polygons. To these people, their entire world consisted of two dimensions. One day someone voiced the belief that there is a third dimension. This was denounced as heresy. Often, formulating our ideas about school science/mathematics, we believe our theory is comprehensive. Then someone comes along and says, "Wait a minute, what you take for granted as problem solving is actually culturally and even politically constituted - it is neither natural or necessary. That in fact, all societies do not act in terms of the problem-solution metaphor. This caution should encourage all of us to carefully examine our tacit assumptions and take a broad perspective. Of course there are times in which we may act as though certain issues are settled but they never are. An ecological perspective may bring into question practices we thought were acceptable until we see those practices embedded in a larger framework and thus find unexpected outcomes in other domains.

There is an innate drive in all of us to make sense of our experiences. I do not say make sense of our environment because some might interpret that to mean the environment is the same for all of us (My brother, Reginald, had Grayson as a brother; I did not). Thus the young child thrives and learns by giving meaning to an experience and then testing the viability of that construction. In a similar way, if one evening you hear a noise as you are sitting in your living room, you try to give it meaning, to interpret it. This is problem solving. As we consider classroom situations and consider problem solving as a goal, I suspect the problem solving we promote in science is not unlike learning as described above - it may be a difference in degree rather than kind. Too often, well defined tasks are presented and the teacher expects a particular "answer." Few learning opportunities arise in such a situation. However, if we encourage students to define their own problems, learning may occur, or more importantly, an individual may experience a perturbation which leads to a problem being formulated and perhaps solved.

Concepts make sense only if we have some abstract schema to organize and give meaning to the concept to be learned (Egan, 1988). This has been called the learning paradox (Cobb, Wood, Yackel, 1992). While attempting to explain concepts is unlikely to be effective, teaching through problem solving has rich potential. As we plan educational experiences for students, problem solving in the broader sense provides a basis for meaning making. Sensory experiences may prove invaluable as the individual is engaged in giving meaning to a problem they have posed but to act as though students learn simply from the "concrete" to the "abstract" may not be useful.

While attention to problem solving heuristics may be helpful, training students to use strategies such as draw a diagram or work backwards may not be. Students profit from constructing their own methods and then explaining their solution methods to the class (Lo, Wheatley, Smith, 1994; Wheatley, 1991). The teacher may wish to assist students in becoming more aware of the strategies they are using and even attaching names to their methods. Explaining and justifying solution methods is part of an instructional strategy called Problem Centered Learning which is ideal for problem-solving instruction (Wheatley, 1991).

Exploration

Perhaps the most important characteristic of successful problem solvers is an exploratory mind set. Helping students overcome a rule orientation should be a major goal of teachers. Until students can adopt an exploratory stance they are unlikely to become effective problem solvers. Successful problem solvers approach a "problem" in a relaxed manner realizing they have a repertoire of things to try. Many actually enjoy the challenge of the unknown while some persons become anxious when they don't immediately see how to solve the problem. Good problem solvers suspend judgment and realize they are likely to try several approaches in coming to know the problem. Wheatley and Wheatley (1982) concluded that adopting an exploratory mind set was the most important factor in successful problem solving. "Children seem to learn a rule oriented mind set in school which inhibits problem solving performance" (p. 116). It was those students who shifted from a rule orientation and explored the problem who were the successful problem solvers. In the Wheatley and Wheatley study, a group of students who continued to study a textbook driven curriculum performed poorly on a five item problem-solving test. In fact, the lower third of the treatment groups out performed the highest third of the control group. The following task was used with the 102 students interviewed.

> I am thinking of two numbers. The sum of the numbers is 33 and their difference is 15. What are the two numbers?

Students in the control group would typically compute 33-15 and write down an answer. Students who had studied problem solving for 18 weeks would often perform the same computation but with no expectation that it would yield an answer; they were just exploring the problem. They followed this exploratory move by other computations in a guess and test mode. This problem proved challenging for all groups but none of the 34 students in the control group solved the problem correctly. For those students who were successful, an exploratory stance was evident.

Problem Solving Viewed As Learning

In this paper, a particular view of learning is taken, one which seems especially appropriate for school science and mathematics in the nineties. Moving away from the logical positivist perspective which has dominated the sciences in this century (Kuhn, 1970), science and mathematics can be thought of as the construction of patterns and relationships - it is a personal activity. The patterns and relationships in physics will be different from those in biology or mathematics. Adopting a Piagetian perspective, learning results from the neutralizing of perturbations and thus necessarily results in some reorganization, or elaboration of existing schemes (Steffe, 1993). Problem solving can be thought of as reducing the healthy tension resulting from a perturbation.

We must recognize that all problem solving occurs in a culture, in a social organization. While the constructivist perspective focuses on individual cognition, there is acknowledgment that problem solving is heavily influenced by social interactions.

It does not seem useful to think of a hierarchy of types of learning such as proposed in previous decades by Gagne (1985). For Gagne, problem solving is a rule governed activity: it is selecting and applying the right rules. "Rules are the stuff of thinking" p. 179. Categorization systems such as facts, skills, rules, concepts, principles, problem solving seem logical until we consider the nature of knowledge and subsequent use of knowledge.

Problem solving is a comprehensive activity which is influenced by intentions, the setting (Brown, Collins and Duguid, 1989; Lave, 1988), our interactions with others, the personal schemes we have constructed, as well as the tasks set by an instructor. It must be recognized that while the schemes are described as personal, this does not imply they are constructed independent of social interactions. We are social beings and a major source of experience and perturbation is our community of others (Cobb and Bauersfeld, in press).

Research Issues

While the evidence for the value of problem solving is strong, too little is known about how one becomes an effective problem solver. What type of experiences will provide the potential opportunities for individuals to develop ways of coping with novelty. In particular, we do not know much about problem solving in practice. While Lave (1988), Saxe (1991), and others have studied individuals using arithmetic in there daily routines, their focus has been on using computational procedures. There is a need for studies which look at problem solving involving abstract concepts as found in mathematics and science. Bransford and his colleagues have effectively used and information-processing orientation in their research program (Bransford and Vye, 1989).

There is controversy about the role of *setting* in determining the problem-solving activity of individuals. To what extent are general heuristics utilized when not explicitly called for? Is it the case as Lave (1988) argues that the setting is all important? Taking the results of Lave and others, the value of general heuristics may be small but there is too much evidence on the other side of this argument to accept this conclusion without further study. Finally, we need to explore how we can organize school science and mathematics courses to enhance problem solving.

Summary

Each person builds a personal set of problem-solving heuristics even if specific heuristics have been experienced in the classroom setting. But central to success in problem solving is an exploratory mind set; students who fail to break out of a school induced rule/formula orientation have little success as problem solvers. Richards (1991) argues for abstract tasks - tasks which are potentially meaningful. He asserts that mathematics (and science) is inherently abstract - that students can and should be involved in abstractions. He questions the assumption that students are only motivated by what they think will be useful to them. Students can and should come to value activities which are abstract in nature (Egan, 1988). As we plan educational experiences for students, problem solving in the broader sense provides a basis for meaning making. As Prawat (1993) says,

> Creating a sense of wonder and awe in students should be our highest priority. I am concerned that the current preoccupation with *practical* problem solving does not advance us very far toward achieving that goal. (p. 14)

Because we live in a fast changing society, the ability to cope with novelty, to deal successfully with non-routine problems is be important. In another sense, becoming a problem solver may enhance the quality of life - it may lead to esthetic pleasure and even enhance the quality of someone else's life. Becoming knowledgeable may be the most important thing persons can do for themselves and society. There are many situations in which a knowledge of science and mathematics may be useful. The question is how we come to know. Problem solving a primary way of knowing.

References

Abbott, E. (1950). Flatland; a romance of many dimensions. Oxford: Blackwell.

Bransford, J. D. & Vye, N. J. (1989). A perspective on cognitive research and its implications for instruction. In L. B. Resnick and L. E. Klopfer (Eds.) , *Toward the thinking curriculum: An overview.* 173-205.

Brown, J, S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-41.

Cobb, P and Bauersfeld, H. (in press). *The emergence of mathematical meaning: Interaction in classroom cultures.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cobb, P., Wood, T. and Yackel, E. (1992). A constructivist alternative to the representational view of the mind in mathematics education. *Journal for Research in Mathematics Education, 23*(1), 1-33.

Egan, K. (1988). *Teaching as storytelling: An alternative to teaching and the curriculum.* London: Routledge.

Gagne, R. (1984). *The conditions of learning and theory of instruction,* 4th ed. New York : Holt, Rinehart and Winston.

Hundeide, K. (1986). Tacit background of children's judgment. In J. V. Wertsch (Ed.), *Culture, communication and cognition: Vygotskian perspective.* London: Cambridge University Press, pp. 306-322.

Kuhn, T. (1970). *Structure of scientific revolutions.* Chicago, IL: The University of Chicago Press.

Larkin, J. (1989). Research on teaching scientific thinking: Implications for computer-based instruction. In *Current cognitive research.* 1989 Yearbook of the Association for Supervision and Curriculum Development. 150-172

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life.* Cambridge, UK: Cambridge University Press.

Lo, J., Wheatley, G., and Smith, A. (1994). The participation, beliefs, and development of arithmetic meanings of a third-grade student in mathematics class discussion. *Journal for Research in Mathematics Education, 25*(1), 30-49.

Mason, J. & Burton, L. (1987). *Thinking mathematically.* New York: Addison Wesley.

Polya, G. (1962). Mathematical discovery: *On understanding, learning, and teaching mathematical problem solving.* New York: John Wiley and Sons.

Prawat, R. S. (1993). The value of ideas: Problems versus possibilities. *Educational Researcher,* 5-16.

Richards, J. (1991). Mathematical discussions. In E. von Glasersfeld, *Radical constructivism in mathematics education.* Dordrecht, Netherlands: Kluwer Academic Publishers.

Saxe, G. (1991). *Culture and cognitive development: Studies in mathematical understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Segal, L. (1986). *The dream of reality: Heinz von Foerster's constructivism.* New York: W. W. Norton.

Sfard, A. (1994). Reification as the birth of metaphor. *For the Learning of Mathematics, 14*(1), 44-55.

Simon, H. & Barenfeld, M. (1969). Information processing analysis of perceptual process in problem solving. *Psychological Review, 76*, 473-483.

Steffe, L. (1993). Interactions and children's mathematics. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Trowell, S. and Wheatley, G. (1994). Establishing an effective learning environment in a mathematics classroom. Paper presented at the Annual meeting of the International Group for the Psychology of Mathematics Education, Lisbon, Portugal.

Trowell, S. D. (1994). The negotiation of social norms in an university mathematics problem solving course. Unpublished doctoral dissertation, Florida State University, Tallahassee, FL.

von Glasersfeld, E. (1987). *The construction of knowledge: Contributions to conceptual semantics.* Seaside, CA: Intersystems Publications.

Wheatley, G. H. & Wheatley, C. L. (1982). Calculator use and problem solving strategies of grade six pupils. Final report submitted to the National Science Foundation. ERIC Number ED219250.

Wheatley, G. H. (1991). Constructivist perspectives on mathematics and science learning. *Science Education, 75*(1), 9-21.

The Cognitive-Processing Nature of Hypothetico-Predictive Reasoning

Derrick R. Lavoie
Department of Biology
University of Northern Iowa
Cedar Falls, Iowa 59614

Reason is the mechanical manipulation of abstract symbols which are meaningless in themselves, but can be given meaning by virtue of their capacity to refer to things either in the actual world or in possible states of the world.

George Lakoff
Women, Fire, and Dangerous Things

Introduction

This chapter is written from a cognitive-science perspective that combines concepts from cognitive psychology, computer science, learning theory, and problem solving. Its fundamental concern is the epistemology of knowlfic reasoning skills of hypothesizing and predicting, referred to as hypothetico-predictive reasoning. This chapter defines hypothetico-predictive reasoning, provides examples, and develops a rationale for its focus in science teaching. It then examines the information-processing nature of how knowledge, associated with hypothetico-predictive reasoning, is processed (i.e., identified, organized, applied, and modified). This leads to recommendations for teaching strategies as well as provocative questions for science-education research concerned with this important scientific-reasoning skill.

Cognitive-Science Framework

In general, the cognitive-science perspective of this chapter assumes that the teaching and learning of problem solving can be improved by understanding and modeling the information-processing nature of the human brain --- just as fixing a car depends, in part, on understanding the inner workings of the internal combustion engine. Success is problem solving is considered dependent upon the structure and organization of knowledge. Two types of knowledge are recognized -- procedural and declarative. Procedural knowledge is process knowledge -- knowing how, what, and when to execute particular cognitive actions. Declarative knowledge is the representation of facts, relationships between facts, analogies, experiences, etc. concerning a particular scientific domain. Procedural knowledge connects and

pter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for ific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

applies declarative knowledge to solve problems. Unlike procedural knowledge, declarative knowledge can be stored quickly and without commitment to where and when it will be applied (Anderson, Boyle, Corbett, & Lewis, 1990).

Improvement in problem-solving ability (i.e., learning) is viewed as a process by which connections are established and networked between procedural and declarative knowledge in increasingly complex ways --- ways that lead to more effective solutions.

The Hypothetico-Predictive Reasoning Process

The hypothetico-predictive process, like any problem-solving process, involves the manipulation of knowledge-based categories. Knowledge categories have a wide continuum relative to their complexity of content and structure. For example, they can be concretely or abstractly associated with pictures of objects (e.g., table, face), conceptual ideas (e.g., conceptual model of planetary motion), or processes (e.g., making bread; constructing a hypothesis). The development, interpretation, and application of knowledge-based categories has significance to all education and is at the heart of the hypothetico-predictive reasoning process.

> There is nothing more basic than categorization to our thought, perception, action, and speech... Whenever we reason about kinds of things -- chairs, nations, illnesses, emotions, any kind of thing at all -- we are employing categories. (Lakoff, 1987, p. 5)

Definition

Hypothetico-predictive reasoning is viewed as a problem-solving endeavor driven by an inductive-deductive reasoning pattern that generates hypotheses followed by associated predictions about scientific phenomena. This process of categorization and re-categorization of knowledge leads to an expansion of knowledge.

While it is obvious that making successful predictions about the stock market, one's marriage, or the bending of light due to large gravitational masses is advantageous, the conditions and processes that lead to such predictions are not so obvious. To develop a hypothesis, specific categories of information (facts, procedures, experiences, concepts, relationships between concepts) are synthesized to form more general or umbrella categories in an attempt to assign meaning to and make sense of something. Anderson (1983) views schema as abstract knowledge structures representing information about relationships between multiple categories.

A hypothesis can be defined as a causal scientific explanation which addresses the "whys," "whats," or "hows" about a scientific phenomenon. For example, a hypothesis concerning what causes green apples to taste sour may reason that:

> Green apples taste sour because the starch in their cells have not yet been converted to sugar through the process of ripening.

The following facts, experiences, and conceptual relationships could have been used to synthesize the above hypothesis.

1). Some green apples I've tasted were sour. (experience)
2). The process of ripening in fruit occurs over time and is often accompanied by a change in color and size. (between relationships)
3). Starch tastes bitter. (fact)
4). Sugar tastes sweet. (fact)
5). Apple fruit is cellular in its composition. (fact)

In science, predicting is a deductive reasoning process that identifies a specific event that might occur in the future, given that a particular scientific generalization (i.e., hypothesis) is true or not true. Testing a prediction can therefore provide: 1) evidence in favor of supporting a hypothesis, or 2) evidence in favor of refuting a hypothesis. Obviously, predictions must be associated with a hypothesis to be scientifically meaningful. A prediction (underlined below) that is derived from the hypothesis (italicized below), can be stated in an if-then format as follows:

> If *green apples taste sour because they have not yet undergone a process of ripening in which the starch is converted to sugar*, then <u>the green apple on the shelf will test negative for sugar and taste sour</u>.

The above prediction focuses on finding evidence that will confirm the hypothesis. Shifting to a focus on finding refuting evidence, one might predict:

> If *the taste of the green apple on the shelf is not affected by a process of ripening in which the starch in its cells have been converted to sugar*, then <u>testing a variety of green apples will result in finding one that will test positive for starch, negative for sugar, and taste sweet</u>.

This latter prediction is a powerful prediction. If only one such apple were found, it would be evidence enough to reject and revise the original hypothesis. To adequately test this prediction will require sampling a variety of green apples in various stages of development. In the event that such an apple is not found, one can tentatively accept the original hypothesis. Seeking to find disconfirming evidence is more logical when one considers that finding supporting evidence in favor of a hypothesis can never prove it. According to Popper (1965), science should proceed by this method of conjecture and refutation.

To identify additional characteristics of hypothetico-predictive reasoning, consider how many high-school and college-level try to explain why when a jar is placed over a burning candle in a pan of water the candle goes out and the water rises up. One hypothesis which students usually generate states that:

> As oxygen is used up by combustion of the flame, the water rises up to fill the void.

Another competing hypothesis states that:

> Heat from the candle causes the air in the jar to expand. As it cools, the air takes up less space, exerts less pressure, and the water rises up to equalize the pressure inside with the outside.

These hypotheses seek to explain why the water rises. Hypothetical explanations often take the form of cause-effect relationships.

> Air that exerts less pressure (cause) takes up less space (effect).

> The outside air pressure (cause) forces the water to rise up (effect).

The number of cause-effect relationships used in a hypothesis can become an index of explanatory power and assessment. Further, as the "why" of the hypothesis is elaborated, the greater the utility of the hypothesis for making specific and testable predictions. Students find it very difficult to develop the "why," and will often make quick predictions based more on guess work than reason (Lavoie, 1992a). If asked to develop a hypothesis following the observation of an unusual or discrepant event students will often develop a description of what happened rather than address why it happened. Students also find it difficult to devise predictions that actually lead to evidence in favor of rejecting or accepting their hypothesis. Before reading on, try to develop a prediction which, if it occurred, would provide evidence refuting the above hypotheses.

In this example, students often determine the "amount of burning" as an important variable affecting oxygen depletion and heat in the jar. They contend that as the flame burns it uses up the oxygen in the jar and leaves an emptiness which the water fills. In the case of the "oxygen is used up" hypothesis, a prediction to refute the hypothesis might be: if more than one candle is used, then the amount of water rise will not change. Logically, since there is a finite amount of oxygen in the jar that is used up, it shouldn't matter how many flames it takes to use it up. In the case of the "expanding air" hypothesis, a prediction might follow that increasing the number of candles (i.e., the amount of heat) should cause an increase in the amount of water rise. If the water level rise does increase, the hypothesis would be supported, if it does not, it would be refuted.

This chapter subsequently discusses the importance of hypothetico-predictive reasoning process and provides additional examples and details of the associated cognitive behaviors.

The Importance of the Hypothetico-Predictive Reasoning

Hypothesizing and predicting are considered to be essential processing skills for scientific inquiry and a terminal objective for science education curriculum development (Good, 1989; National Commission on Excellence in Education, 1983; Butts, Capie, Fuller, May, Okey, & Yeany, 1978). Yore (1992) comments on the relationship between the hypothetico-predictive process and the nature of scientific inquiry:

> Theory building is the end-product of the cyclic, self verifying, constructive process of science that uses informed inquiries, shared values, and communal effort to sharpen problems, focus new questions, and formulate more powerful hypotheses. (pg. 16,17)

The thinking processes involved with hypothetico-predictive reasoning are consonant with those critical thinking skills identified as "Habits of Mind" in the recent *Benchmarks for science literacy: Project 2061* (American Association for the Advancement of Science, 1993) which refer to using reasoned arguments to determine the most appropriate explanation or course of action (prediction) when confronted with problem solving situations.

In sum, hypothetico-predictive reasoning establishes the heart of scientific process, progress, and problem solving. It is a fundamental agent facilitating the discovery and justification of scientific theories.

Cognitive-science research suggests that engaging students' hypothetico-predictive reasoning skills can enhance problem solving (Lavoie, 1989a; Bransford, Sherwood, Vye, and Rieser, 1986; Armbruster, 1985), stimulate peer-group discussion (Good et al., 1988), increase student motivation (Minstrell, 1989; Osborne and Freyberg, 1985), reveal prior knowledge (Lavoie and Good, 1988), and facilitate conceptual change (Lawson, Baker, DiDonato, Verdi, & Johnson, 1993; Jones, 1990). Further, Lawson, et al. (1991) contends the ability to generate alternative hypotheses and test them through their deduced consequences (i.e., engage in the hypothetico-predictive reasoning) leads to the acquisition of scientific concepts.

Making hypotheses and predictions in the science classroom often leads to cognitive commitment -- a desire to know if one's predictions are valid or invalid. To find out, students must learn to design "fair" tests and collect data that can refute their hypotheses. Testing one's prediction is both motivational and essential to learning. Of course, science teachers must give students the opportunities to do so.

Hypothetico-predictive reasoning is most powerful when it leads to anomalous data (i.e., predictions which do not support students' existing belief structures). Holland, Nisbett, & Thagard (1986) expands this idea below.

> Unexpected outcomes provide problems that the system solves by creating new rules as hypotheses. Concepts with shared properties are activated... Covariations among salient and goal-relevant stimuli are being learned: new categories are being formed, and implicit statistical parameters are being revised. (p. 69)

Further, Lavoie (1992a) found that high-school students who made incorrect predictions concerning concepts in biology were significantly more motivated to discuss their hypotheses and predictions, ask higher-level questions, develop alternative hypotheses and predictions (i.e., change their theories or cognitive models), and engage in testing their predictions than students whose data confirmed or apparently confirmed their predictions.

The anomalous effect can be explained by Piagetian theory which contends that, when faced with disequilibrium, an individual learns through accommodating and assimilating (i.e., equilibrating) toward equilibrium. To extend Piagetian theory in more contemporary information-processing terms, an individual learns through assimilating new declarative and procedural knowledge structures, and then, accommodating this knowledge with existing procedural and declarative knowledge structures. Additionally, the ability of students to deal with anomalous data which do not support their predictions is dependent upon their prior knowledge and processing strategies for generating alternative hypotheses or explanations as well as the nature and complexity of the anomalous data (Chinn & Brewer, 1993). Further, Yore (person. comm., 1994) notes that students may make an incorrect prediction for at least three reasons: 1) poor predicting ability, but good conceptual understanding, 2) good predicting skill, but poor conceptual understanding, and 3) poor predicting skill and poor conceptual understanding.

Characteristics of the Hypothetico-Predictive Reasoning

In science teaching, students may move into a hypothetico-predictive processing mode following consideration of a higher-level inquiry question posed by the teacher or the student, often in response to observing a discrepant event, that asks for "why" or "how" something happened or exists. For example, a leading question might be, "How are bees able to find their way back to hive following a day of foraging?"

Students may reason hypotheses such as 1H, 2H, and 3H below:

1H). If bees are able to find their way home, then they do so by using the sun as a reference point (i.e., they remember where the sun is as they fly out, and simply fly back in the reverse direction).

2H). If bees are able to find their way home, then they remember visual landmarks as they fly out and follow a visual pattern back to the hive.

3H). If bees are able to find their way home, then they do it by following a hive specific pheromone trail which they give off as they fly out.

These three hypotheses might lead to the following corresponding and refuting predictions, 1P, 2P, and 3P:

1Pa). If bees go out on a cloudy day, then they will not be able to find their way home.

1Pb). If bees are released from their hive in a large greenhouse with a point source of light located on the ceiling, and the location of the light is changed after the bees are at maximum distance from the hive, then the bees should not be able to find their way home.

2P). If bees are taken away from their hive in a black box and let go some distance away, then they will not be able to find their way home.

3P). If bees fly out on a windy day and the wind direction changes, then they will not be able to find their way home.

The above hypotheses and predictions contain several interesting features. First, notice that each prediction suggests an experiment that is testable. In some cases, this testing requires direct experimental manipulation of the system. Experimentally manipulating a system should provide more control by eliminating possible confounding variables. However, this adds additional bias by removing the natural setting. Developing and testing predictions requires a thorough understanding of experimental design, the scientific process, and those confounding factors that cause bias.

Second, note each prediction and hypothesis uses variations of an "If-then" reasoning format. In the case of the hypotheses, the "if" refers to an outcome or result (e.g., the bees do or don't find their way home) and the "then" refers to a conceptual reason why the proposed outcome would occur (e.g., the bees cue on a hive specific pheromone). In the case of predictions, the "if" refers to the conditions of the experimental design (e.g., cloudy day) and the "then" refers to an outcome or result (e.g., the bees do or don't find their way home). Using this if-then format, the more elaborate the "then" part of the hypothesis, the more likely confounding variables will be eliminated and the scientific validity and control of an experiment will be increased. The more elaborate the "if" part of the prediction the easier to develop a good experimental design that will address the hypothesis. Hypothesizing a conceptual reason and the experimental conditions that will test this reason is at the heart of hypothetico-predictive reasoning and of prime importance for doing good science. It seems possible that science teachers can improve their student's reasoning by requiring them to formulate hypotheses and predictions using the if-then formats described above.

Third, taking a Kuhnian and Lakatosian position, each prediction seeks to find a condition that, if true, would tend to falsify its associated hypothesis. It has been previously argued that seeking to falsify one's hypotheses is better science. However, this still has deficiencies. For example, discovering that bees can indeed find their way back home on a cloudy day does not discount the "sun" hypothesis. Bees may still be orienting to the sun since they can perceive ultraviolet light from the sun through the clouds. Further, bees not finding their way home as a result of the black box experiment doesn't necessarily support the hypothesis that they are cueing on to visual landmarks (there may be many other cues involved).

In sum, the ingredients for success of hypothetico-predictive reasoning, be it leading to confirming or refuting evidence, are to: 1) develop well-formulated hypotheses that explain the "why" of a phenomenon using cause-effect relationships, 2) develop associated predictions which directly address the hypothesis in question with the intent of disproving it, 3) develop associated experimental designs that directly test the predictions, and 4) identify and eliminate the greatest number of confounding variables that could bias the results. Science teachers should encourage students to follow up their hypotheses with appropriate predictions -- and when they provide only predictions, to follow those up with causal explanations. The link between hypothesizing and prediction is vital. If a hypothesis does not have an associated prediction that is testable, it is empty conjecture without substance. Without an explanatory reason behind it, a scientific prediction is random guess work.

Hypothetico-Predictive Reasoning as a Rule-Based Information-Processing System

Hypothetico-predictive reasoning can be viewed, in information-processing terms, as a process by which a goal state (explanatory hypothesis and associated prediction that will test the hypothesis) is attained through the application of specific operators (rule-based knowledge relationships), upon initial problem conditions, through a series of steps, within a problem-solving space (see Figure 1). The problem-solving space can be defined relative to a search through the hypothesis space followed by a search through the prediction space. The hypothesis space consists of all the possible variables and cause-effect relationships between those variables (e.g., linkages between concepts, facts, analogies, experiences, etc.) that are associated with a particular scientific phenomena. The prediction space consists of all the possible outcomes resultiing between the initial state and the final state of a particular system as well as all the possible experimental designs and techniques that might be used to test a given prediction. During hypothetico-predictive cognitive processing there must be a constant search and comparison going on within the problem-solving space as the solver attempts to identify causal schemas that best match the problem conditions.

The success of hypothetico-predictive reasoning depends on the nature of the connections between and within procedural and declarative knowledge (Lavoie, 1992). The hypothetico-predictive space would include all possible connections between procedural and declarative knowledge bases. There is increasing evidence that "expert" problem solvers in one domain are very often "novices" in another (Lesgold, 1988). It could be hypothesized that individuals new to a domain lack linkages or relationships between the domain's conceptual knowledge and their own strategic (procedural) knowledge.

In the hypothetico-predictive system, declarative or procedural rules can be considered as production systems (Anderson, 1985), where if a given condition is satisfied, then the associated action is executed. This is analogous to the if-then propositional format discussed earlier. In the case of the hypothetico-predictive reasoning process, the "if" (hypothesis) part becomes the "condition" and the "then" (prediction) part becomes the "action." Holland et al. (1986) make the following

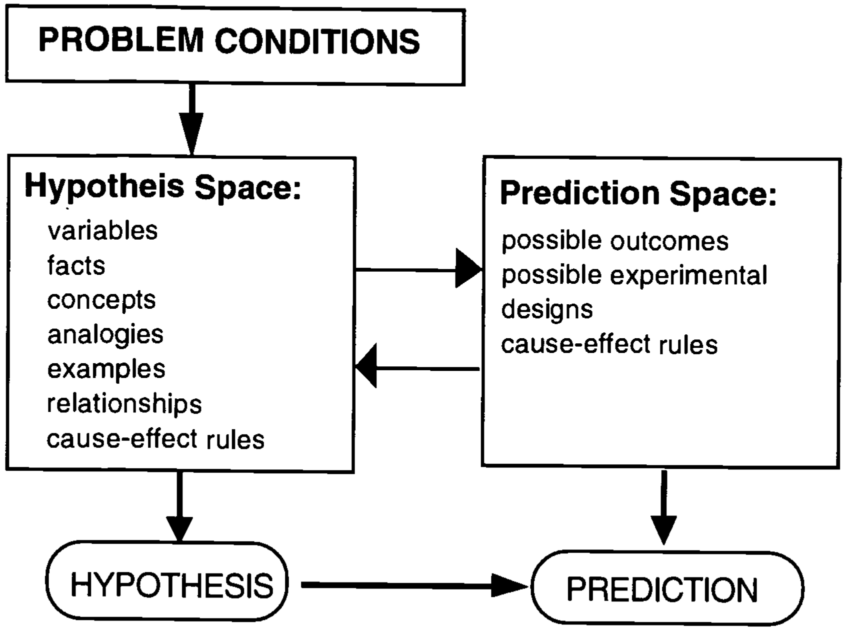


Figure 1. An information-processing view of hypothetico-predictive reasoning.

comment:

> Condition-action rules are obviously well suited for making predictions. A rule that leads to a successful prediction should be strengthened in some way, increasing the likelihood of its use in the future; one that leads to error should be modified or discarded. Predictions about the attainment of goals will normally be the most powerful source of feedback. (p.16)

Holland et al. (1986) goes on to identify three types of production rules.

> Empirical rules are the "bread and butter" performance rules of the system. They describe the environment and its likely next states. Inferential rules provide relatively domain-independent procedures for altering the general knowledge base. Operating principles are innate system-manipulation procedures. (p. 40)

Relative to hypothetico-predictive reasoning, empirical rules represent declarative knowledge. Inferential rules, which develop and alter the empirical rules to particular cases, represent procedural knowledge. The operating principles could be viewed as inferential production rules that direct the formation of linkages between and within procedural and declarative knowledge. Thus, effective use of hypothetico-predictive reasoning depends on the application of all of these rules which lead to the formation, association, and modification of procedural and declarative knowledge.

The hypothetico-predictive information-processing system grows and improves as procedural knowledge acts to apply declarative knowledge to the problem conditions. In fact, procedural knowledge can only be compiled through the application of declarative knowledge (Anderson, et al. 1990). This knowledge compilation process involves developing new rules (e.g., knowledge relationships) and modifying existing rules based on feedback relative to the success of such rules. Holland et al. (1986) comment that:

> Revision of rules based on their predictive success is the key to human's ability to learn complex categories with minimal feedback. (p. 76)

Holland et al. (1989) consider additional complexities of how rules influence hypothetico-predictive reasoning:

> Rules are a natural vehicle for what we take to be the most fundamental learning mechanism: prediction-based evaluation of the knowledge store... There must be mechanisms that evaluate candidate structures, discarding some, storing others, and modifying those that already exist. The evaluation mechanism compares the predicted consequences of applying a knowledge structure with the actual outcome. (p. 16)

Thus, hypothetico-predictive reasoning is dependent upon knowledge generation, knowledge testing, and knowledge modification based on feedback. This must involve procedural cognitive-processing behaviors that search for, identify, select, define, apply, induce, deduce, and evaluate declarative knowledge. This must also involve a continual interplay between specialization and generalization.

> The need for more accurate prediction favors the addition of further specialized rules, whereas the need for efficient prediction favors the addition of general rules to replace a large number of specialized rules. (Holland, et al. 1989, p. 36)

How the information-processing system actually executes its decisions relative to knowledge manipulation is probably parallel in some cases and sequential in others. Holland, et al. (1989) comment on this type of inductive-deductive processing.

> Such restructuring implies that search takes place not only in the space of potential 'next states' along a temporal dimension but also through a space of alternative categorizations of the entities involved in the problem. This type of processing depends on the parallel activity of multiple pieces of knowledge that both compete with and complement each other in revising the problem representation. (p. 12)

It seems likely that the parallel nature of hypothetico-predictive reasoning does not start out parallel, but is formed through trial and error search for an appropriate sequence. Once a successful sequence or strategy is found, the system begins to use it and re-use it as long it remains reasonably successful. When a pathway has been well trodden it may become automatic, tacit, and "chunked." Experts are able to chunk knowledge as cause-effect sequences dependent upon the problem-solving goals at hand (Martin & Szabo, 1990). This ability of the brain to categorize and routinize itself allows it to deal with greater complexity and the coordination of many actions at once (i.e., parallel processing). However, this may reduce the ability of the system to respond in novel ways to generate problem solutions.

The alternative conception literature would argue that procedural-declarative rules are not easy to modify, but tend to persist in the face of competing "new" rules. Students who believe that plants get their energy from the soil will often still believe this after conducting experiments of plants in light and dark. Why individuals cling to particular ideas in the face of contradictory evidence is an interesting question that is not fully addressed by the conceptual change literature. The brain seems to have a remarkable tendency toward persistence and habit. Such persistence should facilitate the chunking of information and increase available space in short-term memory. Behavioral persistence can be explained in information-processing terms in that the more a behavior or pathway is taken the greater the synapse strength of the associated neurons (Anderson, 1992). The greater the synapse strength the greater the likelihood the pathways or behaviors will be executed again given similar conditions.

To change conceptions or cognitive processing behaviors, students must discover that competing alternative conceptions and behaviors are better able to explain and manipulate their world (Posner, Strike, Hewson, & Gertzog, 1982). And, this requires repeated cognitive efforts to apply such alternative conceptions and behaviors to problem situations. Students seem all too often unmotivated to challenge personal beliefs and develop alternative ideas or try out new behavior patterns.

The ability of an information-processing system to build and modify itself is directly related to its cognitive flexibility. From a cognitive-science perspective, cognitive flexibility increases with a greater number of possible solution paths which provide greater access to a relevant procedural and declarative knowledge bases. Cognitive flexibility in information-processing terms would relate to the number of condition-action rules available and the procedural capacity to modify based on feedback. Modification of a hypothesis, for Holland et al. (1989), involves generating new condition-action rules or changing the strength of existing rules based on predictive success. Hypothesizing, in general, can be defined as a process of rule revision based on predictive success.

Thus, cognitive flexibility implies an ability to shift from one conceptualization to another -- and it follows, the ability to shift from one cognitive model, relationship, rule, etc., to another. Lakoff (1987) provides a perceptive example below.

> In fact, learning to become a scientist requires alternative conceptualizations for scientific concepts. Take electricity for example....there are two prevalent ways to metaphorically understanding electricity: as a fluid and as a crowd made up of individual electrons. Knowing how to solve problems in electrical circuitry involves knowing which metaphor to use in which situation. (p. 305)

A crucial element for hypothetico-predictive problem-solving success, and flexibility, is the organization of procedural and declarative knowledge within a cognitive network. This organization is directly dependent upon how the new knowledge is identified, categorized, and connected within already existing knowledge of the cognitive network. Many have argued that prior knowledge is extremely important variable affecting teaching and learning (Ausubel, 1986; Carey, 1986; Finley, 1985; Novak, 1977). Models that represent the structure and function and organization of knowledge used for hypothetico-predictive reasoning are discussed below.

Models of the Hypothetico-Predictive Reasoning

A basic model involving the hypothetico-predictive reasoning process portrays hypothesizing and predicting as part of the scientific method. This model typically involves the following sequential components:

question --> hypothesis --> prediction -->
data collection --> data analysis --> conclusion

This model, typical of science textbooks, is not representative of actual scientific problem solving. In the "real" world scientists follow a model that is much more cyclic with many feedback loops tied to hypothetico-predictive reasoning (Figure 2).

Hypothesizing and predicting establish the basis of wonderment in science as we ask questions and seek out answers about our world. These processing skills also develop scientific explanations and determine experimental design (i.e., determine how data are to be collected and how data are to be interpreted). Based on such interpretation (e.g., comparing actual results with predicted results), hypotheses are often supported, rejected, or modified. This leads to new predictions and the cycle repeats. Thus, science students and scientists should base their understandings of scientific phenomena on not just one hypothesis and prediction but many sequential hypotheses and predictions that involve testing, modifying, and testing again.

Very few models exist for hypothetico-predictive reasoning, per se. Those that do, attempt to define discrete and sequential cognitive steps. Van Joolingen and Jong (1992) determined sub-processes of hypothesis formation to involve identifying variables, selecting variables, and defining the relation that is hypothesized between selected variables. Holland et al. (1989) identified the following steps for engaging in inductive rule-based problem solving (i.e., hypothesizing and predicting):

(1) compare and evaluate one's model with the goal state,
(2) select appropriate rules to apply to the problem condition,
(3) identify associations or connections between and clusterings among rule-based schema, and
(4) evaluate one's model with the goal state based on feedback.

Lavoie (1993a) analyzed students' think-aloud cognitive behaviors while solving hypothetico-predictive problems in biology and identified similar sub-process as above as well as several more sub-process of hypothetico-predictive reasoning (Table 1). During the problem-solving process, actual hypothetico-predictive problem solving involves multiple back-an-forth execution patterns that follow both linear (sequential) and non-linear patterns. As the problem solver moves through the problem-solving space a series of decision are made based on previous decisions and conditions. These decisions have been modeled visually as a tree with many branches and sub-branches. Such trees are used to construct intelligent expert systems (Randolph, 1988). Other attempts to describe complex procedural thinking processes have used flow charts (Yore, 1992), dendrograms (Anderson & Demetrius, 1993), frames (Minsky, 1986), scripts (Schank and Abelson, 1977), and schemata (Rumelhart, 1980). Declarative knowledge has been modeled as concept maps (Fisher, 1990; Malone & Dekkers, 1984; Novak, Gowan,

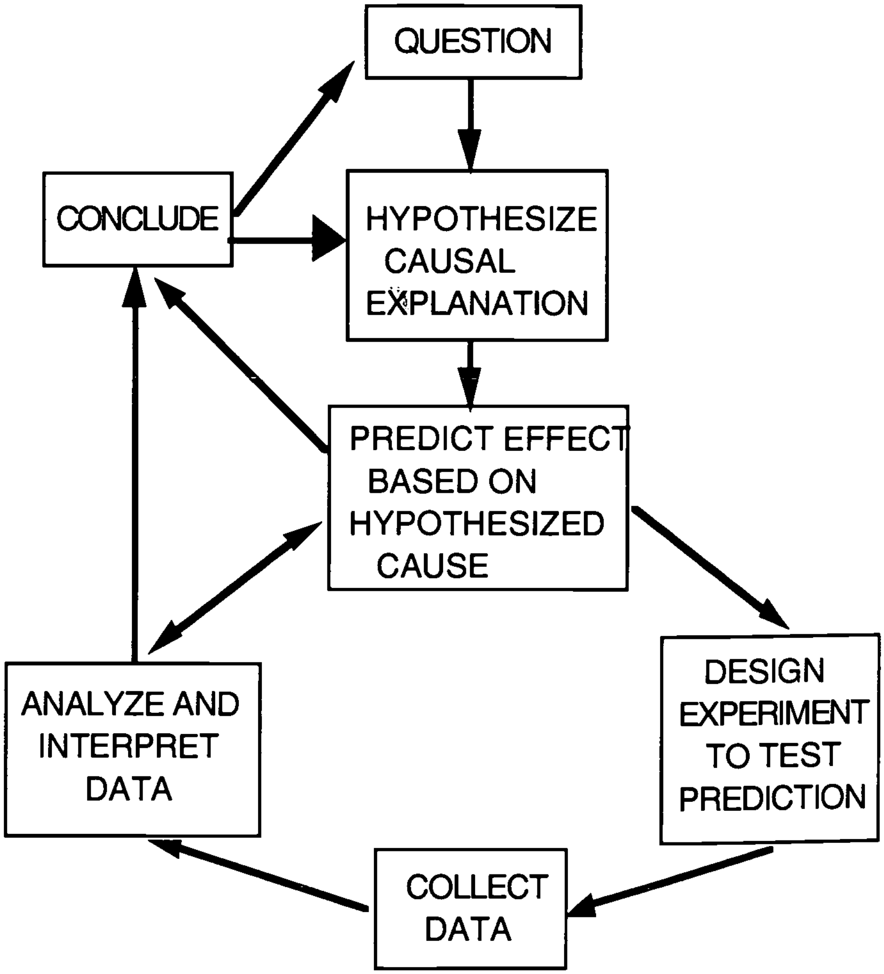


Figure 2. Cycle of scientific investigation showing the central importance of hypothetico-predictive processing.

& Johansen, 1983), word associations (McRobbie, 1991), or simply a list of facts as is common of many science textbooks. Lavoie (1993a) used a dynamic flow chart to model procedural knowledge, declarative knowledge, and the relationships between procedural and declarative knowledge.

During problem solving what is it that causes a particular decision to be made (i.e., a particular rule to be applied, a particular cognitive behavioral sequence to be followed, or a particular branch path to be

Table 1. Cognitive behaviors of hypothetico-predictive reasoning.

1). Identify all information given in the problem situation (terms, conditions, variables, facts, etc.).

2). Identify all goals and sub-goals associated with the prediction problem.

3) Identify any assumptions associated with the problem.

4). Identify and define other relevant terms, variables, and facts not given in the problem in as much detail possible.

5). Using if-then logical reasoning, identify and define cause-effect relationships relevant to solving the problem.

6). Evaluate and modify the relevant cause-effect relationships.

7). Identify and reflect on errors and inconsistencies in logic <u>or</u> knowledge.

8). Identify any gaps (questions, needed information) in the identified knowledge. Identify and take the necessary steps to acquire such information.

9). Identify any relevant experiences you have had that could help solve the prediction problem.

10). Identify any analogies and examples that are relevant to solving the prediction problem.

11). Apply the identified cause-effect relationships to put forth several possible predictions.

12). Evaluate and modify predictions.

13). Select the most probable prediction.

14). Reflect back on the reasoning used to put forth a final prediction.

taken)? Factors affecting this decision involve contextual relevancy and the recency of use (Anderson & Milson, 1989). Contextual relevancy could be defined as the degree of match between the problem-solving conditions and the mental model an individual possesses of the problem solution. Factors that affect such a match probably include problem type, subject-matter knowledge, problem cues, availability of appropriate models, problem-solving strategy, etc. Recency is simply the frequency with which a rule, behavior, or pathway has been applied in the past. The most recently executed is most likely to be used first for the next similar type of problem conditions. In information-processing terms, this likelihood or recency can be explained in terms of a statistical probability -- a probability that is ultimately determined by the firing strength of the associated neurons (see Anderson, 1992; Lawson, 1986).

In sum, the hypothetico-predictive information-processing mechanism involves the application of condition-action production rules to inductively generate hypotheses and deductively generate predictions to test the hypotheses. Such rules are developed and categorized by establishing connections between and within procedural and declarative knowledge bases. This results from procedural knowledge applying declarative knowledge to hypothetico-predictive problems. Engaging in hypothetico-predictive reasoning implies that procedural and declarative knowledge bases undergo many cycles of rule-based synthesis, application, adaptation, and modification. Understanding how procedural and declarative hypothetico-predictive knowledge is organized, connected, and structured is essential for devising effective teaching strategies which are subsequently discussed.

Teaching Hypothetico-Predictive Reasoning

A plethora of research on problem-solving skills provides many ideas concerning how such skills should best be taught (Pestel, 1993; Bunce, Gabel, & Samuel, 1991; Whimbey & Lochhead, 1991; Barba, 1990; Spiro & Jehng, 1990; Dunkhase & Penick 1990; Gabel, 1990; Quellmalz, 1987; Greeno & Simon, 1986; Simon, 1980; Tuma & Reif, 1980). A common constructivist/information processing theme arising from this research is that students should be provided with opportunities to practice solving personally relevant problems that require the application of their declarative knowledge in a variety of increasingly complex ways and in a variety of increasingly complex contexts. Such application of procedural knowledge to declarative knowledge should lead to greater "proceduralization," a process which builds domain-specific declarative knowledge into productions (Anderson, 1987). Constructivist problem solving and constructivist teaching strategies allow students to take an active role in acquiring new skills and knowledge.

In the case of hypothetico-predictive problem solving, variety can be achieved by providing students with different types of problems to solve. Lavoie (1991) identified different problem types associated with hypothetico-predictive reasoning based on cause-effect temporal

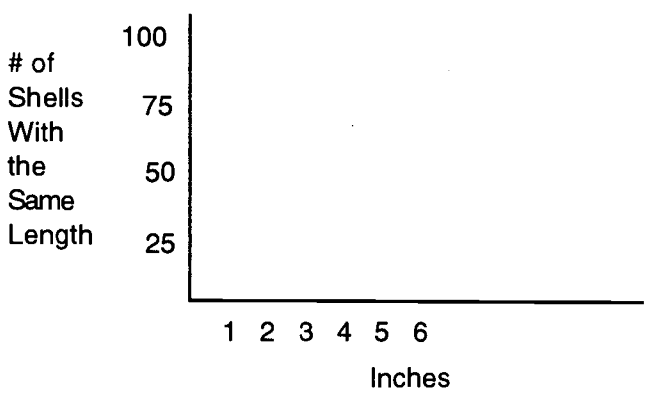
Table 2. Sample prediction problems from biology (from Lavoie, 1993a).

Inheritance

You decide to conduct an experiment. You select the following easily identifiable two-character traits (e.g., tongue roller or not a tongue roller; brown-green eye color or blue eye color; earlobes attached or unattached, left or right handedness, hair on the mid-digit of the fingers hair not on the mid-digit of the fingers). Using the entire school as your sample, you determine the total number possessing each trait. For each trait, *predict* the character that most members of the school will have in common. Please be sure to explain your reasoning.

Variation

One day at the beach you collect a bag of 100 seashells, all of which are the same species of clam. Most appear to be between 2 and 6 inches in length, however, you decide to measure the length of each clam to determine how many clams have roughly the same length. You plot the data as the number of clam shells having the same length. *Predict* the pattern of variation you expect to find on the graph below and explain your reasoning. Please write clearly and completely.



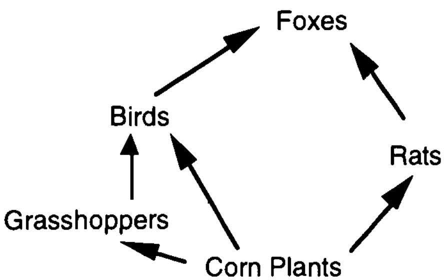
Physiological Responses

You are interested in finding out how light affects the pupil of the human eye. You decide to shine light on only one eye of a friend, and shield the other eye. *Predict* how the pupil of each eye of your friend will respond. Explain your reasoning.

Table 2, continued. Sample prediction problems from biology (from Lavoie, 1993a).

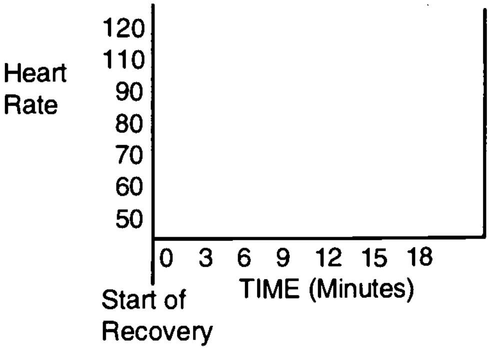
Ecology

Predict what will happen in the following food web if a fatal disease wipes out the birds. Explain your reasoning.



Homeostasis

A recovery period is the length of time it takes for the heart rate to return to a normal rate. *Predict* your own recovery curve the graph below following a 10 minute jog. Explain your reasoning.



patterns (e.g., rate changes; size changes) and conceptual relationships between variables (e.g., foxes eat rabbits). Table 2 provides examples of different types of problems. Variety also implies that students should be provided opportunities to apply or transfer hypothetico-predictive cognitive-processing behaviors across domains to multiple contexts. The often used technique of "do the odd problems at the back of the chapter," while providing needed practice, is largely inadequate for building effective problem-solving skills.

In addition to general problem-solving strategies, the hypothetico-predictive reasoning requires that students: 1) actively formulate problems and questions, 2) identify, analyze, and apply relevant information, 3) develop and carry out goals, and 4) develop, evaluate, and modify product and process knowledge. The entire process must be undergirded by metacognitive monitoring and self-regulation, which is critical to improving performance for any problem-solving skill (Brown, 1978). Thus, hypothetico-predictive reasoning is a constructivist process that requires students to organize and structure procedural and declarative knowledge. This, in turn, influences the accessibility and applicability of such knowledge while engaged in the problem-solving process.

Prerequisite Knowledge

Constructivist/information-processing teaching strategies directed at improving cognitive skills such as hypothetico-predictive reasoning require that the science teacher possess several types of knowledge (Figure 3). First, the teacher must possess a thorough understanding of the conceptual (declarative) knowledge base associated with the particular content domain. Conceptual knowledge implies multiple connections between related bits of information associated with particular concepts. These information bits may include facts, ideas, experiences, feelings, analogies, examples, etc. Information stored as cause-effect (if-then) relationships is particularly useful for hypothetico-predictive problem solving (e.g., if the temperature of a gas increases in a confined space, then the pressure of the gas increases).

Second, the teacher's conceptual declarative knowledge should be firmly interconnected within and between his/her procedural knowledge. This interconnectedness is a primary factor affecting the accessibility and applicability of knowledge for use in problem solving. To acquire declarative-procedural knowledge requires extensive attempts to understand the concepts through exposure to the concepts, and attempts to apply the concepts, in a variety of contexts (e.g., problem solving, hands-on activities, analogies, etc.).

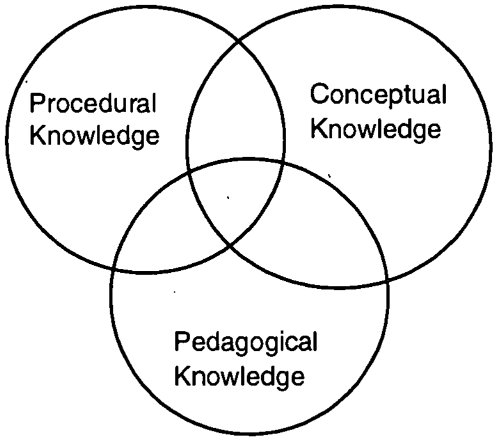


Figure 3. The relationships between different types of knowledge required for effectively teaching problem solving.

Third, the teacher should have pedagogical strategic (procedural) knowledge of successful and unsuccessful cognitive behaviors used for hypothetico-predictive problem solving within a given domain. Experts teachers possess schemata that are more elaborate, interconnected, accessible, and successful than novices (Livingston & Borko, 1989). This allows the "expert" to select appropriate routines, knowledge, and strategies that are most appropriate for helping students learn. Lavoie (1993a) found that successful hypothetico-predictive problem solvers possessed cognitive networks with a greater number of pathways and relationships than did unsuccessful problem solvers. Successful hypothetico-predictive problem solvers used procedural knowledge which defined, evaluated, and applied conceptualized declarative knowledge that was scientifically sound. Successful hypothetico-predictive problem solvers also were better able to plan and reflect on their own problem-solving process. In general, experts were more flexible than novices -- being able to change and adapt their strategies to a variety of problem types and conditions.

Pedagogical strategic knowledge allows the teacher to make appropriate pedagogical decisions about a students' learning such as when to ask appropriate questions, when to provide appropriate feedback, how to structure the lesson sequence, and how to arrange the content of the lesson. Pedagogical knowledge is typically derived from extensive experience working with students as they attempt to solve problems in a given domain. The think-aloud technique combined with qualitative analysis can be used to effectively facilitate the acquisition of pedagogical knowledge. Students' verbalizations are recorded on videotape or audio tape as they think out loud during problem solving. Cognitive behaviors are then identified, categorized and sequenced into behavioral steps or models of the problem-solving process. The behaviors in Table 1 were derived in this way. A teacher having a good understanding of such steps and models should be able to more effectively offer suggestions and ask questions concerning students' problem-solving sequence and problem-solving behaviors. This would allow the identification of scientific misconceptions, processing errors, and unproductive heuristics. Pedagogical knowledge increases the teacher's awareness of the variety of successful ways to solve problems and, it follows, to engage in hypothetico-predictive reasoning.

Finally, the science teacher needs to have pedagogical content knowledge of how to present content (declarative knowledge) so it is most effectively linked with cognitive skills (i.e. hypothetico-predictive reasoning skills). Shulman (1987) refers to structuring the lesson and organizing its content to suit the needs of the learner as pedagogical reasoning and pedagogical-content knowledge, respectively. For hypothetico-predictive instruction, organizing its content could translate to developing different formats for hypothetico-predictive problems (refer back to problem types mentioned earlier) and to determining the amount and structure of the information to be provided to the student either in preceding instruction or in the problem. For example, facts may be presented in isolation or with multiple connections between them. Further, concepts may be presented in cause-effect (if-then) formats. Structuring the lesson may involve several of the strategies discussed below.

Instructional Strategies

Examples of constructivist teaching strategies, which involve hypothetico-predictive reasoning as an essential component, include conceptual change methods (Kyle & Shymansky, 1992; Watson & Konicek, 1990; Hollingsworth, 1989; Osborne & Freyberg, 1985)), the leaning cycle (Abraham and Renner, 1986; Lawson, Abraham, & Renner, 1989), and the generative learning model (Osborne & Wittrock, 1983). Sub-strategies which can be used with the above strategies to more directly focus on hypothetico-predictive processing skills include combinations of cognitive apprenticeship (Collins, Brown, & Newman, 1989), think-aloud problem solving (Barba & Rubba, 1992), systematic modeling (Rubin & Norman, 1992), and explicit cognitive-strategy training (Symons, Snyder, Cariglia-Bull, & Pressley, 1989; Wong, 1989).

The goal of each of the above strategies is to restructure the learner's schema so as to improve problem-solving ability. The focus is not on the final answer (i.e., prediction) but on the process by which to reach the answer. These strategies facilitate hypothetico-predictive processing by stimulating the use of cognitive behaviors such as identifying, organizing, and applying relevant facts, causal-relationships, experiences, analogies, and examples to the problem situation.

Cognitive apprenticeship. Cognitive apprenticeship is a model of teaching and learning in which an expert (teacher) coaches a novice (student) attempting to solve a problem (Farnham-Diggory, 1992; Collins, Brown, & Newman, 1989)). Following demonstration or modeling of expert problem solving by the teacher, the student is given simple problems to solve. Based on questions posed to the student, observations of their behavior, and questions asked by the student, the teacher provides hints and supports (scaffolding) which the student needs to solve the problem. This could include discussion of alternate problem-solving strategies, multiple forms of representing declarative knowledge associated with the problem, and alternate conceptualizations of the problem (Roth, 1991). Relative to hypothetico-predictive reasoning, the cognitive coach monitors the students' problem-solving sequence and behaviors, responds to students questions, and provides feedback relative to the use of appropriate procedural or declarative knowledge. In subsequent attempts to solve problems the teacher moves the student toward increasingly complex problems while trying to reduce the need of the student to rely on outside assistance (fading). To facilitate feedback to the teacher the student can be required to think out loud as they solve hypothetico-predictive problems.

Group problem solving. While most classroom teachers do not have the time to analyze recordings of students' think-aloud behaviors, there are alternatives. For example, students can become cognitive coaches through pair or small-group cooperative strategies. Pair problem solving might involve an observer recording on paper the cognitive process behaviors of a solver while they think out loud during problem solving. Then both solver and observer analyze the problem-solving strategy with the intent of improving it -- identifying gaps in knowledge, identify successful and unsuccessful cognitive behaviors, outlining the sequence, etc. The teacher can look over student notes to gain greater insight into their cognitive behaviors and problem-solving strategies. Requiring students to "write it down" also increases their accountability. Lavoie (1992b) tested a pair problem-solving method for hypothetico-predictive reasoning and found it to significantly improve students' problem-solving success as well as the quantity and quality of their procedural processing behaviors. Behaviors that the observer and solver could use to facilitate pair problem-solving success were also identified with this technique.

Small groups of students might interact in a variety of ways to solve problems. For instance, a team of four students might be assigned to solve a problem individually while recording the cognitive processes they are using in as much detail as possible. Then, the group gets together and shares their ideas for the purpose of synthesizing a problem-solving model. Another group method might have student teams collaborate on a problem solution until every member of the group can explain the solution process (Kagan, 1990). The teacher then picks a student at random to explain the process. If s/he can do so, the entire team gets points. Pea (1993), as cited in Yore (1994), provides an interesting look at the dynamics of a cooperative learning situation:

> Imagine a classroom. But instead of having a teacher in front of 30 students, imagine small groups using artefacts, such as optical devices including mirrors, light sources, lenses, and a computer tool kit that lets one build dynamic models of different optical situations. imagine the students talking animatedly with one another, comparing predictions and arguing about how to frame and solve problems by creating simulations of optical situations established with hands-on materials. And they are interacting with other groups. The teacher is an additional resource and interpreter who comes around and who the students may request information from when they feel blocked in their inquiries. (p. 267)

Systematic modeling. Systematic modeling is the teacher-centered version of group problem-solving strategies and cognitive apprenticeship described above. The teacher leads students through the cognitive processing involved with solving a problem by thinking out loud while he/she solves a problem. The students presumably learn the cognitive processing involved by following and imitating the teacher's behaviors. To increase effectiveness of this strategy, students should record the teachers' cognitive steps and then discuss them with themselves and the teacher. Such reflection should result in better overall perspective on when, where, and why to apply particular cognitive actions. Further, hypothetico-predictive processing may be improved by simply encouraging students to verbally use "if-then-because" words while attempting to solve hypothetico-predictive problems.

Explicit cognitive strategy training. It seems probable that procedural knowledge has a strong generic character in a similar manner that the generic tools can be used to fix the car, the kitchen sink, or build a house. The same generic questions arise when encountering any problem; what tool do I use, where do I begin, what sequence should I follow, how do I test what I've done, etc.? If procedural knowledge has domain-free qualities, then strategies for teaching such skills should also be relatively domain free. Anderson et al. (1990) believe that the basic learning principles are domain free. Research is needed to examine the degree to which domain-free procedural knowledge can be acquired and transferred. It could be hypothesized that those possessing generic problem-solving skills but lacking domain knowledge will become more proficient at problem solving in a new domain faster than someone who lacks both general skills and domain knowledge.

One way to help students acquire problem-solving success is to explicitly provide them with the cognitive behaviors (i.e., the tools) necessary for solving problems in a given domain. Cognitive-strategy research has supported that problem-solving instruction needs to be more explicit (Pressley, Goodchild, Fleet, Zajchowski, & Evans, 1989; Symons et al., 1989; Nickerson, 1987; Segal, Chipman, & Glaser, 1985; Larkin, 1980). Sternberg (1983) recommended explicit instruction in both executive and non-executive information processing. Bransford (1990) and Prawat (1989) demonstrated the transfer of strategic procedural knowledge when subjects were given specific hints and directions.

Several explicit problem-solving strategies have been developed which seem to extend the systematic modeling strategy described above. These cognitive training strategies break the problem-solving process down into a series of steps to be learned. The Good Strategy User Model (Pressley, et al. 1989) involves direct explanation of strategies as the instructor behaviorally models a series of steps followed by practice with feedback. A teacher using the Strategy Intervention Model (Deshler and Schumaker, 1986) first describes a cognitive strategy to students then models through it while thinking out loud. Students gain skill with the strategy through teacher led demonstrations followed by guided practice with corrective feedback in which students must verbally rehearse each step or behavior until it is mastered. The Training Arithmetic Problem-Solving Skills (TAPS) strategy focuses on improving students' abilities to solve mathematical word problems (Derry, 1989). Students learn to apply strategic techniques such as forward, backward chaining, and metacognitive monitoring. This strategy emphasizes diagnosing problem-solving performance to identify existing behaviors that should be encouraged, existing behaviors that should be eliminated, and new behaviors that should be acquired.

The Explicit Prediction Teaching Strategy (EXPRTS), developed by Lavoie (1993b), provides the cognitive tools that students can use to facilitate the hypothetico-predictive problem-solving process. To implement EXPRTS involves explicitly teaching students when and how to apply the cognitive behaviors identified in Table 1 to hypothetico-predictive problems. Lavoie (1993b) provided students with explicit behavioral sheets which listed each behavior in sequence. Students were required to elaborate their knowledge for each behavior, in writing, before making a final prediction. Van Joolingen and De Jong (1992) used a similar type of structured scratch pad to facilitate hypothesis formation. EXPRTS was shown to significantly improve problem-solving performance relative to students' predictive accuracy, logical reasoning, and the identification and application of successful cognitive scripts when compared to a non-explicit (traditional) strategy. EXPRTS forces students to access what might otherwise remain "inert" prior knowledge.

EXPRTS encourages students to generate several testable hypotheses and to make decisions regarding competing hypotheses. Evaluating competing hypotheses is a good exercise in critical thinking. It helps students to identify with the origins of their beliefs and to examine the relevancy, legitimacy, and full extent of their knowledge. Learning with EXPRTS should be enhanced if students are encouraged to debate their personally developed hypotheses and predictions. To do so, students must think carefully about how to logically present their commitment to a particular hypothesis and to look for inconsistencies in the hypothesis of their competitors as well as their own.

EXPRTS could become more constructivistic by having the students determine their own cognitive behavioral sequence for problem solving, perhaps using one of the group problem-solving strategies discussed earlier. Once determined, hypothetico-predictive sheets could be developed and used on various types of hypothetico-predictive problems dealing with different contexts. Further, having students determine their own contexts within which to embed the problem-solving task would increase its relevance and motivation.

Assessing Hypothetico-Predictive Reasoning

Dynamic instructional methods demand dynamic assessment methods. This section will briefly address assessment of students' performance relative to hypothetico-predictive processing. Assessment is viewed as diagnostic, formative, and summative. In each case, evaluating students' information processing mechanisms is the primary goal, taking precedent over the traditional focus on the final answer (e.g., prediction is validated).

A teacher employing diagnostic assessment could assign students hypothetico-predictive problems prior to instruction requiring the application of yet to be studied concepts. Lavoie (1992a) used hypothetico-predictive sheets to assess students' hypothetico-predictive reasoning. Each sheet consisted of a problem statement followed by a request that the student make a prediction and then provide an explanatory reason (hypothesis) for his/her prediction in writing (Figure 4). Lavoie (1992a) found hypothetico-predictive sheets increased students' feelings of accountability and led to the identification of a greater number of cognitive behaviors compared to students not using the sheets. Another method of assessing students' cognitive skills could involve assessing students' think alouds either individually or in

POPULATION DYNAMICS

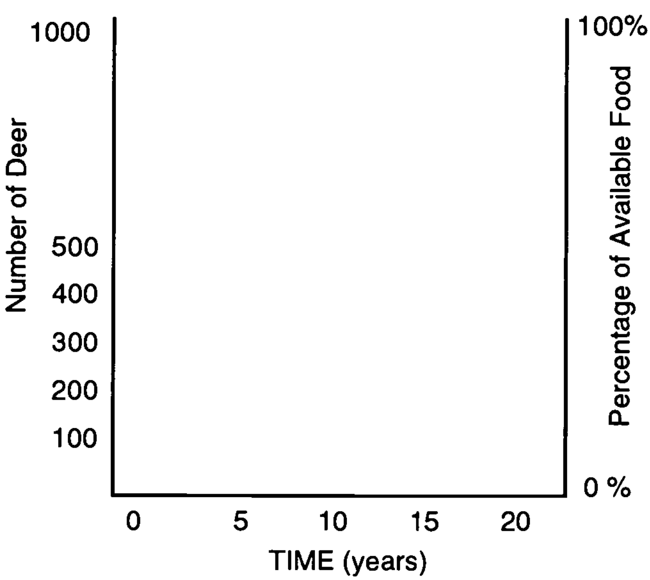
Prediction Problem Sheet **Name:** ____________

Problem One:

Congratulations! You just landed a job as Environmentalist I for the Department of Natural Resources. Your first major task is to introduce 20 deer (10 males and 10 females) into a game preserve in Northern Michigan. It has been estimated that the food plants in the preserve could maintain a unchanging population of 500 deer per year. Females normally produce one offspring per year and it takes about 2 years for offspring to reach reproductive maturity.

Predict how the population of deer and amount of available food plants will change over the course of 20 years, and fully explain your reasoning. Let a straight line represent the population of deer, and a dotted line the amount of food plants available. Begin both predictions at the left side of the graph and proceed over time.

Food Available .... # of Deer ________



Fully Explain Your Reasoning Below:

Figure 4. Sample prediction problem sheet (from Lavoie, 1992a).

pairs. Lavoie (1992b) used a pair problem-solving method to improve pre-service secondary science methods students' hypothetico-predictive processing. Students were assessed on their degree of peer interaction and number of cognitive behaviors which they identified and were able to apply to the problem solution. Models and cognitive behavioral sequences developed by students individually or in groups could be analyzed relative to Table 1 and other criteria that is yet to be determined. Further, visual schematics of the thinking process, such as the cognitive-network model developed by Lavoie (1993a), would seem to have important applications as assessment tools. Perhaps students could be trained to explicitly develop and assess their own cognitive-network models.

Formative assessment of hypothetico-predictive reasoning should be viewed as an on-going process concerned with collecting data and providing feedback for modification of students' cognitive skills. Formative techniques would include various combinations of self monitoring, peer monitoring, and teacher monitoring strategies followed by feedback and modification of hypothetico-processing behaviors and sequence.

Summative assessment can be viewed as a final formative assessment designed to show the extent to which students have developed their hypothetico-predictive reasoning abilities. This assessment should focus on the procedural-knowledge processing behaviors and the associated declarative-knowledge base that students use for developing hypotheses and predictions as well as the validity of those hypotheses and predictions.

Future Research

This chapter has discussed cognitive-science aspects of hypothetico-predictive reasoning relative to logical format, information-processing architecture, and strategies for teaching and assessing this important skill. Several research questions and directions can be identified for each component.

Logical Form

The format for hypothetico-predictive reasoning was explained in terms of if-then conditional logic. Conditional or propositional logic is represented by a combinatorial system which leads to the determination of causality based on the relationships between variables and outcomes. For example, an "if p, then q" system of propositional logic will produce 16 combinations involving the relationship between p and q (Good, 1977). Examination of the rules and combinations of conditionalized (propositional) logic could lead to the identification of mini-strategies for the application and evaluation of hypothetico-predictive reasoning. Lawson (1990) has identified several forms of conditional logic relative to hypothetico-deductive processing which should also be applicable to the hypothetico-predictive reasoning process. Lastly, Piaget's logical operations of identify, negation, reciprocal and correlative (i.e., the INRC group), which also explains relationships between variables and outcomes, should be applicable to if-then hypothetico-predictive reasoning.

Information-Processing Architecture

Research must continue to explore methods of representing and explaining the structure of hypothetico-predictive processing knowledge. In particular, research needs to develop and test operational models that represent the relationships between procedural and declarative knowledge. To develop such models requires continuing efforts to identify the details of how students build production systems and knowledge networks as they acquire and apply hypothetico-predictive processing skills?

Teaching and Learning

Research needs to examine the impact of a variety of variables affecting the teaching and learning of hypothetico-predictive reasoning skills. For example, how is the acquisition of procedural or declarative knowledge, associated with hypothetico-predictive problem solving, affected by students' prior conceptualized knowledge, prior proceduralized knowledge, learning styles, and cognitive preferences. What are the effects of varying problem context, format, and complexity? What mechanisms lead to the formation of linkages between procedural and declarative knowledge? What factors allow teachers to acquire the necessary prerequisite knowledge for training students to use hypothetico-predictive reasoning skills? If such factors and mechanisms can be identified, then it should be possible to develop teaching/learning strategies that can insure the optimal growth of hypothetico-predictive reasoning skills.

This chapter has pointed out that hypothetico-predictive reasoning, by its nature, leads to the manipulation and application of declarative knowledge (i.e., science concepts). Research needs to develop and examine strategies by which hypothetico-predictive reasoning can be used to more effectively teach and learn science concepts? This chapter has also suggested that hypothetico-predictive processing skills have a strong generic nature. Research should examine the extent to which such skills can be transferred from one problem-solving context to another and from one subject domain to another. Identifying factors that facilitate or hinder such transfer should result in more effective instruction.

Finally, this chapter has identified a variety of teaching strategies for hypothetico-predictive reasoning, many of which have not been formally tested in science teaching classrooms. Subsequent studies should continue to develop, test, and compare these strategies with others, yet to be developed, under a variety of conditions. For example, can hypothetico-predictive problem solving be explicitly taught to high school, middle school, or elementary-level students?

Conclusion

This chapter has explored the nature of hypothetico-predictive reasoning, which is considered to be an integral component of scientific problem solving, from a cognitive science information-processing perspective. This mechanistic view of learning and teaching attempts to open the "hood" and inspect the "engine" to see how it runs -- be it a car, brain, or hypothetico-predictive reasoning. Logically, understanding how information is developed, related, and applied during hypothetico-predictive reasoning should lead to the development of effective hypothetico-predictive teaching strategies.

The hypothetico-predictive reasoning process was shown to depend on a cognitive network of if-then (cause-effect) production rules that arise between and within the declarative and procedural knowledge bases. Hypotheses and predictions are developed, tested, and modified via procedural rule-based behaviors that identify, structure, and apply declarative knowledge. The greater the number of connections established between and within declarative and procedural knowledge bases the greater the system's cognitive flexibility and information processing power -- essential elements for conceptual change and hypothetico-predictive processing success. It was argued that effective strategies for teaching the skill of hypothesizing and predicting should train students to apply hypothetico-predictive processing behaviors that establish links within and between students' procedural and declarative knowledge bases.

This chapter has provided the science education researcher with a framework for continuing useful investigations concerning the nature of hypothetico-predictive reasoning. The chapter has also endeavored to provide the science teacher with suggestions for teaching and assessing this important cognitive tool. The fact that many of these suggestions have yet to be formally tested should not eliminate their application. As teacher/researchers (Kincheloe, 1991; Tobin, 1991), science teachers have the power and the opportunity to develop, investigate, and modify the use of hypothetico-predictive teaching/learning strategies in their own classrooms. This author welcomes collaboration with science teachers interested in investigating some of the methods described in this chapter.

References Cited

Abraham, M. R., & Renner, J. W. (1986). The sequence of learning cycle activities in high school chemistry. *Journal of Research in Science Teaching, 23*, 121-144.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy: Project 2061*. Washington, DC: American Association for the Advancement of Science.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1985). *Cognitive psychology and its implications (2nd ed.)*. New York, NY: W. H. Freeman.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94*, 192-210.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review, 96*, 703-719.

Anderson, J. R., Boyle, F., Corbett, A. T., & Lewis, M. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence, 42*, 7-49.

Anderson, O. R. (1992). Some interrelationships between constructivist models of learning and current neurobiological theory, with implications for science education. *Journal of Research in Science Teaching, 29*, 1037-1058.

Anderson, O. R., & Demetrius, O. J. (1993). A flow-map method of representing cognitive structure based on respondents' narrative using science content. *Journal of Research in Science Teaching, 30*, 953-969.

Armbruster, B. (1985). Content area textbooks: a research perspective. In J. Osborn, P. Wilson, & R. Anderson (Eds.), *Reading education: Foundations for a literate America*. Lexington, MA: Lexington Books.

Ausubel, D. (1986). *Educational psychology*. New York, NY: Holt, Rinehart, and Winston.

Barba, R. H. (1990). Problem solving pointers: Techniques that can be taught. *Science Teacher, 57*, 32-35.

Barba, R. H. (1992, March). *A qualitative and quantitative analysis of expert and novice earth and space science teacher's declarative, procedural, and structural knowledge*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, Massachusetts.

Barba, R. H., & Rubba, P. A. (1992). Procedural task analysis: A tool for science education problem-solving research. *School Science and Mathematics, 92*, 188-192.

Bransford, J. D. (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher, 19*, 2-10.

Bransford, J. D., Sherwood, R. D., Vye, N. J., & Rieser, J. (1986). Teaching thinking and problem solving: Suggestions from research. *American Psychologist, 41*, 1078-1089.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology*. Hillsdale, NJ: Erlbaum.

Bunce, D. M., Gabel, D. L., & Samuel, J. V. (1991). Enhancing chemistry problem-solving using problem categorization. *Journal of Research in Science Teaching, 28*, 505-521.

Butts, D., Capie, W., Fuller, E., May, D., Okey, J., & Yeany, R. (1978). Priorities for research in science education. *Journal of Research in Science Teaching, 15*, 109-114.

Carey, S. (1986). Cognitive science and science education. *American Psychologist, 41*, 1123-1130.

Chinn, A. C., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*(1), 1-49.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser (pp. 453-494)*. Hillsdale, NJ: Lawrence Erlbaum.

Demetrius, O. J., & Anderson, O. R. (1992, March). *A recursive dendrographic method of representing cognitive structure based on respondent's narrative using science content*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, Massachusetts.

Derry, S. J. (1989). Strategy and expertise in solving word problems. In C. B. McCormick, G. Miller, & M. Pressley (Eds.), *Cognitive strategy research: From basic research to educational applications.* New York, NY: Springer-Verlag.

Deshler, D. d., & Schumaker, J. B. (1986). Learning strategies: An instructional alternative for low-achieving adolescents. *Exceptional Children, 52*, 583-590.

Dunkhase, J. A., & Penick, J. E. (1990). Problem solving in the real world. *Journal of College Science Teaching, 19*, 367-370.

Farnham-Diggory, S. (1992). *Cognitive processes in education.* New York, NY: HarperCollins.

Finley, F. N. (1985). Variations in prior knowledge. *Science Education, 69*, 697-705.

Fisher, K. M. (1990). Semantic networking: The new kid on the block. *Journal of Research in Science Teaching, 27*, 1001-1018.

Gabel, D., Ed. (1990). *What Research Says to the Science Teacher, Vol 5 (Problem Solving).* Washington, D.C.: National Science Teachers' Association.

Good, R. (1977). *How children learn science.* New York: MacMillan Publishing Co.

Good, R. (1989, March). *Toward a unified conception of thinking: A cognitive science perspective.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Good, R., Strawitz, B., Franklin, B., Smith, V., Roberts, F., & Moncado, G. (1988, April). *Using prediction in a science learning cycle: A pilot study and proposed research.* Paper presented at the 61st annual meeting of the National Association for Research in Science Teaching, Lake Ozark, MO.

Greeno, J. G., & Simon, H. A. (1986). *Problem solving and reasoning: Handbook of cognitive science.* Princeton, NJ: Lawrence Erlbaum.

Holland, J., Nisbett, R., & Thagard, P. (1989). *Induction: Processes of inference, learning, and discovery.* Cambridge, MA: MIT Press.

Hollingsworth, S. (1989). Prior beliefs and cognitive change in learning to teach. *American Educational Research Journal, 26*, 160-189.

Johnson-Laird, P. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Jones, G. (1990, April). *Cognitive conflict and cooperative learning*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.

Kagan, S. (1990). The structural approach to cooperative learning. *Educational Leadership*, January, 12-15.

Kincheloe, J. L. (1991). *Teachers as researchers: Qualitative inquiry as a path to empowerment*. New York, NY: Palmer Press.

Kyle, W. C., & Shymansky, J. A. (1992). Conceptual change teaching and science learning. *NARST Monograph, 5*, 29-40.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago, ILL: The University of Chicago Press.

Lavoie, D. R. (1991). Enhancing science instruction with hypothetico-predictive reasoning: A cognitive science perspective. *Proceedings of the Montana Academy of Sciences, 51*, 1-15.

Lavoie, D. R. (1992a, March). *The effects of adding a prediction/discussion phase to a science learning cycle*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, MA.

Lavoie, D. R. (1992b, March). *Improving a pair problem-solving strategy for developing students' hypothetico-predictive reasoning skills*. Paper presented at the annual meeting of the Association for the Education of Teachers in Science, Boston, MA.

Lavoie, D. R. (1993a). The development, theory, and application of a cognitive-network model of prediction problem solving in biology. *Journal of Research in Science Teaching, 30*, 767-785.

Lavoie, D. R. (1993b, April). *The development and evaluation of an explicit cognitive strategy for teaching prediction problem solving in biology*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.

Lawson, A. E. (1986). A neurological model of sensory-motor problem solving with possible implications for higher-order cognition and instruction. *Journal of Research in Science Teaching, 23*, 503-522.

Lawson, A. E. (1990). Use of reasoning to a contradiction in grades three to college. *Journal of Research in Science Teaching, 27*, 541-551.

Lawson, A. E., Abraham, M. R., & Renner, J. W. (1989). A theory of instruction: using the learning cycle to teach science concepts and thinking skills. *NARST Monograph*, 1, National Association for Research in Science Teaching.

Lawson, A. E., Baker, W. P., DiDonato, L., Verdi, M. P., & Johnson, M. A. (1993). The role of hypothetico-deductive reasoning and physical analogues of molecular interactions in conceptual change. *Journal of Research in Science Teaching, 30,* 1073-1085.

Lawson, A. E., McElrath, C. B., Burton, M. S., James, B. D., Doyle, R. P., Woodward, S. L., Kellerman, L., & Snyder, J. D. (1991). Hypothetico-deductive reasoning skill and concept acquisition: Testing a constructivistic hypothesis. *Journal of Research in Science Teaching, 28,* 953-970.

Lawson, A. E., McElrath, C. B., Burton, M. S., James, B. D., Doyle, R. P., Woodward, S. L., Kellerman, L., & Synder, J. D. (1991). Hypothetico-deductive reasoning skill and concept acquisition: Testing a constructivist hypothesis. *Journal of Research in Science Teaching, 28*, 953-970.

Lesgold, A. (1988). Problem solving. In R. J. Sternberg, & E. E. Smith, Eds., *The Psychology of Human Thought.* New York, NY: Cambridge University Press.

Livingston, C., & Borko, H. (1989). Expert-novice differences in teaching: A cognitive analysis and implications for teacher education. *Journal of Teacher Education, 40*, 36-42.

Malone, J., & Dekkers, J. (1984). The concept map as an aid to instruction in science and mathematics. *School Science and Mathematics, 84*, 220-233.

Martin, J. S., & Szabo, M. (1990). The dialogue and computer-assisted instruction in chemistry. *Journal of Computer-Based Instruction, 17*, 41-45.

McRobbie, C. J. (1991). Cognitive styles and cognitive structure. *Science Education, 75*, 231-242.

Minsky, M. (1986). *The society of mind.* New York, NY: Simon and Shuster.

Minstrell, J. A. (1989). Teaching science for understanding. In L. B. Resnick, & L. E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research.* Washington, DC: Association for Supervision and Curriculum Development.

National Commission for Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, D.C.: U.S. Government Printing Office.

Novak, J. D. (1977). *A theory of education.* Ithaca, NY: Cornell University.

Novak, J. D., Gowan, D. B., & Johansen, G. T. (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education, 67*, 625-645.

Osborne, R., & Freyberg, P. (Eds.). (1985). *Learning in science: The implications of children's science.* Portsmouth, N. H.: Hinemann.

Osborne, R., & Wittrock, M. (1983). Learning science: A generative process. *Science Education, 67*, 489-508.

Pea, R. D. (1993). Learning scientific concepts through materials and social activities: Conversational analysis meets conceptual change. *Educational Psychologist, 28.*

Pestel, B. C. (1993). Teaching problem solving without modeling through "thinking aloud pair problem solving". *Science Education, 77*(1), 83-94.

Pines, A. L., & West, L. H. T. (1986). Conceptual understanding and science learning: an interpretation of research within a sources-of-knowledge framework. *Science Education, 70*, 583-604.

Popper, K. (1965). *Conjectures and refutations: The growth of scientific knowledge.* New York: Harper & Row.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*, 211-227.

Prawat, R. S. (1989). Promoting access to knowledge, strategy, and disposition in students: A research synthesis. *Review of Educational Research, 59*(1), 1-41.

Pressley, M., Goodchild, F., Fleet, J., Zajchowski, R., & Evans, E. D. (1989). The challenges of classroom strategy instruction. *Elementary School Journal, 89*, 301-342.

Quellmalz, E. S. (1987). Developing reasoning skills. In J. B. Baron, & R. J. Sternberg, (Eds.), *Teaching Thinking Skills: Theory and Practice.* New York, NY: W. H. Freeman and Company.

Randolph, G. L. (1988). Expert systems as cognitive models for intelligent tutors. *Journal of Computers in Mathematics and Science Teaching, 7*, 16-20.

Roth, W. M. (1991, April). *Aspects of cognitive apprenticeship in science teaching.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Lake Geneva, WI.

Rubin, R. L. and Norman, J. T. (1992). Systematic modeling verses the learning cycle: Comparative effects on integrated science process skill achievement. *Journal of Research in Science Teaching, 29*, 715-727.

Rumelhart, D. (1980). Schemata: The building blocks of cognition. In R. B. Spiro, & W. Brewer, Eds., *Theoretical issues in reading comprehension.* Hillsdale, New Jersey: Lawrence Erlbaum.

Schank, R., & Abelson, R. P. (1977). *Scripts, plans, goals, and understandings: An inquiry into human knowledge structures.* Hillsdale, New Jersey: Lawrence Erlbaum.

Shank, R. (1990). *Tell me a story: A new look at real and artificial memory.* New York: Scribners.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-22.

Simon, H. A. (1980). Information processing explanations of understandings. In P. W. Juseqyk, & R. M. Klein, (Eds.), *The Nature of Thought: Essays in Honor of D. O. Hebb.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, H. A. (1980). Problem solving and education. In D. Tuma, & F. Reif, (Eds.), *Problem solving and education.* Hillsdale, N.J.: Erlbaum.

Spiro, R. J., & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix, & R. Spiro (Eds.), *Cognition, education, and multimedia: Exploring ideas in high technology.* Hillsdale, NJ: Lawrence Erlbaum.

Sternberg, R. J. (1983). Components of human intelligence. *Cognition, 15*, 1-48.

Symons, S., Snyder, B. L., Cariglia-Bull, T., & Pressley, M. (1989). Why be optimistic about cognitive strategy instruction? In C. B. McCormick, G. Miller, & M. Pressley (Eds.), *Cognitive strategy research: From basic research to educational applications*. New York, NY: Springer-Verlag.

Tobin, K. (1991, April). *Constructivistic perspectives on research in science education*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Lake Geneva, WI.

Tuma, R., & Reif, F., Eds. (1980). *Problem solving and education: issues in teaching and research*. Hillsdale, NJ: Erlbaum.

Van Joolingen, W. R., & De Jong, T. (1992). *Supporting hypothesis generation by learners exploring an interactive computer simulation*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, California.

Watson, B., & Konicek, R. (1990). Teaching for conceptual change: Confronting children's experience. *Phi Delta Kappan*, May, 680-685.

Whimbey, A., & Lochhead, J. (1991). *Problem solving and comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wong, B. Y. C. (1989). Musing about cognitive strategy training. *Intelligence*, *13*, 1-4.

Yore, L. (1992, March). *Plausible reasoning in elementary and middle schoo science: Abduction, induction, deduction, and hypothetico-deduction* Paper presented at the annual meeting of the National Association fo Research in Science Teaching, Boston, Massachusetts.

Yore, L. (1994, May). *Elementary school science teacher education at th University of Victoria*. Paper presented at the B.C. Science Teachers Association Elementary Leadership Conference, University of B.C. Vancouver.

Knowledge Structures And Successful Problem Solving

Patricia L. Hauslein, Ph.D.
Department of Biology
St. Cloud State University
St. Cloud, MN

Mike U. Smith, Ph.D.
Mercer University School of Medicine
Macon, GA

Introduction and Definitions

In recent years both science teachers and science education researchers worldwide have become increasingly interested in problem solving. Educators are interested in enhancing student problem-solving skills as tools students can use for increasing and consolidating their understanding of scientific content. More generally, most if not all science educators hope that these skills can be learned in ways that will promote their transferability to tasks outside the classroom.

Early research (Simon & Simon, 1979; Larkin, McDermott, Simon, & Simon, 1980; Simon, 1980; Greeno, 1980; Larkin, Heller, & Greeno, 1980; Larkin & Reif, 1979) focused on the nature of successful problem solving (usually as performed by educator experts in the field) and how this activity differed from unsuccessful problem solving (usually studied in novices, i.e., neophytes in the field). Perhaps the most pervasive (if not unexpected) findings in these studies was that experts have much more knowledge of the domain, which they bring to bear on the solution process (Simon, 1980; Chi, Glaser, & Rees, 1982). Subsequent studies revealed, however, that experts (and also successful problem solvers with less experience) differ from unsuccessful solvers, not only in the quantity of their knowledge (i.e., how many things they know), but also in the structure of their knowledge (e.g., in cohesiveness, congruence, interrelatedness, accessibility, and organization). It is this structure that we usually are referring to when we speak about "understanding" a certain domain. The purpose of this chapter is to review the techniques employed in analyzing this cognitive structure, to review the more salient findings of this research, to summarize our current understanding of the nature of cognitive structure, and to consider some of the implications of this understanding for instruction and further research. This chapter assumes the theoretical framework of constructivism within the information-processing paradigm. We also believe that a complete understanding of problem solving must include at least a basic model of brain function. Later in the chapter we

apter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for ntific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

will discuss a model of brain function referred to as the connectionist model. We feel this model is a good physiological fit to the pedagogy and psychology of problem solving, i.e. the cognitive science of problem solving.

Within the information-processing paradigm, knowledge structure (also referred to as cognitive structure or conceptual framework) refers to the internal organization of knowledge. Cognitive structure may be defined as "a hypothetical construct referring to the organization (relationships) of concepts in long-term memory" (Shavelson, 1972, pp. 226-227). Cognitive structure is assumed to include both the identifiable elements of knowledge (the facts, concepts, principles, and procedures) and the relationships among those elements (Shavelson, 1972). This term is sometimes used synonymously with knowledge structure, although the latter term is more commonly used to refer to the structure of an individual's internal knowledge or to the formal conceptual structure inherent in a particular domain as evidenced by writings in the field, that is a structure of the discipline (Schwab, 1978). Cognitive structures can also be equated with mental models, in the broadest sense of the term as mental representations of external phenomena. In more typical usage, however, mental models are commonly understood to be not the total sum of knowledge about a field but some representation created with the express purpose of abstracting the most critical components of a limited phenomenon to facilitate understanding and/or explanation (Johnson-Laird, 1983). In this sense, one may hold a mental model of the atom as analogous to our solar system, a model of electricity as analogous to the flow of water through a pipe, or a model a cell membrane as a bowl of Jell-O with fruit imbedded within. (Much of the mental models research has also focused on the way in which language can be understood as strings of symbols that are mental representations of meaning; Johnson-Laird, 1983). We prefer, therefore, to view mental models as part of the broader cognitive structure that encompasses an individual's entire understanding of a domain.

This view of cognitive structure is congruent with constructivist epistemology. It implies that coming to understand, i.e., learning or sense-making, is a process in which the brain intentionally attends to and selects information from which to construct meaning, interpreting incoming information in light of the current cognitive structure. This construction results in the creation of cognitive links between what is perceived and relevant aspects of long-term memory. Learning from an experience may involve adding new pieces of information (nodes) and/or new linkages within the extant cognitive architecture or may require destroying or remodeling old structures which are inconsistent with the new information. Piaget called these two types of learning assimilation and accommodation, respectively. Recent research into the latter has focused on the alternative conceptions that must be replaced, the process has come to be called conceptual change or exchange (Posner, Strike, Hewson, & Gertzog, 1982).

Cognitive maps, another term frequently encountered in this literature, are simply two- (or even three-) dimensional graphical representations of this mental organization. In these maps, the smallest bits of knowledge (which can be either declarative or procedural) are typically represented as nodes which are, in turn, linked to each other, representing the relationship between the two nodes. Perhaps the most widely known type of cognitive map in science education is the "concept map," popularized by Joe Novak. (For a summary of recent understanding and use of concept maps, as well as concept circles and semantic nets as employed by James Wandersee and Kathleen Fisher, respectively, see *Journal of Research in Science Teaching*, special issue 10, volume 27, 1990).

These collections of nodes and links are related to, but not to be confused with, the sets of interconnected nodes problem-solving researchers (especially cognitive scientists) often use to represent a solution pathway taken by an individual. The nodes in the latter representations are taken as time points or positions held during the problem-solving process. The solver selects paths within these nets as s/he perceives particular concepts contributing to the solution of the problem (e.g. Simmons & Lunetta, 1993) Such representations are superficially very similar to cognitive maps, but the solution pathways represent the possible sequences of various steps that may be taken in the solution process, not the interrelationships among elements of a person's internally structured knowledge.

In the following pages we propose to describe cognitive structure as it relates to scientific problem solving. As science teachers, our interest in this regard is in the problem solving that occurs in the science classroom and not in the solution of problems in the scientific research laboratory. The latter open-ended, less structured problem solving is, of course, more difficult to understand and has been less studied. Although the field of situated cognition is currently a prominent area of research (see Rogoff & Lave, 1984 for an introduction) we will focus on the formal knowledge of a domain and the problem solving in the classroom with reference to "real world" problems where appropriate.

The terms "problem" and "problem solving" have also been used with disparate meanings in the literature. We take a problem to be "any task that requires analysis and reasoning toward a goal (or "solution") (Smith, 1991). A task would be defined as a problem because of its innate characteristic to cause analysis and reasoning for a solution. For an extensive treatment of the two sides of this issue, see Smith, 1991 and Bodner, 1991.

Knowledge Structures and the Brain

We are a long way from completely understanding the working of the human brain. The complexity of the system especially as we explore higher cognitive functions is awesome. However, in an extremely useful synthesis of neurobiology, cognitive science, and the constructivist philosophy Anderson (1992) suggests that our current understanding is adequate to "establish the foundations for a complete neurocognitive theory of human behavior, especially in relation to perception and knowledge construction" (p. 1039).

At least two characteristics of the brain are important to our understanding of problem solving. The first is the active processing of information the brain accomplishes as impulses are generated in response to the environment. Sensory input (i.e. nerve impulses) pass through a number of brain structures before reaching the cerebral cortex, an area associated with higher cognitive function. For example, before reaching the cerebral cortex the input may pass through the limbic system to be evaluated for an emotional response and/or to be compared to prior knowledge. These structures filter and mold incoming information, determining how it is encoded or even if it is encoded into memory.

The second characteristic of the brain useful to understanding problem solving is its ability to autogenerate impulses within the system. Many of the cells excited following an external stimulus are actually excited by nerve fibers coming from the cortex or other intermediate brain centers and not directly from the receptor cells. The impulse from the receptor cell begins a cascade of neural activity, a literal spread of activation throughout the brain (Anderson, 1992). However, this spread of activation must be limited, otherwise every stimulus response would put us into a frenzy of brain activity.

The effect of one neuron on another is determined by its rate of firing, that is, the number of impulses it transmits per second. A neuron is said to have an activation level analogous to its firing rate. The interaction between neurons are of two types. A neuron can increase the level of activation of another cell or, as importantly, it can reduce the activation level thus inhibiting it from firing. Since any one neuron can be synapsed with many others, only those neurons receiving more excitation than inhibitory messages are activated. In this connectionist model of knowledge construction (McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986), it is the pattern of connections, the neural net, which forms an internal representation of knowledge and experience. For example, the concept of DOG is represented by a particular pattern of neural activation, a pattern which also includes the inhibition of other perhaps related networks. It is a common experience of many parents to have their toddler identify all small creatures as dogs. Within this connectionist model of information processing, as the child matures a second related neural net is established which corresponds to CAT. Thus after the stimulus of a dog barking is received, the neural net associated with DOG is active whereas the CAT net is inhibited.

Connectionist modeling explains the ability of the human brain to actively construct meaning from external stimuli and store knowledge in useful patterns, as well as reconstruct prior knowledge as it is challenged by new knowledge. This model may also explain the persistence and interference of misconceptions. The connectionist model of cognition describes learning not as a simple stimulus-response link, but as the establishment and modification of a network of interconnected neurons with multiple input and output pathways. With each cycle of learning the connections between some neurons are strengthened while others are weakened. A rehearsed misconception can become well established, and therefore difficult to restructure.

Understanding the organization of incoming information and its organization into concepts has been the specific focus of a substantial number of cognitive scientists. Much of this research has focused on an individual's ability to form concepts or categories (such as DOG) and to recognize specific instances as belonging to those categories (cf. Estes, 1991, 1993; Rosch & Lloyd, 1978; Smith & Medin, 1981). In an effort to determine how categorized knowledge is organized within the brain, Knowlton and Squire (1993) studied categorization in a group of amnesic patients who had severely impaired declarative memory due to limbic or diencephalic brain damage. Although the brain-damaged subjects could not recall seeing a pattern only moments after seeing it, they could properly categorize a new pattern as belonging to a pre-trained category or not, as well as the control subjects. Thus, damaging the ability to encode declarative knowledge (the experience of a stimulus) does not interfere with the development of categories or with the subsequent recognition of members of those categories. This study suggests that although encoding individual stimuli or experiences is dependent on the limbic and diencephalic regions of the brain, categorization-level knowledge is not dependent on the integrity of these structures. In the context of the current discussion, the implication is that concept (or categorization) knowledge is distinct from knowledge of its exemplars or component parts. Learning of concepts may, therefore, not occur as a byproduct of perception of individual inputs. The organization of knowledge into concepts or constructed categories may require a different (and explicit) form of instruction.

Problem-solving difficulty can occur of course for many reasons, however, the ability to delineate the characteristics of problems seems to lie at the foundation of successful problem solving. Perfetto, Bransford, and Franks (1983) showed that subjects did not spontaneously use prior knowledge (i.e. activate a specific neural net) to solve a problem unless they were explicitly informed of the relationship between prior declarative knowledge and the problem-solving situation. In addition, when these subjects were given a second chance to solve a problem, inadequate self-generated solutions interfered with subsequent access of relevant knowledge (i.e. lack of inhibition of inappropriate neural nets). Using the connectionist model, it would seem from the research that correct categorical-level activation is essential for problem-solving success. For example, top-down information processing (Anderson, 1985) suggests that high-level general knowledge (e.g. categories) will determine the interpretation of incoming information. If inappropriate or less useful categories are activated the student may spend time traversing neural nets that will not aid in the solution of a problem. In fact, this model might suggest that inappropriate neural nets may become strengthened if the student continues to make the same mistake, thus promoting or reinforcing the misconception.

The active processing and autogeneration of impulses of the human brain make our understanding of problem solving that much more complicated. Problem situations have their own characteristics but it will also be our understanding of the internal problem situation that will lead to better understanding of the nature of problem solving.

Knowledge Structure Assessment

Internal mental organization of knowledge is, of course, a hypothetical construct. We can use microscopes to study the interconnections of neurons in the brain, but we have no cognitive microscopes with which to literally see the ways in which a person's knowledge is structured. Like a physicist who deduces the behavior of sub-atomic particles from traces left in a cloud chamber, we are left to deduce the internal knowledge structure from externally visible evidence. A wide variety of assessment tools is currently in use, some of which are more appropriate than others to some domains or certain types of research questions. Understanding the differences and the relative strengths and weaknesses of each is of prime importance before we can judge the merits of the available studies in this domain and before we can intelligently select tools for our own use.

Most assessment tasks request that the subject explicitly organize some information provided. The sciences are recognized as semantically-rich domains, that the essence of the sciences is embedded in the language used and in the unique meaning ascribed to each term. Many current assessment techniques have therefore focused on the concepts and conceptual labels used in a particular domain. In these tasks the subject is provided with more or less limited groups of terms and asked to sort or otherwise organize them in some way that is assumed to reflect his/her mental organization. The terms may represent declarative or procedural knowledge nodes. The researcher may or may not provide the categories into which the terms are to be organized. Concept sorting has been used effectively in studies of knowledge structure in biology (cf. Hauslein, Good, & Cummins, 1992), geology (cf. Champagne, Klopfer, Desena, & Squires, 1981), chemistry, (cf. Gorodetsky & Hoz, 1985), physics (cf. West, Fensham & Garrard, 1985, and math (cf. Roehler, Duffy, Hermann, Conley, & Johnson, 1988).

Other prominent research has provided the subjects with collection of problems to organize. Problem sorting has been studied in a diverse set of domains, including genetics (cf. Smith & Good, 1984), physics (cf. Chi, Feltovich, & Glaser, 1981) computer programming, (cf. Weiser & Shertz, 1983), trouble shooting (cf. Perez, 1991). As in the concept-sorting tasks, subjects may be instructed to group the items into a pre-established set of categories (cf. Veldhius, 1990) or into categories that they devise according to some criterion provided. In the latter case, problems may be grouped into sets "that would be solved in similar ways" or terms grouped "according to similarity." In addition to the terms provided, some studies have asked subjects to generate other related terms. The prime example of this procedure is the word association task (e.g., Thro, 1978; Jackson, Grassia, & Wolfe, 1988); others request that the subject produce superordinate and/or subordinate concepts or terms to the ones provided (Hauslein et al. 1992; Adelson, 1985). Some researchers have asked that subjects directly produce node-link maps or other graphical representations of their knowledge (Champagne et al. (1981). Still other studies provide a text for the subject to summarize (Chi et al., 1982).

The second component of most assessment tasks is the requirement that the subject provide information about the nature of the relationships among the various elements (usually after they have been sorted). Subjects may be asked to explain what each of their groups represents or the criteria upon which membership is determined (cf. Hauslein et al, 1992, Smith, 1990; Smith, 1992). Subjects may be asked to rate the similarity or the relatedness of pair-wise combinations of terms, as on a scale of 1 to 10 (Fennker, 1975; Diekhoff, 1983; Rips, Shoben, & Smith, 1973) or relatedness may be estimated by the order in which the terms are generated (as in word association tasks, e.g., Adelson, 1981; McKeithen, Reitman, Rueter, & Hirtle, 1981; Jackson et al., 1988; Thro, 1978). In yet other studies (e.g., Santos-Gomez, Wooliscroft, & Mennin, 1990), subjects are asked to rate the degree to which one phenomenon is typical or characteristic of another (as in typical presenting symptoms for a given disease). Roth, Gabel, Brown, and Rice (1992) used a proximity index from multidimensional scaling to mathematically determine relationships between concepts. Subjects can also be asked to solve a collection of problems and their performance analyzed to deduce the rules used in their solution (Maloney, 1987; Siegler, 1978). The unique collection of these rules and the characteristics that appear to determine their use are therefore considered to be one view of cognitive structure.

Researchers have also found it useful to ask subjects to simply describe in a less structured way "everything they know" about a particular topic, most often using "think-aloud" interviews and/or protocol analysis (Chi et al., 1982). Anderson and Demetrius (1993) have used a modification of this procedure which they call "flow mapping" which asks the subject specifically to describe the major elements involved in the problem situation and their functions.

The products of these assessment tasks may be thought of as a more or less accurate snapshots of a portion of an individual's cognitive structure (understanding). Individual cognitive maps have two general purposes. First, there is a growing body of evidence that they can be powerful metacognitive tools for meaningful learning. Students asked to produce such maps must necessarily focus, not only on the individual pieces of information which they might be prone to memorize in a non-meaningful way, but also on the structure of the knowledge--the way in which the pieces are related. Second, cognitive maps can be used as formative or summative evaluation tools. Most of the above techniques, because of their methodological and computational complexity are used primarily in research settings. However, the simpler techniques of concept mapping can be used in classroom situations providing a teacher with a powerful tool to assess student learning (Novak, 1990).

While classroom research has focused on the use of cognitive maps of individuals, most psychological research to date has focused on characterizing the knowledge structures common to defined groups of individuals such as novice students or Ph.D. educator experts. These sorting task analyses result in an outcome which describes characteristics of the different groups, that is, the analyses generate a group cognitive structure. The group cognitive structures are then compared. For educators, the underlying assumption is that a detailed understanding of the knowledge structure of experts and of the differences in the knowledge structures of good and poor problem solvers in a knowledge domain will be powerful information for designing instruction. Recent research has provided many insights into these typical knowledge structures, but their utility when applied to the classroom largely remains to be tested.

Combining the extremely rich data from a group of cognitive maps can be a daunting challenge. Most studies to date have employed either cluster analysis or factor analysis techniques. In brief, cluster analysis techniques measure the tendency for a group of subjects to associate one item with another. When many such pair-wise comparisons are calculated, groups of items are typically observed to "cluster" together. Each cluster can be considered a node in the collective cognitive structure and the degree to which two items tend to be connected by the group is a measure of the relatedness of these items in the group's cognitive structure. When subjects are asked to group items into categories established *a priori* by the experimenter (or when acceptable similarity dimensions can be identified *a posteriori*), each category can be considered to be a variable, and the frequency with which an item is placed by the subjects into that category can be considered the value of the item on that variable. Thus, the use of two prescribed categories produces data that can be used to place each item in a two dimensional Euclidean space. The use of more than three dimensions (categories) is difficult to visualize but simple to analyze by multidimensional scaling (MDS) which permits the computation of the scalar distance (relatedness) of any two items within the n-dimensional space. These results may also be presented as dendrograms (such as Figure 1 below) in which the sorted items are represented as equidistant points along the X axis and the measure of similarity is plotted on the Y axis. The height of the vertical line joining any two items is the measure of the relatedness. When preestablished categories are not employed, the cluster analysis can be performed by preparing similarity matrices which present the frequency at which each item was paired with every other item. (For more detailed explanations of these and related techniques, see Kruskal, 1977; Schvaneveldt et al.,1985; Schvaneveldt, Durso, & Dearholt, 1989; and Veldhuis, 1990.)

The results of the analyses of multiple interrelationships of the many elements of the cognitive map can, of course, be difficult to interpret, especially as the number of dimensions studied increase. In addition to cluster analysis, therefore, the categorizations of the research subjects (expressed as item pairings) are often subjected to factor analysis, a set of statistical procedures employed for the purpose of explaining these multiple relationships "in terms of a few conceptually meaningful, relatively independent factors" (Kleinbaum & Kupper, 1978, p. 376). Item pairings by the subjects can also be subjected to factor analysis which groups the test items into factors analogous to clusters. Correlations ("factor loadings") between each of the original variables and the underlying common factors identified by factor analysis are then computed. The variables that correlate highly with a given emergent factor are compared against those variables that do not. The researcher uses this information to conceptually interpret each factor.

These findings are often visualized by employing multidimensional scaling (MDS) techniques. MDS considers each factor as a continuous variable and the factor loading as the value for each original variable on that dimension. Plotting the factor loadings on two factors (as the X and Y axes) then results in a two-dimensional matrix on which the proximity of any two items indicates how closely they are related in the original data collected from a group of subjects. Most MDS programs have a pre-established goodness-of-fit index to mathematically determine the number of dimensions. Of course, programs can be forced to generate more or fewer dimensions, at the discretion of the researcher. As with any data reduction technique information is lost with each reduction therefore the number of dimensions and their

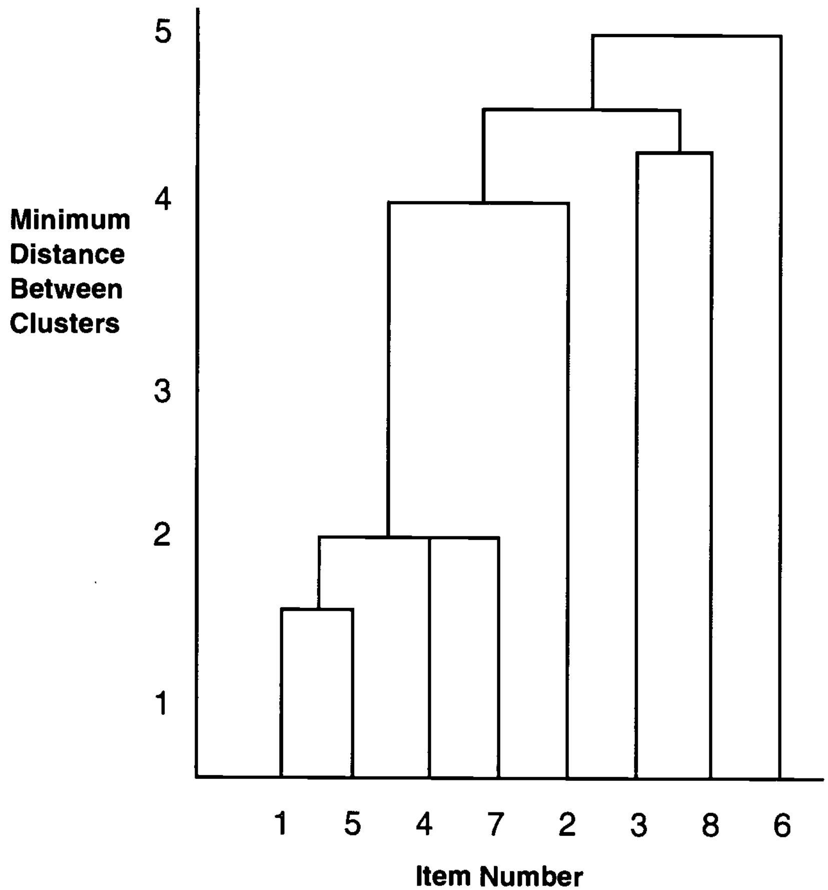


Figure 1. Dendogram derived from association matrices.

interpretation can be very subjective and intuitive. For example, two items in close proximity in two dimensions may become separated on a third dimension. It is up to the researcher to determine if a third or even fourth dimension is interpretable and/or useful in a particular project. A straightforward example of how these procedures can be applied can be found in Hauslein et al. (1992).

Statistical procedures, such as ALSCAL on SAS (see Takane, Young, Lewyckj, & de Leeuw, 1978) also produce weight scores that allow for the comparison of the extent to which different groups of research subjects employ the dimensions identified. These weights indicate how influential (salient) that dimension is to the decision of the subject group to place an item in the internal cognitive architecture on the basis of that factor. For example in their study, Hauslein et al. (1992) found that preservice teachers tended to categorize concepts based only on the community-to-individual dimension, whereas novice teachers tended to consider both dimensions, community/individual and cellular/sub-cellular.

Subjects of different types (e.g., experts versus novices) can be further compared and the factors underlying the differences between the categorizations made by these groups identified by computing cognitive distances between groups and using these as the input data for another multidimensional scaling analysis. Cognitive distance is based on the number of individual item-pair differences between groups (Fillenbaum & Rapoport, 1971; Gorodetsky & Hoz, 1985). When the Hauslein et al.(1992) cognitive structure data were subjected to this analysis, the MDS plot presented in Figure 2 resulted. Two factors emerged as explaining differences among the subject groups. The authors interpreted these factors as representing deep versus surface concept attributes and fixed versus fluid cognitive structure. This plot demonstrates that the experienced teachers and the scientists were similar in their deep understanding of the concepts. However with consideration of a second dimension, these two groups were found to be quite different in the fluidity of their cognitive structures.

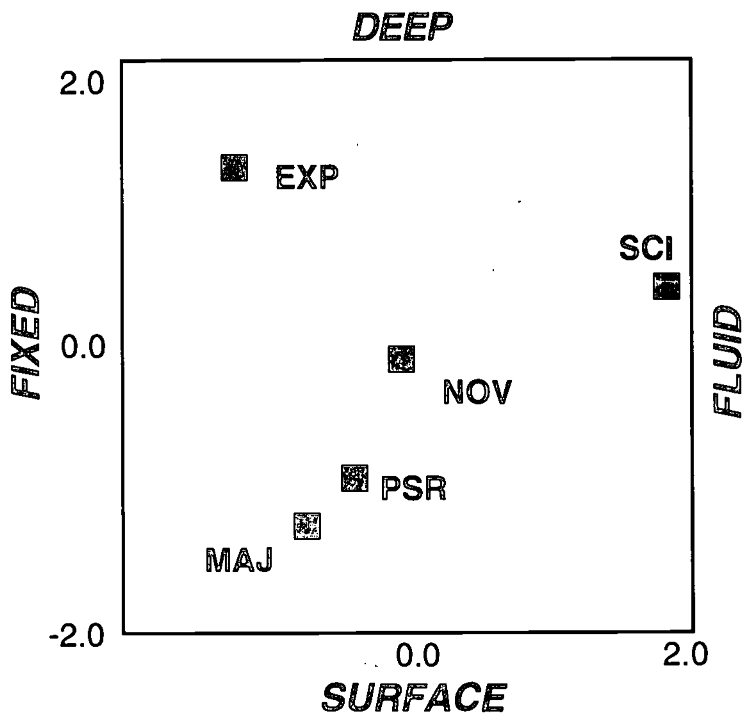


Figure 2. Multidimensional representation of cognitive distance among groups.

Methodological Concerns in Knowledge Structure Assessment

As with any research tool, the issues of validity and reliability need to be addressed. In particular to cognitive structure assessment we need to ask (1) How accurately does the picture obtained reflect the internal brain reality (their validity) and (2) To what extant are the results of our assessments repeatable (their reliability)?

The concern for the general validity of knowledge structure assessment can be indirectly addressed in three ways. The first is by using triangulation, i.e., the collection of information about knowledge structures using two or more different tools or perspectives. The primary way in which triangulation is applied in this research is to perform both qualitative and quantitative analyses of the research subjects' behavior and to compare the findings. The often congruent findings from both quantitative and qualitative measures typically found in most knowledge structure studies, to some extent explains why researchers have not been overly concerned with this issue. A stronger case could be made for validity however, if researchers would employ at least two completely different tools (e.g., a sorting task and a word association task) for assessing the same knowledge structure. We are aware of no such studies, but we strongly recommend this experimental design for future research.

A second procedure for enhancing the validity of a study's findings employed by at least two investigations (Chi et al., 1981, & Smith, 1992) can be considered as an assessment of predictive validity. In both of these studies, an initial round of assessments of the knowledge structures of selected groups of individuals (teachers, professionals, and students) was used to design a second sorting task that focused more sharply on the differences that were found to distinguish among the groups in the first assessment. The sortings obtained with the revised task were then compared against those predicted on the basis of the first experiment. This is a very robust experimental approach, which we recommend to other researchers.

The third way to address the concern of validity in these types of studies is perhaps more obvious. As with other types of research, congruence of the findings of many studies by different researchers strongly supports the common conclusions, especially if they arise from divergent types of research (e.g., qualitative and quantitative).

Reliability should also be a concern in studies of knowledge structures. It seems logical that the organization produced by a given subject at a given time might be affected by a host of external and internal subject variables as well as by minor variations in the task itself. Many studies (cf. Champagne, et al 1981; Champagne, Gunstone, & Klopfer, 1985, Gordetsky & Hoz, 1985; Uche, 1987; West, Fensham, & Garrard, 1985) have shown a shift in cognitive structure due to instruction but few studies have looked at the repeatability of the original cognitive structure when treatment is withheld. Using a word association task, Roth et al. (1992) found little change in cognitive structure of the control group after a 4 week interval. However, we also have anecdotal data suggesting that the first assessment of an individual's knowledge structure may differ substantially from a repeated assessment. The identification of internal and external variables and the nature of their effects on cognitive structure assessment should continue to be a source of concern and a topic that merits further research

Each of the techniques used to assess knowledge structures has its own set of strengths and weaknesses, and care is required when a teacher or researcher is choosing the appropriate tools to be used in a given situation. We suggest that the selection be guided by the following questions.

(1) Is the primary purpose of the assessment to determine the knowledge structures of an individual or to generalize about the structures of a group of similar individuals (e.g., novices teachers)? Assessment techniques that employ quantitative analyses which allow for the collation of data from many subjects may be required for studies of group cognitive structure. However, quantitative measures can gloss over important individual differences, as when a pair of items is grouped together by one subject for a reason that is very different from the reason a second subject uses to group the same two items. Qualitative research generally gives a richer, more detailed picture of the individual.

(2) What kind of resources are available? The various techniques require differential amounts of time, equipment (especially computers for sophisticated statistical analyses such as MDS), and expertise in designing, performing, and analyzing the assessment.

(3) Am I concerned about the nodes in the knowledge structure (the terms and concepts) and how closely they are linked and/or am I concerned about the nature of the connections between the nodes? Some techniques (e.g., word association tasks) assess the degree of relatedness between concepts but provide no direct information about the nature of the relationship. In contrast, some techniques (e.g., factor analysis) facilitate deducing the relationship, while others ask the subject to explicitly name and label each link.

(4) Am I seeking to discover new modes of organization or to confirm previously identified structures? In other words, how much do I already know about this knowledge structure in the target individuals? The answer to this question will determine

how much structure you want to impose on the outcome by the design of the task itself. Some tools, for example, simply ask for sorting items into piles and thus proscribe the production of hierarchical structuring. Similarly, tasks that require subjects to sort items into a prescribed set of categories assume that those categories are the ones employed by the subjects. For example, a study that seeks to confirm that Ph.D. faculty members use conceptual (or "deep") categories compared to the more "superficial" categories of novice students in the field, would employ a less open-ended assessment than a study of knowledge structures in an area where such a typical distinction was being sought.

Knowledge Structures and Problem Solving

Nearly to the point of definition, knowledge structures are the foundation of either success or confusion in problem-solving situations. Problem-solving success depends on whether or not the solver possesses the requisite declarative and procedural knowledge, whether s/he can identify which components of this knowledge are the most applicable in the current problem-solving task, and the motivation/personal inclination to expend the effort required to attempt to solve the problem. Most work in cognitive structure and problem-solving research has revolved around the differences found between experts and novices of a knowledge domain. Across various content domains the characteristic differences found between these two groups can be summarized into three categories:

(1) perception of the problem,
(2) approach to and solving of the problem, and
(3) cognitive means brought to the problem.

The difference in the perception of a problem between experts and novices is best described in the classic study of Chi et al., (1981). They observed that experts sorted problems according to conceptual ("deep") characteristics of the problem statement whereas novices focused on a literal reading (or "surface" characteristics) of the problem (e.g., inclined planes versus pulley problems).

The difference in the approach to a problem between experts and novices has been described as working forward versus working backward or means-ends analysis (Larkin et al. 1980; Larkin, 1981). Using the means-end approach, novices compare the initial state of the problem to the desired result and take actions (at times hit-or-miss) to decrease the distance between the two states (see Anderson, 1985). Novice problem solving is therefore often a superficial formula-matching task. On the other hand, experts use a forward working or knowledge development approach, using the givens of the problem to generate the information needed in successive problem steps until the desired result is achieved. Experts tend to spend a considerable amount of initial time on a low-detail qualitative analysis of a problem which is then used in planning the approach to the problem.

Both of these differences (perception of and approach to a problem) are dependent upon the third characteristic, that of the cognitive means one brings to the problem. As discussed by Chi, Glaser, and Rees (1982), the problem-solving enterprise is mediated by the mental processes defined by the knowledge structure. Experts are found to have knowledge of their field more hierarchically arranged and stored in larger functional units (Larkin, McDermott, & Simon, 1980). It is inferred, therefore, that the expert's greater success at problem solving is due to the ability to activate the appropriate problem schemata (Chi et al., 1981) and therefore to more efficiently access and process useful knowledge (Roehler et al., 1988). In addition to having a greater wealth of declarative knowledge, the expert also has a greater store of domain-specific procedural knowledge. This is in contrast to the novices' use of less powerful but broadly applicable heuristics (Larkin, 1981). The results of the studies mentioned above and others have resulted in the identification of characteristics associated with expert versus novice problem-solving behaviors. Useful summaries of these are presented in Simmons and Lunetta (1993) and Smith (1991).

The Knowledge Organization By Use Hypothesis

The continued focus on expertise as a unitary construct for explaining knowledge structure and its relation to problem solving is becoming increasing difficult to support. First, there is now increasing evidence that expertise is a continuum of knowledge and skills rather than the simple expert/novice dichotomy. Experts do indeed tend to have more knowledge of the domain, to be more successful problem solvers in the area, etc., but few researchers would maintain that there is an absolute line of demarcation between expert and novice. In some studies, some experts have been found to be less successful problem solvers than some novices, depending presumably on many situation-specific and subject-specific variables. Second, current research suggests that experts within a domain may vary in important ways, especially in the manner in which they have organized their knowledge. These different types of experts appear to have structured their knowledge in ways that reflect the manner in which they are called upon to access and apply that knowledge. We believe this "Organization of Knowledge by Use Hypothesis" to be more accurate portrayal of the nature of expertise than the simple surface/depth dichotomy.

The first indication that the real-world application of knowledge may affect cognitive structure and thus problem-solving ability came from the work of Weiser and Shertz (1983). These researchers requested that three groups of subjects sort computer programming tasks into categories based on how they might solve them. The three groups included undergraduate computer majors (novices), computer science graduate students (experts) and professionally employed programming managers (a second type or level of expert). Comparison of the sortings of novices and the graduate student experts revealed the now-familiar deep versus surface dichotomy. What is unique to their study, however, is evidence that the programming managers used a representation of the problems very different than that of both the other groups. Interviews with the managers revealed that they solve problems by delegating them to programmers, therefore their categorizations reflected the kinds of programmers to whom they would assign the problem.

In a more direct attempt to look at the differences in cognitive structure between different types of experts, Smith (1992) administered classical genetics problems to biology faculty members who teach genetics, licensed genetic counselors, and undergraduate science majors. These subjects were asked to arrange genetics problems into categories based on how they would solve them. Subjects were also asked to solve four problems. Once again faculty experts were found to sort problems according to conceptual primitives while novice students sorted by superficial characteristics of the problem statement. Surprisingly, both the counselor experts and the novice subjects tended to sort the problems according to more superficial characteristics (e.g., human versus non-human). This type of organizational scheme did not, however, lead to a lack of success on the problem-solving tasks for the expert counselors. Unlike the novices, neither group of experts tended to focus on the verbatim wording of the problem statement. These findings are consistent with the hypothesis that as expertise is attained, a person restructures his/her knowledge of the domain into a framework that is based on critical dimensions that facilitate the daily use of that knowledge, the "organization of knowledge by use" hypothesis. These data also suggest that while certain cognitive structures are supportive of problem solving, they are of little use if the subject does not apply that knowledge in appropriate situations. This ability to recognize when certain knowledge is appropriate and to apply that knowledge successfully appears to depend on how well the successful problem solver's declarative knowledge and procedural knowledge is integrated (cf. Lavoie, this Monograph).

This integration of procedural and declarative knowledge was also identified as a critical aspect of cognitive structure by Hauslein et al. (1992). In this study Hauslein et al. determined the cognitive structures of university level biology faculty, upper division biology majors, preservice teachers, and novice and experienced inservice teachers. As in the Smith study (1992), the cognitive structures of the two kinds of experts (university faculty and experienced teachers) were quite different. Using an index of difference between groups (cognitive distance) they demonstrated again the deep versus surface structure organizational dichotomy distinguishing experts from novices. Both the scientists and the experienced teachers were found to have a deeper conceptual understanding than that of the other groups. However, they also were able to differentiate the two types of experts from each other by a dimension they called fixed versus fluid thought. They inferred that this dimension reflects the degree of cross-linking in a cognitive structure (analogous to cross-links in a concept map). The scientists were characterized as having a much greater degree of cross-linking than any of the other groups. This degree of cross-linking resulted in a fluidity of thought, which was manifested in the scientists' ability to see multiple relationships among a particular concept and other categories and concepts. On the other hand, the experienced teachers had a more fixed cognitive structure. Their constant references to the teaching of each concept indicate a cognitive structure fixed on the procedural knowledge of how to teach the concepts, pedagogic content knowledge (Shulman, 1987). It is inferred from these data that the different structures were developed based on the groups particular need and use of the knowledge.

From their research Simmons and Lunetta (1993) suggest that problem-solving ability can be described on a continuum of a knowledge structure characteristic considered fragmented to integrated. They infer that if unsuccessful subjects "used fragmentary knowledge during problem solving, then it is reasonable to assume that the nature of their knowledge consisted of more generic concepts than the more specific concepts (and strong concept linkages) which are associated with integrated knowledge" of successful problem solvers (p. 168). They also concluded that successful problem solvers exhibited a superior sense of how and when certain problem-solving behaviors should be used.

These results are similar to those found by Hauslein et al.(1992), with scientists and teachers. Simmons and Lunetta's dimension of fragmented versus integrated seems to be a blend of the deep/surface and fixed/fluid characteristics. However, we can infer from both these studies that knowledge structures of successful problem solvers are easier to access and appropriately apply because of their deep and/or integrated nature. Presumably this organization allows successful problem solvers to identify the crucial aspects of a problem and to apply the appropriate conceptual and procedural knowledge to the solution of the problem. In contrast, unsuccessful problem-solvers have cognitive structures which are more content shallow (superficial) with limited connections between related ideas and concepts (fragmented). This would be observed in the use of weak heuristics and inappropriate or imprecise concepts. Unsuccessful problem-solvers in the Simmons and Lunetta (1993) study spent less time on crucial initial description and re-description behaviors which would seem to be most dependent on this interconnected knowledge. In simple terms, certain problem solvers tend to be unsuccessful, not only because they lack certain declarative and/or procedural knowledge, but also because their knowledge is not richly interconnected. They cannot distinguish relevant from irrelevant problem information that facilitates recognition of problem types. They do not have sets of procedures and concepts linked to problem types, and thus encountering a problem does not result in triggering successful solution pathways.

The connectionist modeling of these successful and less successful (expert/novice) cognitive structures would suggest that the integrated deep structure of the expert relies as much on inhibitory neural responses as on excitatory ones. The expert would relatively quickly establish an inhibitory field around an activated neural net. Like a dichotomous decision tree, the direction of useful information is found quickly, eliminating large portions of the knowledge base in favor of more fruitful routes. The successful problem solver is then able to focus on necessary information and appropriate procedures. For the less successful solver with some degree of content specific knowledge, the nodes associated with the concepts may be linked to each other with far more excitatory pathways than inhibitory ones, as was seen with the preservice teachers in the Hauslein et al. (1992) study. This situation would not enable the problem solver to discriminate useful from less useful knowledge as easily. It is therefore less likely that useful procedural knowledge would be available, such that more mental effort must be exerted in order to determine a route to useful information.

Experts whether experienced teachers or scientists, genetic counselors or university faculty would be seen, to a degree, as having a deep, integrated knowledge structure. That is, having a hierarchically arranged cognitive structure of useful and accessible domain specific knowledge and procedures. The difference observed between these experts seems to be one of development, (i.e., learning and experience) of this knowledge, considering the effect purpose and use has on the structure and restructuring of that knowledge.

Knowledge Structures and the Development of Problem-Solving Skills

Inherent in the research on problem solving is the desire to determine the best methods of teaching both transferable problem-solving skills and domain-specific skills to students. As mentioned earlier, most expert/novice studies presume that the delineation of expert cognitive structure and problem-solving behaviors will identify a cognitive structure model of expertise. The tacit assumption is that, once the differences between novice (students) and expert (faculty) are understood, the task will simply be to design instruction that decreases those differences, that is, that helps students become more like experts.

Recommendations Based on the Surface/Depth Dichotomy

As described previously, early research suggested that differences between expert and novice knowledge structures were determined by the surface/depth (or fragmented/integrated) dichotomy. Some studies have further suggested that graduate students (or others who are somewhere in between experts and novices) have knowledge structures and exhibit problem-solving behaviors that are intermediate to the two extremes. Within this paradigm, problem-solving ability therefore is assumed to progress along a single dimension and instruction focuses on helping students learn in such a way that their knowledge is structured around the recognized conceptual primitives of the discipline. This is, in fact, the basis for "Concepts of . . ." courses so common in college curricula and the basis for the instruction provided by many intuitively good teachers. Such instruction focuses, not only on the concepts themselves, but also on the connections between them. This is in direct contrast to the frequently lamented form of science teaching and learning which is essentially the memorization of literally thousands of terms, definitions, and other bits of unconnected information.

Knowledge structure research suggests that this intuitively good instruction can be improved in at least four ways. First, we suggest that the development of a preferred structure for a body of knowledge is enhanced if it is taught explicitly. The teacher in such a classroom not only teaches the content in an organized manner focusing on basic concepts, but s/he also externalizes that organization. S/He spends time making his/her internal knowledge organization transparent and helps students understand how the pieces fit. In other words, the instruction focuses explicitly on both the conceptual primitives and on the relationships among them. In times past, instructors have attempted to achieve these goals by sharing with students outlines for their lectures or by making frequent allusions to previous lessons. That teachers have not always explicitly focused on these connections (e.g., the connection between meiosis and Mendelian genetics) has been widely criticized.

Second, teachers can enhance the development of knowledge structures by using tools that allow them to be visualized and manipulated. Most of these tools are related in some way to the "concept maps" first popularized by Novak who calls such well connected knowledge "meaningful understanding" (Novak & Gowin, 1990). Concept circles, for example, are a simplified method used with elementary students (Wandersee, 1990). More recently, computer tools such as SemNet (Fisher, 1990) are available for this purpose. A growing body of evidence is now available demonstrating the value of using such tools (see JRST, special issue 10, volume 27, 1990).

Third, we suggest that the understanding of concepts can be enhanced by practicing the categorization of exemplars of those concepts. A student's understanding of the term "gene" for example is clearly enhanced and the meaningful connections between the related nodes in his/her cognitive structure are increased when that student is able to properly distinguish monohybrid (one-gene) problems from dihybrid (two-gene) problems. In order to perform such a task, the student must first learn the essential characteristics exhibited by members of each category/concept and then must learn to recognize those attributes in specific instances These characteristics then become the nodes in the knowledge structure being built, each linked in a specific way to the concept being learned. Like many of the recommendations that arise from this body of research, the utility of categorization training in actual classroom settings has not been carefully tested. Our anecdotal experience with this tool, however, suggests that this is an area worthy of future investigation.

Fourth, knowledge structure development is enhanced by testing that requires the development of explicit knowledge structures with well developed connections. "What you test is what you get" is a truism of curriculum development. Using concept maps and computer mapping tools is likely to have a very limited effect on learning if the students know that their exam will only require matching terms with definitions. Implementing this kind of testing, of course, may be difficult for some teachers. To begin with, the teacher might simply use "compare and contrast" test items or other questions that ask the student to explain the relationships between two concepts. At the other end of the spectrum, as a final examination Novak sometimes asks students to concept map the entire course (Novak, personal communication, 1992). Assuming that a well-constructed knowledge structure is equivalent to a meaningful understanding, we could also recommend the use of an occasional transfer problem as an assessment item.

Recommendations Based on the Organization of Knowledge By Use Hypothesis (Smith, 1992)

As described previously, in contrast to the early surface/depth dichotomy findings, the "organization of knowledge by use" hypothesis proposes that experts organize their knowledge on the basis of critical dimensions that facilitate the daily use of that knowledge (Smith, 1992). Thus there are different kinds of experts, each of which can be expected to have a knowledge organization that differs based on their respective experiences and the ways in which they typically use that knowledge. As an individual accumulates experience employing his/her declarative and procedural knowledge, that person's cognitive structure is fine tuned to meet the specific needs of the environment, i.e., becoming a genetics counselor or a teacher. For the teacher that knowledge must be functional in the social context of the classroom and the teachers' lounge; for the scientist that knowledge must be functional in the community of peers.

With continuing experience in which the current cognitive structure is adequate to the demands made upon it, the individual's knowledge organization tends to become fixed. Hauslein et. al. (1992) point out that all experienced teachers in their study referred to concepts in a pedagogical context. Concepts presented in the sorting task were seen as things to be "talked about" in class. It was suggested that as a teacher gains knowledge and experience the specific cognitive structures used for teaching are built upon a framework of purpose, called pedagogic content knowledge. On the other hand, research biologists used the superordinate framework of evolutionary theory to organize their cognitive structure while the teachers saw evolution more as a separate topic among other topics taught in the biology classroom (Hauslein, et al., 1992). Practicing scientists would be expected to appear more fluid in their cognitive structure because the superordinate framework of their cognitive structure would be the current philosophy of that science, which they are often reshaping.

The organization by use model (Smith, 1992) when applied to pre-professional and professional education implies that instructors need to have a practical understanding of the ways in which their current students will be required to use that course knowledge in the future. Case-based learning which is widely employed in law schools and problem-based learning which is widely touted in medical education are two examples of curricular designs that carefully attend to such concerns.

On the other hand, what does the "organization of knowledge by use" hypothesis have to say about teaching non-science majors and K-12 instruction? Given that the pace of change in modern society is so rapid that it is impossible to predict how our students will be called upon to use their knowledge, how are we to encourage them to build that knowledge? First of all, it seems important for science teachers to recognize that our goal must be to prepare our students with understandings and problem-solving skills that will not only enhance their success in subsequent course work but also prepare them to use that knowledge as appropriate in their lives outside the classroom. The exact form such instruction should take is unclear but should clearly include a focus on (a) the most basic explanatory concepts of the field, (b) how these concepts can be applied to making decisions in the real world, and (c) thinking and reasoning skills that are transferable to non-classroom contexts. (For example see Roth this Monograph.)

Another recommendation is to focus not only on the content but also on structuring and restructuring knowledge to meet the demands being placed on the individual. We propose that the goal of instruction is a knowledge structure of the domain where concepts are encoded as they relate to each other and where the importance of the organization of that knowledge is understood. Curriculum and instruction should be developed which would not only allow for the acquisition of knowledge, but concentrate on opportunities for the students to structure and restructure their knowledge, building a more and more complex hierarchical structure with each learning cycle.

Considering how this research should be applied to the instruction of individuals who are not pre-professionals raises another crucial question. Much of this research is based on the tacit assumption that the science instructor's goal is to produce experts in science. In point of fact, of course, the vast majority of our students will not become scientists. Furthermore, it is now clearly documented that the development of true expertise in most fields requires several thousand hours- much more time than most of our students will spend studying science. Therefore, we believe that the time has come to question this essentially unquestioned (and mostly unrecognized) goal (see Smith, 1992, March; Willson, 1990).

We propose that the goal of our instruction is not to produce experts but to produce competence, that we should seek to help students develop cognitive structures in such a way that the knowledge and skills they acquire can be effectively applied as needed in situations beyond our classrooms. In other words, our aim is for (non-preprofessional) students to learn science in such a way that it will help them to be successful in both their personal and professional lives as non-scientists (non-experts).

Summary and Conclusions

Whatever the future context in which individuals may be required to apply their knowledge and whatever organization of that knowledge may be most conducive to navigating that situation. Individuals are more likely to be successful if they understand the basic concepts involved, the relationships among those concepts, and both the generally applicable and domain-specific procedures and reasoning skills that are appropriate to any given situation. The research reviewed here strongly supports the conclusion that successful problem solving not only requires that this knowledge be adequate, but also that it be organized appropriately. Many instructional recommendations have been made for how one might address this task.

Many of these recommendations derive from the two primary understandings of expertise reviewed--the surface/depth dichotomy and the organization of knowledge by use hypothesis. It is important to understand that the two explanations are not mutually exclusive. We propose that the organization of knowledge by use hypothesis is broader than but also includes the surface/depth dichotomy. In this view, educator experts organize their knowledge according to conceptual primitives because this structure is conducive to the way they use that knowledge (in teaching). Other non-educator experts have also built cognitive structures that are conducive to successfully applying that knowledge in their context, but that organization may or may not center around conceptual primitives. Expertise and problem-solving success clearly does require an understanding of the basic concepts of the domain, but these concepts may or may not form the framework of an expert's cognitive structure.

We have also proposed that the goal of instruction in science problems solving is to produce competent citizens more than to produce experts. This goal revision raises some important questions. How do successful students differ from experts (and from unsuccessful students)? How should we define success? What level of competence do we aim for our students to achieve? These questions also impact research design. Perhaps it is not the Ph.D. expert that we should be studying so much as the successful non-expert. What kinds of tasks should we be using as measurements of this "success?"

Although there is clearly no definitive route to this instructional goal, we propose that our students are more likely to make successful use of their knowledge if we have helped them to understand the following:

1) that knowledge is organized or interconnected and that they come to understand a body of knowledge and to be able to use that knowledge only when they grasp those interconnections;

2) that different professions or situations place different demands on the way in which knowledge is to be accessed and used;

3) that some knowledge organizations may be more useful in one situation than in another so that a new situation may require that you reorganize what you know;

4) that rote memorization, i.e., knowledge that is not organized so as to be useable in some future setting, may have no value outside of the anticipated examination;

5) that learning how to organize knowledge in effective ways is as important as learning the meanings of the individual concepts involved.

Acknowledgement

We wish to thank Mr. Marcus Durham for assistance with manuscript preparation.

References

Adelson, B. (1981). Problem solving and the development of abstract categories in programming languages. *Memory & Cognition, 9*, 422-433.

Adelson, B. (1985). Comparing natural and abstract categories: A case study from computer science. *Cognitive Science, 9*, 417-430.

Anderson, J. R. (1985). *Cognitive psychology and its implications.* New York: W.H. Freeman & Company.

Anderson, O. R. (1992). Some interrelationships between constructivist models of learning and current neurobiological theory, with implications for science education. *Journal of Research in Science Teaching, 29*, 1037-1058.

Anderson, O. R., & Demetrius, O. J. (1993). A flow-map method of representing cognitive structure based on respondents' narrative using science content. *Journal of Research in Science Teaching, 30*, 953-969.

Bodner, G. M. (1991). A view from chemistry. In M. U. Smith (Ed.). *Toward a unified theory of problem solving: Views from the content domains* (pp. 21-34). New Jersey: Lawrence Erlbaum Associates.

Champagne, A. B., Gunstone, R. F., & Klopfer, L. B. (1985). Effecting changes in cognitive structure among physics students. In L. H. T. West & A. L. Pines (Eds.), *Cognitive Structure and Conceptual Change* (pp. 163-187). Orlando, FL: Academic Press, Inc.

Champagne, A. B., Klopfer, L. E., Desena, A. T., Squires, & D. A. (1981). Structural representations of students' knowledge before and after science instruction. *Journal of Research in Science Teaching, 18*, 97-111.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Diekhoff, G. (1983). Testing through relationship judgements. *Journal of Educational Psychology, 75*, 227-233.

Estes, W. K. (1991). cognitive architectures from the standpoint of experimental psychologists. In M. R. Rozenweig, & L. W. Porter (Eds.). *Annual Review of Psychology, 42*(1), 1-28.

Estes, W. K. (1993). Concepts, categories, and psychological science. *Psychological Science, 4*, 143-153.

Fennker, R. M. (1975). The organization of conceptual materials: A methodology for measuring ideal and actual cognitive structures. *Instructional Science, 4*, 33-57.

Fillenbaum, S., & Rapoport, A. (1971). *Structures in the subjective lexicon.* New York: Academic Press.

Fisher, K. M. (1990). Semantic networking: The new kid on the block. *Journal of Research in Science Teaching, 27*, 1001-1018.

Gorodetsky, M. & Hoz, R. (1985). Changes in the group cognitive structure of some chemical equilibrium concepts following a university course in general chemistry. *Science Education, 69*, 185-199.

Greeno, J. G. (1980). Theory and practice regarding acquired cognitive structures. *Educational Psychologist, 10*, 117-122.

Hauslein, P. L., Good, R. G., & Cummins, C. L. (1992). Biology content cognitive structure: From science student to science teacher. *Journal of Research in Science Teaching, 29*, 939-964.

Hauslein, P. L., Good, R. G., Cummins, C. L. (1992). Biology content cognitive structure: From science student to science teacher. *Journal of Research in Science Teaching, 29*, 939-964.

Hayes, J. (1980). *The complete problem solver.* Philadelphia: The Franklin Institute Press.

Jackson, N. E., Grassia, J. , & Wolfe, D. (1988, April). *Knowledge structures and problem-solving performance of expert statisticians and expert and semi-novice users of statistics.* Paper presented at the annual convention of the American Educational Research Association, New Orleans, LA.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, Massachusetts: Harvard University Press.

Kleinbaum, D. G., & Kupper, L. L. (1978). *Applied regression analysis and other multivariable methods.* North Scituate, MA: Duxbury Press.

Knowlton, B. J., & Squire, L. R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science, 262*, 1747-1749.

Kruskal, J. B. (1977). Multidimensional Scaling and other methods for discovering structure. In K. Eslein, A. Ralston, & H. Wilf (Eds.), *Statistical Methods for Digital Computers (Vol. 3)* (pp. 296-339). New York: Wiley.

Larkin, J. H. (1981). Cognition of learning physics. *American Journal of Physics, 49*, 534-541.

Larkin, J. H., Heller, J. I., & Greeno, J. G. (1980). Instructional implications of research on problem solving. *New Directions for Teaching and Learning, 2,* 51-65.

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

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

Maloney, D. P. (1987). *Effects of surface structure and deep structure on novice rules of physics.* Unpublished manuscript, Creighton University, Omaha, NE.

McClelland, J. L., & Rumelhart, D. E. (Eds.). (1986). *Parallel distributed processing, explorations in the microstructure of cognition. Vol. 2. Psychological and biological models.* Cambridge, MA:MIT Press.

McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology, 13*, 307-325.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*, 937-949.

Novak, J. D., & Gowin, D. B. (1990). *Learning how to learn.* New York: Cambridge University Press.

Perez, R. S. (1991) A view from trouble-shooting. In M. U. Smith (Ed.) *Toward a unified theory of problem solving: Views from the content domains.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Perfetto, G. A.., Bransford, J. D., & Franks, J. J. (1983). Constraints on access in a problem solving context. *Memory and Cognition, 11*(1), 24-31.

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accommodations of scientific conception: Toward a theory of conceptual change. *Science Education, 66*, 211-227.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12*, 1-20.

Roehler, L. R., Duffy, G. G., Hermann, B.A., Conley, M., & Johnson, J. (1988). Knowledge structures as evidence of the 'personal': Bridging the gap from thought to practice. *Journal of Curriculum Studies, 20,* 159-165.

Rogoff, B., & Lave, J. (Eds.). (1984). *Everyday cognition: Its development in social context.* Cambridge, Massachusetts: Harvard University Press.

Rosch, E., & Lloyd, B. B. (Eds.). (1978). *Cognition and categorization.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Roth, W.-M., Gabel, D., Brown, L., & Rice, D. (1992). A combined method of INDSCAL and cognitive mapping for describing changes in students' in structure. *Educational and Psychological Measurement, 52*, 769-779.

Rumelhart, D.E., & McClelland, J.L. (Eds.). (1986). *Parallel distributed processing, explorations in the microstructure of cognition. Vol. 1. Psychological and biological models.* Cambridge, MA:MIT Press.

Santos-Gomez, L.., Wooliscroft, J. O., & Mennin, S. P. (1990, April). *Clinical reasoning processes and knowledge structures of medical students from a problem-based and a traditional curriculum.* Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.

Schwab, J. J. (1978). Education and the structure of the disciplines. In I. Westbury, & N. J. Wilkof (Eds.), *Science, Curriculum, and Liberal Education: Selected Essays* (pp. 229-272). Chicago: The University of Chicago Press.

Schvaneveldt, R. W., Durso, F. T., & Dearholt, D. W. (1989). Network structures in proximity data. *The Psychology of Learning and Motivation, 24*, 249-284.

Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., Cooke, N. M., Tucker, R. G., & De Maio, J. C. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies, 23*, 699-728.

Shavelson, R. J. (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology, 63*, 225-234.

Shulman, L.S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*, 1-22.

Siegler, R. S. (1978). The origin of scientific reasoning. In R. S. Siegler (Ed.). *Children's thinking: What develops?* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Simon, H. A. (1980). Problem solving and education. In D. T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, H. A., & Simon, P. A. (1979). Trial and error search in solving difficult problems. In H. A. Simon (Ed.), *Models of thought.* New Haven: Yale University Press.

Simmons, P. E., & Lunetta, V. N. (1993). Problem-solving behaviors during a genetics computer simulation: Beyond the expert/novice dichotomy. *Journal of Research in Science Teaching, 30*, 153-173.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, M. U. (1990). Knowledge structures and the nature of expertise in classical genetics. *Cognition and Instruction, 7*, 287-302.

Smith, M. U. (1991). A view from biology. In M. U. Smith (Ed.). *Toward a unified theory of problem solving: Views from the content domains* (pp. 1-19). New Jersey: Lawrence Erlbaum Associates.

Smith, M. U. (1992). Expertise and the organization of knowledge: Unexpected differences among genetic counselors, faculty, and students on problem categorization tasks. *Journal of Research in Science Teaching, 29*, 179-205.

Smith, M. U. (1992, March). *Characteristics of the successful novice problem solver: Implications for theory and Instruction.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, MA.

Smith, M. U., & Good, R. (1984). Problem solving and classical genetics: Successful versus unsuccessful performance. *Journal of Research in Science Teaching, 21*, 895-912.

Takane, Y., Young, F. W., Lewyckj, R. J., & de Leeuw, J. (1978). How to use ALSCAL 3: *A nonmetric multidimensional scaling program with several individual differences options.* Chapel Hill: University of North Carolina at Chapel Hill.

Thro, M. P. (1978). Relationships between associative and content structure of physics concepts. *Journal of Educational Psychology, 70*, 971-978.

Uche, S. C. (1987). *Students' cognitive structure of some general biology processes.* (Doctoral Dissertation, State University of New York at Buffalo). Dissertation Abstracts International, 48, 1158A.

Veldhuis, G. H. (1990). The use of cluster analysis in categorization of physics problems. *Science Education, 74*(1), 105-118.

Wandersee, J. H. (1990). Concept mapping and the cartography of cognition. *Journal of Research in Science Teaching, 27*, 923-936.

Weiser, M. & Shertz, J. (1983). Programming problem representation in novice and expert programmers. *International Journal of Man-Machine Studies, 19*, 391-398.

West, L.H.T., Fensham, P.J., & Garrard, J.E. (1985). Describing the cognitive structure of learners following instruction in chemistry. In L. H. T. West & A. L. Pines (Eds.), *Cognitive structure and Conceptual Change* (pp.51-59). Orlando, FL: Academic Press.

Willson, V. L. (1990). Methodological limitations for the use of expert systems techniques in science education research. *Journal of Research in Science Teaching, 27*, 69-77.

WHAT HAPPENS TO A ROCK WHEN YOU THROW IT IN THE WATER? DOING HIGH SCHOOL PHYSICS THE PHYSICISTS' WAY

Wolff-Michael Roth
Faculty of Education
Simon Fraser University
Burnaby, B.C., Canada, V5A 1S6

Introduction

Vignette

Miro stared out of the window overlooking the city. He could see some of the old university buildings that dated back to the early fifteenth century, others that housed the labs where several Nobel prize winners had worked. He had just finished a meeting with Kling, one of the two leaders of Floss Labs, concerning yet another problem Miro had identified in their tooth brushing problem. Floss Labs, one of approximately 100 federally funded special research institutes focused its investigations on the oral cavity. One of their current projects studied whether it was better to brush teeth with a vertical on a horizontal stroke. For months the research team has been working on this problem, and the best way Miro could describe their research was "mucking about." This was a strange experience, for he had expected something different from the lab headed by Kling and Ganiel both of whom had doctorates in physics and in dentistry, and who had made a number of published discoveries. Miro thought that his "mucking about" was so different from his school and university experience of physics.

He remembered particularly that one physics exam. It was the rocket problem, which, as he found out many years later, has been used repeatedly to learn about students' ideas about mechanics (Clement, 1982). Slide rule and pencil were the only things the policies at Bohr College allowed him to bring. Miro knew that there was a right answer to the problem, because the teacher would not have asked it otherwise, but the solution he thought of looked too simple. It was obvious that the rocket should move sideways, then go down for a while, and then go sideways again once the engines are turned off. After pondering for a while where the trick was in the question,

Chapter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for scientific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six*. Manhattan, KS: Ag Press.

he decided to write down his answer on the space provided. Although he generally found physics difficult and regretted his decision that got him into this course, he had not expected to receive only a few part marks on this problem. Much of his experience of physics had been like this, with the exception of some advanced laboratory courses where they did some "real" research.

Now, however, things were different. During this first visit to Floss Labs, he had already found that the work of the lab assistants, technicians, and mechanics was not different from that of lab assistants, technicians, and mechanics in other fields. He had thought to see the real discovery work and problem solving once he started working with Kling. But things didn't seem to change when he began on his first, the Brushing Problem. First there had been all the technical problems which didn't seem to end. It had taken them weeks to decide about the materials for the dentures and the "plaque" and problems seemed to emerge from everywhere. They had started with the plaster models dentists use, but the spray paint used as "plaque" penetrated and did not rub off quickly enough. Over the course of several weeks, they tried a number of different materials for dentures, reduced the model to a representation of front teeth only, and had decided on a phosphorescent paint, so that they could record the remaining plaque on photographs after illuminating the teeth with UV. Although they drew on the knowledge and experience of all the people in the lab and that of many others such as profs at the university, suppliers of materials, published literature, computer programmers, and statisticians, an answer to their Brushing Problem still was not in sight. Thinking of the present again, Miro sighed and called it a day.

A few years later, long after he had left, Miro found out that the Floss Labs team, although successful on other problems, eventually had abandoned the brushing problem which had become one of the team's blind alleys.[1]

---

[1] The vignette is autobiographical and combines my experience of physics as a high school, college, and university student with my later experience in research and as teacher. My research on problem solving in physics was strongly affected by the discrepancy I had experienced between the traditional problem-solving exercises throughout most of my schooling and the laboratory research I did later.

This vignette points out two concerns which a new brand of cognitive scientists have voiced in recent years. First, there is a significant difference between the cognitive activity of people when they face word problems, puzzles, or games with rather well-defined rule systems (e.g., chess) versus problems which they face in the mostly ill-defined settings of everyday lives (Brown et al., 1989; March, 1991; Forsythe, 1993; Gooding, 1992; Hutchins, 1991; Knorr-Cetina, 1981; Latour, 1992; Lave, 1988; Lestel, 1989; Starling, 1992; Suchman, 1987). Second, the research on the cognitive activity of scientists showed that a special rationality which distinguished scientists from most ordinary people did not exist, findings which stand in contrast to many classical studies of cognitive science. Much of the evidence collected by researchers in this relatively new area of cognitive science called situated cognition (which has become a largely interdisciplinary field) shows that theories of problem solving based primarily on information processing poorly modeled problem solving in ill-defined settings although they still are reliably modeling problem solving in highly structured settings. In addition, as a whole, transfer studies are inconclusive whether aspects of problem solving can be transferred between contexts even if source and target domains are highly structured; transfer from well-defined to ill-structured domains is even less likely to occur (Lave, 1988).

Based on my autobiographical experience glossed in the opening vignette, and the research results of recent studies on problem solving in ill-structured settings, I felt that besides working more traditional word problems, high school students needed opportunities to frame and solve problems in ill-structured settings. In this way, they would have experiences similar to scientists who work on the frontiers of their knowledge rather than on familiar and well-structured problems which are the basis of many expert-novice studies. That is, in their respective settings, students and scientists define and solve problems for which they do not possess algorithmic and routine answers. In part I of this chapter, I first provide a sketch of the research approach and results of classical studies on problem solving in physics which will allow me later to show the limitations of this approach in problem-solving settings of my physics classroom. Following an outline of the new, multi-disciplinary research on physicists' cognition, I give a research-based rationale for including open-inquiry in physics teaching. Part II of this chapter provides a description of the physics course I taught and in the context of which I conducted my research on students' problem solving. In Part III I present some illustrative research findings on students' problem framing and solution finding. I conclude this chapter with some comments on specific issues arising from this new approach to teaching physics and researching students' cognition.

Information-Processing Paradigm

Classical studies on problem solving in general and in physics more specifically conceptualized the fundamental issues in the field in terms of information to be processed (Larkin, McDermott, Simon, & Simon, 1980; Larkin & Rainard, 1984; Roth, 1991a). The text and images of word problems provided all the information needed to answer the question. Analyses of problem solving were interested in how well research participants decoded the semantics of the text in terms of its informational value and then processed the found information. In order to decode texts, problem solvers need to have content relevant knowledge which was thought to be stored in long term memory (LTM); the important component for processing stored knowledge and information provided was short term memory (STM) (Roth, 1990a, b).

Classic studies on the organization of physics knowledge in LTM indicated differences between experts and novices (Chi, Feltovich, & Glaser, 1980; Ferguson-Hessler & de Jong, 1987; Heller & Reif, 1984; Roth, Gabel, Rice, & Brown, 1992; Shavelson, 1974; Thro, 1978). These studies showed that experts had organized their *declarative* content knowledge (definitions, principles, formulas, facts) in strongly hierarchical form, which differentiated between concepts on the basis of deep rather than surface structures, and integrated related knowledge very well. The knowledge organization of novices was less hierarchical, less integrated, and less differentiated so that they tended to rely on semantic surface features of the word problems. Experts and novices less frequently differed in their *procedural* knowledge regarding the manipulation of equations. Sometimes investigators use a third kind of subject matter knowledge, *problem situations*, as an important component of physics problem solving. This form of knowledge permits solvers to extract from more or less implicit situation descriptions in the word problems those features which serve as cues in searching knowledge of the *declarative* and *procedural* type (Ferguson-Hessler & de Jong, 1987).

Before individuals can solve a given problem, they need to re-present the information in a form amenable to solution. Four representations appear in the work of many solvers with physics knowledge (Chi, Feltovich, & Glaser, 1980; Larkin & Rainard, 1984). Beginning with (1) the collection of words and sentences, a solver builds (2) a new representation including the objects and relations between them specified in the text. This new representation is then converted into a third kind which is based neither on words from the vernacular nor on real objects but on (3) scientific objects (force, acceleration) and drawings (e.g., free body diagram). Finally, this representation is converted into (4) a computational one, constituted by algebraic symbols related by operators and equalities. It appears that the step from (2) to (3) is the most crucial in solving physics word problems (Champagne, Klopfer, & Gunstone, 1982; Heller & Reif, 1984).

In a classic article on problem solving in kinematics, Larkin et al. (1980) described the differences between an expert and a novice. In this domain, the amount of information required for mastery consists of about 10 "things" — concepts and laws. The comparison was done on the basis of a typical word problem:

> A bullet leaves the muzzle of a gun at a speed of 400 m/s. The length of the gun barrel is 0.5 m. Assuming that the bullet is uniformly accelerated, what is the average speed within the barrel?

There were four major differences in the problem-solving approaches. First, the expert required about one-quarter of the time needed by the novice. Second, while the novice predominantly used the working-backward heuristic (from the unknown problem solution to the given quantities) the expert the working-forward heuristic (from the givens to the solution). This was surprising, because working backward had traditionally been considered the more sophisticated heuristic. This difference, however, appeared to arise from the fact that the expert worked in a domain very familiar to him, while the novice needed to generate goals and subgoals in her search. The management of goals and subgoals requires time, which made for the differing solution time. Both approaches can be modeled in a production system (rules for producing answers according to an IF... THEN pattern) characteristic for the modeling of problem solving in classical cognitive science. The production system modeling the novice was constituted by rules such as:

IF the dependent variable in an equation is the desired quantity
    THEN try to solve the equation.
IF the values of the independent variable are not known
    THEN create a goal to find the values of these variables.

The expert on the other hand was successfully modeled by a production system with rules such as:

IF you know the values of all the independent variables in any equation
    THEN try to solve for the dependent variable.

The third difference between expert and novice was that the latter mentioned each equation to be used and then substituted values, while the former reported only the numerical results of the substitutions. Larkin et al. (1980) likened the difference in these approaches to that between compiled and interpreted computer programs. In interpreted programs (e.g., BASIC), a central control system orders the execution of each step translating each instruction in a step by step fashion into one which the machine can understand. Compiled programs first translate all the instructions into machine-understandable code, making on-line translation unnecessary. As a result, compiled programs run about ten times faster. The fourth difference between expert and novice was the representation on which they operated. The novice functioned primarily through syntactic translation, while the expert immediately generated a physical representation including objects of type (2) and (3).

Situated Cognition and Embodied Laboratory Practices

While traditional cognitive science (and artificial intelligence) has been successful in modeling problem solving of well-structured word problems in introductory mechanics or statics, researchers in their own ranks have begun to realize that it has largely failed to model everyday ill-structured problem solving (Anderson, 1990; Lave, 1988; Suchman, 1987). This failure arose from the fact that models of expert problem solving was derived in highly structured set of signals and clues, and very limited domains well known by the solver. In their everyday laboratory and discovery work, however, scientists work in ill-defined contexts and it is often unclear what the problems or relevant frames are, which aspects of the setting to conceptualize as dependent and independent variables, or which familiar solution processes to invoke. Traditional cognitive-science theories of problem solving also failed to model the situated, distributed, and embodied nature of knowing (Brown, Collins, & Duguid, 1989; diSessa, 1993; Lave, 1993; Pea, 1988; Resnick & Levine, 1991; Scardamalia & Bereiter, 1994; Suchman & Trigg, 1993; Varela, Thompson, & Rosch, 1993). As a consequence, a new branch of multi-disciplinary cognitive science has emerged. Recent studies of cognition in scientific laboratories and everyday situations have been conducted from anthropological, cross-cultural, ethnomethodological, organizational, philosophical, and sociological perspectives.

This research in cross-cultural cognition, cognition of everyday work practices, scientific laboratories, schools, and in knowledge-intensive organizations shows that (1) there is a fundamental difference between solving textbook problems and solving authentic problems and (2) physicists (including astronomers) are not endowed with a special intelligence, scientific process skills, and problem-solving heuristics that allow them to construct solutions in a rational process leading from initial (problem) to goal (solution) states (Collins, 1982; Drake, 1990; Garfinkel, Lynch, & Livingston, 1981; Gooding, 1992; Pickering, 1985; Pickering & Stephanides, 1992; Pinch, 1985; Traweek, 1988; Woolgar, 1990). These studies describe scientific work in physics as situationally contingent and circumstantial, showing innumerable false-starts, ad hoc procedures to assure the efficacy of a method, improvised repairs of prior actions and talk, and situated inquiries when trouble becomes apparent; scientific work is ridden by “superstitions”, that is, actions chosen for no other reason than that they work; and problems and solutions are negotiable without single correct solutions. In this, scientists are not different from shoppers, workers in dairy factories, or bookies and candy sellers in street markets. Scientists' selections of research questions, equipment and materials are largely determined by their immediate laboratory contexts, their specific embodied skills and local practices, the availability of financial and material resources, equipment, and information. Thus, the processes and products of science are hybrids which are marked by the same *situation-specific logic* that characterizes all other, non-scientific everyday problem solving. The models of problem solving that emerged from these studies are ethnographic and holistic, accounting as much as possible for the common-sense aspects of knowing and their embodied laboratory skills that physicists bring to work.

Bringing Real Physics into the Classroom

As pointed out above, there is mounting empirical evidence of the fundamental differences engendered in cognitive activity on textbook and real-world problems. The notion of "authentic" problem solving (problem solving with high degree of similarity to out-of-school settings) has been used in a number of programmatic statements about the improvement of teaching and school learning to narrow these differences (Brown et al., 1989; Scardamalia & Bereiter, 1994; Turkle & Papert, 1991; Wiggins, 1989). Others are more cautious about the benefits of a prescriptive value of the notion of *authentic*, or reject it as mere rhetoric. Some of the problems arise because few have actually defined conditions that make classrooms activities at least *quasi*-authentic.[2] In my work, I operated under the assumption that school physics needed to include activities with the potential of giving rise to cognitive activity resembling that of physicists who work at the frontiers of their fields but which are absent in normal physics classes. Accordingly, to be authentic, school activities had to have at least the following five features in common with scientific work (below referred to as Authenticity Features): (1) participants learn in contexts constituted in part by ill-defined problems; (2) participants experience uncertainties, ambiguities,

---

[2] I am actually more cautious about the notion of *authentic activities* in schools than the present chapter might indicate. For, one of the fundamental characteristics of everyday scientific and non-scientific activity is the purposefulness with which they are conducted. Its is usually some product, whether this is a best-deal in a supermarket, a deer for feeding the family, a new big-bang, cancer, or AIDS theory, or an increase in a company's output, and so forth. In the course of reaching their goals, people apply old and learn new useful practices and resources (knowledge). Thus, learning in everyday settings is the outcome of purposeful activity. On the other hand, schools are designed to transmit culture and impart skills for their own sake. As such, schools are antithetical to purposeful learning.

and the social nature of learning and knowing in science; (3) learning is driven by the current knowledge state of the participants; (4) participants experience themselves as members of learning communities in which knowledge, practices, and discourses are shared; and (5) in these communities, members can draw on the expertise of more knowledgeable others, and on any suitable resource which could enhance their learning. In the examples from my classroom presented below, all five conditions were met. First, students defined problems and solutions on their own; in the course of constructing solutions, students identified new problems which they had to solve as a matter of course before they could go on to find an answer to their focus question. Second, in the process of constructing problems and solutions, students dealt collaboratively with the uncertainties inherent in investigations of a *first-time-through* nature. Third, students' learning was automatically driven by their current knowledge, because the research problems were of their own design. Fourth, the classroom procedures which I established encouraged students to interact with each other within and across groups and to make active use of the school's resources (classroom, library, physical plant, and kitchen). Finally, students were encouraged to draw on the expertise of anyone they might consider more knowledgeable on a relevant topic, including myself, the laboratory assistant, other science and mathematics teachers, carpenters and mechanics, and so forth. The following are descriptions of students' framing and solving of physics problems in such an environment derived from a series of investigations (Roth, 1992, 1993, 1994, in press; Roth & Roychoudhury, 1993; Roychoudhury & Roth, in press). These descriptions will depict students' problem solving in physics from a cognitive-science perspective which emphasizes the situated nature of knowing and learning. That is, the students' problem-solving activities are the results of interactions between cognitive agents, real objects, and the physical environment. The robustness of understandings arising from these interactions comes not only from individual minds (as traditional cognitive science assumed) but from the nature of the learning system that includes social organization (in small groups and classroom communities) and the mediational means of tools, materials, and language (Ueno & Arimoto, 1993).

An Alternative Physics Course

At the beginning of each unit, the students received a handout, outlining some suggested laboratory activities, readings, and a minimum number of qualitative and quantitative word problems to be done from any of the various textbooks. Besides the student investigations that constituted the core of the course around which all other activities revolved (about 70% of class time), there were four other types of activities. First, students read relevant sections from the main textbook and at least one additional source; at the end of a unit they summarized key ideas by constructing concept maps. Then, they submitted each week six to eight textbook problems of their own choice, provided that problems came from different textbook sources but were on the current topic.[3] Third, they prepared short notes on the biographies of scientists or wrote essays on special topics not normally covered in the textbook. Finally, we met from time to time as a whole class to answer questions, discuss the designs of experiments, to introduce new equipment, and so forth.

The students in this course spent much of their time (6-7 of the 9 periods in a 2-week cycle) on experiments of their own design (Authenticity Features 1 and 2). Outside of the scheduled classroom time, the students used the lab and its facilities during their spare periods, after school, in the evening, or during the weekend. Because of the residential nature of the school, many students worked on their own about one evening per week. The classroom was always open in the sense that students from any section could work on their experiment unless it interfered with the work of the students scheduled for that period. This arrangement provided yet another avenue for students to see and talk about different ideas (Authenticity Features 4 and 5). Because of these arrangements, the physics laboratory was a communal place bustling with activity throughout the day and the evening. It was a place for learning, and many students from different groups came there to work on projects for other subject areas (mathematics in particular because of the software available on the computers). Students usually worked in groups of 3 to 4 which remained very stable throughout the year. Each year, there were only a few instances in which two or three students had to be asked to change groups in order to improve their level of participation, or the quality of student-student interactions.

The social aspect of communities of learning (Authenticity Features 4 and 5) was further stressed in that one period per 2-week cycle was used for whole class discussion, review, sharing research results with other students, and for introducing new tools such as measurement instruments or computer programs. In my role of expert in this learning community, I often began a new unit with a range of demonstrations during which I also introduced new instruments, apparatus, or software (Authenticity Feature 5). Immediately after the demonstrations, students began to "play" with the equipment and where encouraged to think about an investigation, to formulate a research problem (often in the form of a focus question) and planning the data collection; they usually took one class period for planning an investigation. The students then set up the apparatus, collected the data, and submitted the data to a computer-based mathematical and graphical analysis (2-4 periods). Each group discussed its results,

[3] By assigning readings and a minimum number of textbook problems I ascertained that the formal curriculum requirements set by the province were satisfied.

consulting with other groups, myself, or the assistant (1 period), and then prepared a report.

The science laboratory was well equipped, and the operating budget of the science department was quite generous when compared to many other schools (resources, Authenticity Feature 5). There were a few air tracks, and Macintosh and Apple II computers available in the physics laboratory. The data collection interfaces included temperature probes, motion timers, advance interface equipment for multiple voltage inputs, a force meter and additional kits for building photo gates. After building instruments from kits, students could interface these with computers for fast data collection and processing. The Macintosh computers were loaded with mathematical modeling software, statistics packages, simulation programs, programming languages, spreadsheets, word processors, and graphing/painting programs. These tools allowed students to focus on ideas and patterns across repeated cycles of measurement rather than on the mechanics of collecting one set of data. Most importantly, I kept a steady supply of handyman tools, masking tape and Scotch tape, glue, cardboard, styrofoam, string, wire, paper, various liquids such as cooking oil, coolant, various motor oils and other odds and ends which could be used to make things "on the fly" or to repair equipment. Students also brought materials from home or organized the purchase of special materials and items, such as liquid nitrogen for experiments in superconductivity, yo-yos for the study of motion with changing acceleration (Authenticity Feature 3). Most of this material and equipment was kept directly in the classroom so that it was easily accessible to the students. A small in-class library provided easy access to written reference materials such as old textbook series and science encyclopedias.

In this setting, the students in consultation with the laboratory assistant and myself developed very interesting and quite complex investigations. In every instance, however, the investigations began with a phenomenon of interest identified by the students themselves. The interested reader will find a comprehensive description of some of these activities and how they were realized in Roth (1991b).

The illustrative data presented below are based on the video- and audiotaped classroom and interview material (footage showing the students in all stages of their experiments, interviews regarding their views of science, knowing and learning, and teaching strategies) and student- and teacher-produced artifacts (laboratory reports, feedback, research memos). A teaching and laboratory assistant (who held an MS in physics) helped in the collection of the video materials and constituted an additional resource person to the students. For two graduating classes, data collected during their junior and senior year physics courses provided extensive evidence for the development of their embodied laboratory skills, views about science (physics in particular), and the processes and products of scientific knowledge.

Most students came from middle and upper-middle class families of professionals (doctors, lawyers, and teachers) and business people. The school also attracted students from lower income brackets, in part to strengthen its competitive sports program by offering bursaries. The junior year qualitative physics course attracted a large proportion of the grade level (55 - 69% in the three years of research) mainly because of two reasons. First, for many programs in science, engineering, and medicine, the local universities had a high school senior level physics requirement; taking the junior level course kept students' university options open. Second, this physics course was known, because of its structure, as a "fun course." However, although generally college bound, few students from this school eventually select university programs in science or mathematics, but predominantly opt for careers in business, law, and engineering.

Problem-Framing and Solution-Finding Processes

The data presented here were collected in the context of a series of investigations concerned with objects falling under the effect of buoyancy and friction. "What happens to a rock when you throw it in the water?" was a question that came up repeatedly during the class discussion of Carl, Jim, and Pete's (CJP) experiments. CJP investigated this topic for several weeks during the fall of their junior year and returned to the same topic during their senior year.

With little or no experience in a domain such as Newtonian mechanics, students' research problems were stated simply and often related one dependent and one independent variable. Over time, their research problems became increasingly complex. At the same time, students pursued sets of interconnected questions which turned their investigations into research programs. In the process of implementing an inquiry, the students framed many locally emerging problems which they needed to resolve before they could continue with their global problem. At times, students framed research problems for which they did not achieve the results they expected, a phenomenon I labeled "blind alley." The many choices students made affected the data they collected and the conclusions they could draw from the experiment. In some instances, students had constructed solutions before the problems they solved. Thus, the problems followed the solutions in a reversal of traditional problem solving. The construction of problems was not always confined to the immediately local, but could also happen over longer cycles of data collection, data interpretation and redesign of an experiment. The students' prior plans and *a posteriori* descriptions of their activities left their actual problem-solving activities undetermined in interesting ways. That is, their plans could have been realized in many different (and contradictory) ways, and their after-the-fact accounts were reconstructed (and still underdetermined) accounts of what they had actually done. Especially in the initial stages in a new domain under investigation, students often used concrete objects to analyze complex situations rather than explicitly stating variables. Their understanding of situations was elaborated in terms of the behavior of material objects. Later, they developed more "abstract" forms of representing the phenomena of interest. Besides the use of concrete objects in elaborating an initial understanding of phenomena, students began by narratively describing phenomena they wanted to observe. Through their negotiations, often during the experiment itself, and especially with increasing familiarity in a domain, students analyses became more abstract.

In the following two sections I will describe in more detail a few characteristics of students' work during the phases of setting problems and solving them. The significance of these descriptions resides in the yet-unknown problem-solving behaviors of students in ill-defined settings and in the research approach which does not limit cognition exclusively to decontextualized mental processes.

Processes of Problem Setting

Over the duration of their two physics courses, many students developed very interesting series of investigations. Given that the students always framed their problems on the basis of their current knowledge, the research problems initially related only one dependent and one independent variable. Over time, the research problems became increasingly complex. The following example illustrates (a) the process of framing problems in this open-inquiry lab, (b) the evolving nature and coherence of students' research questions, and (c) the meaningfulness of problem framing from a student perspective. (Detailed analyses of the physics students problem-related activities and beliefs have been reported in Roth, 1992, 1993, 1994, in press; Roth & Roychoudhury, 1993; Roychoudhury & Roth, in press.) Table 1 is composed of three excerpts from my data base.[4]

The episode in Data Set 1.a was recorded as part of a 1-hour session during which CJP were to frame a problem, plan a solution, design an experiment, and, if they had time, to pilot various aspects of their investigation. CJP had already discussed for some time the

---

[4] I used the following transcription conventions:

- [...] for pauses; each triplet of dots corresponds to one second.
- (??) for words that could not be deciphered; each question mark stands for one word.
- ,.?! punctuation to indicate whether the utterance was heard as a question or exclamation or whether the utterance came to a full or part stop.
- :::::: for indicating the omission of irrelevant talk.

Table 1. Three different types of data illustrative of the work done by students relating to the framing of problems in this open inquiry environment.

Data Set 1.a.

During a planning session, the following conversation ensued between Jim and Carl (fourth experiment in Data set 1.b.):

AUDIO	VIDEO
1. J: What are we going to do?	
2. C: We are going to do different liquids, same mass, same volume.	
3. J: That has nothing to do with what we did.	(doodles)
4. C: We cannot (??) [......] accurately [...] we cannot continue this experiment using the same liquid [...] compare the cubic with the tear drop that wasn't quite a tear drop	(doodles)
5. [...] this compare to that for these are not constant shapes.	(draws two different tear drops)
6. We are going to take a constant shape, change the liquid it goes through [...] we can measure how this liquid changes [...]	
7. J: This is based on the shape [...] this in general should be slower than this in general.	(points to a less ideal tear drop)
8. C: This isn't constant shaped, though.	
9. J: Why?	
10. C: Because we can't make a perfect square [...] cube with plasticine.	

Data Set 1.b.

CJP conducted a series of experiment on falling objects guided by the overarching theme of "What happens to a rock when you throw it in the water?" They conducted a series of experiment including:

1. What is the motion of a rock in free fall?
2. What is the motion of an object on an inclined plane?
3. How does hydrodynamics affect the free-fall through water?
4. How viscosity of liquids affects the free-fall of a rock.
5. To investigate harmonic oscillation damped by various liquids.
6. To investigate the process of a yo-yo in motion under the force of gravity.

7. What is the relationship between acceleration and air friction?

Data Set 1.c.

Excerpt from students laboratory report for experiment #3

> We found this lab interesting because we were pursuing something of our own interest. It was a lab of our own design, and creation, and it was a lot of fun to do. This lab taught us the value of design in things like air planes and race cars, even the aerodynamics of our own cars, and how that affects fuel consumption. It also brought up more questions concerning form, and shape, and the effect of density of liquids.

---

implications of their previous experiment (Data Set 1.b., problem #3, in which they investigated objects of 4 different shapes fall in a column of water), but could not come to an agreement how and in which direction to proceed. In response to Jim's question of what to do next, Carl suggested to take one object ("same mass, same volume") and drop it through various liquids (2). Jim objected, claiming that this new problem had nothing to do with their original idea of testing the hydrodynamics of different shapes (3). Carl countered by elaborating his reasons for changing the focus of their problem (4-6). First, the use of plasticene as a material for the dropping object limited the accuracy with which they could make "ideal" cubes and tear drops (4). In support of his argument, he drew two tear drops of different shape, questioning their ability to produce perfect shapes on which to base their analysis (4-5). By questioning the possibility for deciding which of two tear drops was the ideal one, Carl pointed to the impossibility of making a fair comparison between a cube and a tear drop in an experiment. To avoid this dilemma, he suggested a slight change in topic. By choosing to work with one object, they would avoid the question of shape and instead study the effect of liquids on the motion of the object (6). However, Jim did not give up and persisted in making the object's shape problematic[5] (7). Carl reiterated the problematic production of "perfect" shapes (8, 10).

The episode illustrates several aspects of problem framing. First, students framed new problems on the basis of prior experiments. Because of the problems with producing ideally-shaped objects, Carl suggested a continuation of their experiments by investigating the effect of liquids. That is, he rendered problematic the environment of the falling object. Second, just what an interesting and feasible problem for research was arose from the negotiations in which students rendered problematic various aspects of their previous experience. Here, Carl suggested to abandon the comparison of the hydrodynamics of various shapes because of the problems with producing the ideal shapes to be

---

[5] I use the term "problematic" in Wheatley's (1991) sense. Accordingly, a situation/text is not a problem *a priori* but has to be constructed as such; in that case it has become problematic.

compared. Jim, on the other hand, found such a direction problematic because it changed the focus of their inquiry from the question of an object's shape to that of environmental effects. Both of these aspects give rise to a new view of the nature of problems. According to this view, authentic problems arise from the interaction of individuals with others (social setting) and with the environment (physical setting). Thus, ontologically, problems do not exist per se, but are constituted in the relationship of individuals who act in physical and social settings. For example, a problem framed by students in this classroom may not be a problem at all for a physicist-in-training with some experience in hydrodynamics. Yet it could be a problem for another physicist who studies the molecular aspects of objects moving through liquids.

The episode illustrates two other phenomena. First, when students began their investigations in new domains, they often did not analyze these in terms of (dependent, independent, controlled) variables. Rather, they began with descriptive accounts of what materials and events were to be involved; e.g., here Carl suggested to study different liquids (2). Only through their negotiations and later in the context of their developing expertise with various liquids did they begin to focus on density and viscosity as two possible distinguishing features of liquids. Second, the students used various means to facilitate their negotiations during problem framing. Here, Carl drew two "tear drops" of different shape (5). These drawings became inscriptions to which the students' talk could directly refer, that is, the object of their talk was immediately available for indexing.

Data Set 1.b. shows the range of investigations conducted by CJP over a total of 14 weeks during the fall semesters of 1990 (experiments 1-4) and 1991 (experiments 5-7). From a physics perspective, the investigations are all structurally connected in that they investigate motion in a gravitational field and affected by friction generated by varying parameters. Among the parameters CJP changed were the shape of the falling objects, the size of a friction-generating sail, the friction-generating medium (air, various liquids), and the type of motion (straight fall, oscillations of yo-yo or spring). Such evidence indicated to me that students, when free to design their own investigations, begin to establish research agendas which investigate specific phenomena in quite some detail. Students' investigations become more refined because of their developing competence in the use of data analytic methods and statistical software, conceptual understanding, and the increasing familiarity within the topic of their own interest.

Data Set 1.c. illustrates students' responses and attitudes towards open-inquiry. They valued the freedom and motivation that comes with formulating their own problems and designing experiments to find solutions. Because they can choose problems of their own interest, they bring prior experience, interest, and out-of-school experience to the classroom which helps them to connect their experiences. In this way, the problems they framed, the solutions they came up with, and what they learned in terms of knowing and doing science was meaningfully connected to their other experiences ("This lab taught us the value of design in things like air planes and race cars, even the aerodynamics of our own cars"). Third, the data show that students perceived their investigations as self-supporting in that they generate more questions to be answered ("It brought up more questions concerning form, and shape, and the effect of density of liquids").

As most of their peers, CJP wanted to do experiments "where [they] knew that it works." They anticipated specific relationships. Even if I had known that the students would not observe the expected pattern, CJP was a group that seemed to handle well those "experiments where they could not observe their effect." I had termed these experiments *blind alleys*. The number of *blind alleys* was in part determined by my decision to prevent students from becoming too frustrated, and to allow them to adjust slowly to uncertain outcomes (which are a hallmark of original research). In other circumstances, the decision to prevent a *blind alley* was determined by the hazards involved, e.g., when CJP considered liquefying a salt and then preparing a solution with water. Sometimes, we (the teaching assistant and I) engaged with students in the pursuit of a *blind alley* because we too did not know the outcome. In one such example, a group decided to measure the heat of fusion of water by immersing a small amount of it in an alcohol bath of which they monitored the temperature in a continuous fashion; in another case, two students attempted to build a gas chromatograph by continuously monitoring the temperature of the combustion of the carrier gas with and without test materials. In spite of many variations in the experimental parameters and consultations with various expert sources (such as chemistry teacher, university library), the expected outcomes were not observed. In these cases, I assisted students in viewing their experiments from a positive perspective rather than as failures. I encouraged them to (a) regard their ideas as reasonable within their current understandings, (b) use the experiment as evidence for procedures which do not work, and (c) to report these negative findings.[6]

---

6 In my physics courses, students were never penalized for their data. If they came to conclusions which differed from that of other students and/or currently accepted theory, I asked them to discuss possible reasons for these discrepancies. Any other approach seemed counterproductive, supporting the development of academic dishonesty. In one instance, two very gifted students with scholarships to major US and Canadian universities, fudged their chemistry data by orders of magnitude in order to get the "right" answers for a teacher who emphasized the accuracy of measurement and penalized students for not getting close.

On the basis of these and similar data, I discerned several dimensions in the problem setting of students. First, students began their research with problems framed in more general (less specific) form. Second, students continued framing problems which focused on the same phenomena, but with increasing detail as their familiarity with the phenomena increased. Third, because of the continued focus on the same phenomena, they developed entire research programs. Fourth, because students framed problems themselves, these were motivating (at the appropriate level for the students current understanding), meaningful, and connected to students prior and out-of-school experience. Because of these elements, students could take ownership of both problems and solutions. Fifth, problems do not exist *a priori* but arise out of the interaction of individuals acting in social and physical settings. Just what a problem is depends on the relationship of individuals or groups with each other and with the physical environment. All of these features of students' problem-framing activity parallels work in scientific laboratories and everyday mathematical problem solving as corroborated by recent biographical, anthropological, and sociological reports (Knorr-Cetina, 1981; Lave, 1988; Suzuki, 1989). The quality of students' inquiries evolved from the locally and temporally situated group decisions and, I found, could not be predicted by any set of variables such as general achievement in physics, mathematics, or overall GPA. As the students proceeded, they constructed hosts of contingent problems which they resolved as a matter of course in the process of achieving a solution to their over all, global problem. From the interactional work with their peers and the setting, solutions to contingent problems emerged oftentimes spontaneously.

Framing problems is an important skill in everyday environment where problems are generally undefined or ill-defined; in contrast, the textbook problems that students encounter in schools are relatively well-defined and of extremely limited context. For example, what is a best-buy in a supermarket is not simply a question of buying the product with the smallest unit price (a ratio problem) but depends on a large number of considerations such as available package sizes, age of the product, shelf-life, size of the family, visitors, relative costs of small and large packages, etc. Thus, educators have called for explorations through the pursuit of questions in a stable domain as a key ingredient of teaching, implying that students are the originator of essential questions (Collins et al., 1989; Wiggins, 1989). In the pursuit of their own questions, students learn not only to gain pleasure from inquiry but they also feel in control over problems and solutions ("We found this lab interesting because we were pursuing something of our own interest"). Such ownership implies, as Lave (1988) indicated that "individuals experience themselves as in control of their activities, interacting with the setting, generating problems in relation with the setting and controlling problem-solving processes" (p. 69). As the present data illustrated, the students in my studies experienced such ownership and control. In traditional school settings on the other hand, students often experience themselves as objects with no control over problems and solutions. *Blind alleys* were an integral part of this process, but are not common in most high school and university science teaching. In part, this is due to the view of "efficient" science teaching as training in getting facts and procedures right.

Processes during Solution Finding

When students do textbook word problems, they are mainly concerned with getting the right answer, finishing their work and pleasing the teacher (Wheatley, 1991). When students frame their own problems, however, solution finding takes on different aspects. The choices students make affect the data they collect and the conclusions they draw from the experiment. In this episode (Table 2), CJP decided to measure the viscosity of several liquids for their hydrodynamics experiment.

CJP dealt with the problem of how to measure the viscosity of liquids. In the course of their discussion, the three actually constituted the problem as a composite of a number of problems: Should viscosity be determined relative to some substance?, Which substance should they take?, and Should the volume of liquid to flow through their instrument be measured or merely kept constant (without measuring it in absolute terms)? In lines (1-4), Pete and Carl framed the problem, how to measure viscosity and whether air was an appropriate standard for comparison (they had previously done a free-fall experiment in air, see Data 1.b.). Carl thought that they could not measure the viscosity of air, implying that it was not a useful standard (4). However, Jim proposed a method that would allow them to use air as a standard (vignette in line 5). Although they abandoned the idea of using air as comparison, they kept Jim's suggestion of using a funnel for determining viscosity. Their subsequent discussion no longer questioned the use of the funnel, but how to refine their method, and whether to make absolute or relative measurements. Jim asked whether they should use the rate at which liquid volumes (milliliter/minute) flowed through their device (6); but Carl suggested to avoid the absolute volume measurement by marking a specific level in the funnel and use the same amount of all liquids leaving them with time as a measure of viscosity (7-8). At this point, I entered the discussion and suggested to use a burette allowing for both methods (9). Jim responded that using a burette would increase the time for determining the viscosities (10); but I suggested that it would help them in determining it more accurately (both the flow rate and the liquid volume).

In his explanation for measuring relative viscosity, Jim did not merely provide a verbal description, but accompanied his talk with concrete manipulations of the apparatus. He indicated how to hold the funnel while filling it with water and how to close the narrow end with his

Table 2. Students' approaches to problems of their own framing.

AUDIO	VIDEO
1. P: How do you measure viscosity?	
2. C: This has to be relative.	
3. P: Viscosity relative to air?	
4. C: Air we can't test.	
5. Jim then makes a suggestion for measuring viscosity of air: Fill the funnel with water, put paper on top of surface. Turn the funnel upside	(Manipulates a funnel)
down and let the liquid fall out. The air has to come in through the end. Take the time it takes the liquid to fall out and you have the time it takes the air to fill the funnel through the nozzle.	
6. J: Should we measure the viscosity in milliliter/minute?	
7. C: Just fill it up to a certain, the same for each liquid.	(indicates how high to fill funnel)
8. Then set it over the sink, or whatever [...] measure the time it takes to fall out.	(holds funnel over a graduated cylinder in lieu of a sink; shows finger on bottom of funnel)
9. T: Or we could just take burettes and just time a certain amount, how long it takes.	(comes in from background)
10. J: But this one will be faster. 11. T: But it will be more exact.	

index finger; and he gestured with the palm of his hand the placement of a sheet of paper over the wide opening, followed by the turning of the instrument and the opening of the narrow end to allow the water to flow (5). In (7), Jim gestured the level up to which he suggested to fill the funnel. He accompanied his instructions with detailed gestures how to hold the funnel, empty it over the sink (for which he used a graduated cylinder as a place holder), and how to release the flow of water.

This episode illustrates the framing of a problem (measuring viscosity), and the subsequent choices for resolving it. We can see that a specific choice along the way (the funnel), moved the students' investigation in a certain direction. After Jim's suggestion for using a funnel (5), the group no longer considered other options but worked to refine their funnel method. Thus, their problem frame (viscosity as a flow rate) had entailed a specific solution, i.e., viscosity as determined by the flow rate through a funnel. Such choices determined the investigations and ultimately the learning trajectory in ways the traveler's choices determined his life in Robert Frost's poem, *The Road not Taken*. In this way, some problems and solutions were more likely than others and some content learning was achieved over others, independently of confined textbook presentations of "motion." As I showed elsewhere (Roth, 1994), when students shared their problems and solutions with their peers, a wide range of content was "covered" that went beyond, and was far more integrated, than any textbook and teacher-centered presentation of content.[7] Of course, in some instances these choices led to *blind alleys*, with their own corresponding pedagogical problems.

This episode illustrates another important aspect in students' problem-framing and solution-finding processes: thinking as embodied action on concrete materials. In addition to, and often hand in hand with the description of phenomena, students used concrete materials, instruments in whole or part, and drawings to think about and to communicate their ideas. These concrete materials were also used for ad hoc simulations which helped students to formulate their ideas. Out of these manipulations arose the descriptions of experiments and the formulation of focus questions. Each student group reasoned by referring to the actual objects or events. In a sense, these objects became mediators of meaning within the group. For example, for a proposal to observe the forces acting during deceleration, students used a spring scale to describe the event and how to measure the forces; the spring scale was central to the discussion of this idea and helped to overcome those moments when they did not understand each other ("You lost me there!"). Or, while discussing the instrumentation for their experiment on the thermal expansion of liquids, students secured pieces of equipment such as a test tube, glass tube, or a temperature probe; these facilitated their planning effort during which they talked about such details as the placement of the temperature probe for accurate readings or the placement of the "amplifier" glass tube. Thus, the concrete materials provided students with something that they could manipulate, point to, or talk about. As such, these materials did not only facilitate

[7] The process by which this occurs has little to do with transmission. Rather, it seems to be a characteristic of learning communities (I highlighted earlier the communal aspects in my classrooms) an example of which we have described in some detail elsewhere (Roth & Bowen, in press).

planning, but also became mediators of meaning within the groups. In all of my studies, the concrete conceptual objects focused the discourse of the participants. Concrete objects created a shared interactional space which facilitated the negotiation of meaning and lead to the point that concepts and ideas could be taken as shared. These objects also allowed participants to run various scenarios which then could be examined for their structure. Thus, when students discussed an experiment with a spring scale they presented, criticized, and modified their ideas about this experiment by running different measurement scenarios accompanied by narrative explanations. Facilitated by the concrete objects and the processes, the students constituted the shared understanding that the experiment was not viable and should be abandoned.

This episode (Table 2) also illustrates that the students were not simply concerned with the collection of one set of data which they then interpreted to make it fit some theory. Rather, the students began to run their experiments and compared the results to their expectations. In the case of discrepancies, they began to construct "errors," that is, reasons for the discrepancies. Through iterations of re-designing experiments, "improving" specific aspects of the phenomena to be observed, changing the data recording mechanisms, and reformulating their models of the phenomena, they attempted "to change" the data (quality) in specific ways. These changes were always the result of experimental manipulations rather than *a posteriori* changes of the data. Because there are few high school level treatments of complex real-life motion phenomena which include friction, buoyancy, turbulence, and so forth, the students did not feel the need to make their data fit to an *a priori* theory. Rather, they engaged in a description of the observed objects and events, and attempted to describe relationships with mathematical functions by using available software.[8]

The student behaviors I observed were markedly different from problem solving in traditional labs. When the students do "cookbook" laboratories, they are less concerned with meaning as they go along because they can expect to be able to fit data and theory once they are done, even if this means that they have to fit the data in a Procrustean

---

8 Some of their data could easily be fitted with a polynomial of first or second degree. In other cases, I suggested possible transformations of the data before attempting a fit. Finally, using a mathematical modeling program, I prepared a least square algorithm which allowed students to fit any function to their data by adjusting the function's parameters on the basis of visual inspection and the improvement of the goodness-of-fit index, R^2 (Roth, 1992, 1993).

manner;[9] students are more concerned with completing the step-by-step procedure and with finishing the exercise (Gallagher & Tobin, 1987). In the present open-inquiry learning environment, the students' problem solving was driven by epistemic rather than practical concerns. They knew what they wanted to measure; they knew what data to collect; and in many instances, they had specific hunches of what to expect. Thus, when the data did not confirm their expectations, they constructed possible "errors" which led them to change parts of their set-up, data collection modes, or models for the observed phenomena. In other words, "errors" were problems of some sort that led to discrepancies between expectations and observed data or resulting interpretation. In the end, the resulting graphs and interpretations were interactively constructed achievements in which both the social and physical settings bore on the ultimate solutions.

Framing their own problems and developing corresponding solutions left ownership over problems and solutions with students. The students were aware of the difference between problem solving in open-inquiry and traditional laboratory exercises (which most of them conducted in a concurrent chemistry course). My students did not like the step-by-step approach of traditional verification laboratory exercises because the purpose of most steps remained hidden from them and they often fudged their data in the chemistry labs to conform with the expected values. Problem framing and solution finding in the open-inquiry physics laboratory, on the other hand, provided them with new opportunities. They knew each step and why they were taking it; they saw problems as challenges to be resolved with locally available (physical and social) resources; and in some cases they even decided to abandon the line of inquiry they had taken. Such approaches are the hallmarks of effective everyday problem solving and show much cognitive flexibility in the framing, transformation, resolution, and abandonment of problems and solutions. Like Lave (1988), I found that this flexibility allows individuals to take control of their activities, interact with the setting, generate problems in relation to the setting, and to control the problem-solving process. On the other hand, word problems given in school and psychological experiments create contexts in which participants experience themselves as objects, with little control over problem frame or choice of problem-solving process. If school problem solving is to be a

---

[9] The mythical inn keeper Procrustes forced all his guests to fit his beds; using his axe, he shortened those that were too long; those that were too short, he stretched until they measured up to the beds' size. I have observed many students in high school and college to "work" their numbers in similar ways to fit the relevant theory. Even scientists have been reported to "work their numbers." When this becomes public, it usually leads to the disgrace of those concerned. I believe that science laboratories which emphasize the "right" answers/measurements contribute to the phenomenon of "working" the numbers.

preparation for out-of-school situations, then we need to prepare contexts in which they can develop knowledge-in-practice and the embodied skills for dealing with the complexity of real-life problem solving rather than that of well-defined, puzzle-like, and decontextualized riddles.

Information Processing and Situated Cognition

The illustrations of students' problem framing and solution finding hint at the problems an information-processing approach will have in modeling students' activities. While the space limitations of this chapter does not allow for a full discussion of the issues, the following comments may suffice to point out some of the difficult issues. These comments are not designed to debunk traditional cognitive-science perspectives (which are valid for problem solving in confined and well-structured domains) but to bring to the level of awareness those aspects of problem solving in open-inquiry environments which cannot be modeled by the traditional approach.

First, traditional cognitive-science research on problem solving works in well-defined domains allowing for a tight control of variables, prerequisite knowledge, information provided, or possible solution paths but not for ill-structured real-world problems. The problems are taken as given and solvers' activities are judged against some standard solution. My students worked in laboratory settings in which problems were not preframed, but were one of the students' accomplishments. That is, the information which is fixed and predetermined in laboratory studies, is open to interpretation by the participants; what counts as important information and is foreground, and what counts as background noise and remains unspecified is elaborated differently by each and every group as a function of the present state of the setting. Given this variability, standard solutions equivalent to those on word problems cannot be specified.

Second, while the four representations (words and sentences, objects and relations, scientific objects and drawings, algebraic symbols) identified by traditional cognitive science are still relevant, they are so in a different way. In the situated cognition view that accounts for the present and similar data, there are no longer unproblematic one-to-one, *a priori* correspondences between representations but mutually constitutive, socially constructed, and shared conventions which associate dissimilar semiotic elements (phenomenon, different representations).

Third, working on the edge of their knowledge, the fundamental problem-solving processes of expert scientists and students are more similar than dissimilar. This arises from the fact that scientists, if they could be modeled by production systems, would use rules more similar to those characteristic of the students. However, in addition to my own studies on students, there is ample evidence from ethnographic studies of scientists' cognition showing that important aspects of their knowledge is tacit. Furthermore, much of my students' or scientists' commonsense knowledge which allows them to function in this world is never made explicit. By definition, this knowledge cannot be modeled with production systems.

Fourth, my illustrations show that important aspects of students' activity occur in the interplay of social, physical, historical, and individual cognitive factors. The interplay of these factors is so complex that any aspect of students' work, processes and products cannot be understood unless the entire system is considered. Here, the setting of individual activity can no longer be considered stable, but as continuously changing. Traditional theories of cognition do not model these situations, even in its newer versions of connectivism, because they do not account for the extra-individual aspects of problem framing and solution finding.

Fifth, the recent focus on knowing as mutually constitutive interplay of discourse and world has illustrated that information is not out there and uninterpreted. Rather, perception (visual, aural, or otherwise) is always interpreted in the light of an individuals prior experience. From this view, information cannot be taken as an unchangeable given but is always relative to the individual and the setting. The information-processing perspective does not model this aspect of my students' problem-solving activities.

Conclusions and Implications

In this chapter, I provided examples from my research on problem solving in physics. These results lead to a different view of problem solving, a problem solving which is much more akin to that in which everyday folks and scientists engage. Because of the difference with traditional views of problem solving in junior high and high schools, there are some important implications for teaching and evaluating science.

New Principles of Problem Solving in Physics

The findings reported here illustrate a substantial difference between problems set by textbooks and teachers and those set by students themselves. When problems are not set by students, their main problem is the search of hidden solutions already known to teachers; students attempt to get these right answers rather than doing problems out of an epistemic interest. On the other hand, when the students frame their own problems, solution processes and products are often entailed in the problems. In this case, students own problems and solutions (both as processes and products). That is, students design the problems and, much like scientists, construct solutions and the criteria for accepting them as appropriate, even if this means that they abandon specific solutions or problem frames.

The present descriptions also illustrate that problems do not exist independently of the individual; that is, the ontology of a problem is a function of individuals, their experience and their (social and physical) settings. First, textbook word "problems" are not problematic for physicists and competent physics teachers who can often provide answers without reflecting or calculating. On the other hand, these "problems" are often insurmountable for beginning physics students; that is, students do not provide the same answer as physicists, and are thus judged deficient. Second, the motion of a yo-yo is a problem if all one has available is a stopwatch or a ticker timer, but is a much smaller problem if one has a sonar-based motion detector and a force probe. Third, just what is a problem often depends on viewpoints different students take on a certain phenomenon. Through students' negotiations, local and global problems emerge from their interactions in physical and social settings. Thus, "problem" indicates the relationship between an individual or group of individuals and some situation (e.g., "What is the answer to the question in the word problem?" or "What is the best buy for some item on my grocery list?"); it is not a property of this situation.

What emerges from this and similar research is a view of problem solving in ill-structured settings which includes three key elements of the new cognitive-science perspective. Knowledge is (1) not a factual commodity or collection of facts, but (2) has a process character expressed in the dialectic of individuals acting in settings; and (3), this dialectic of dilemma-resolution, setting, and activity suggests the inseparability of problem framing and the means and ends of problem solving. The presented view constitutes a radical departure from traditional normative models of problem solving with their unexamined presuppositions of rationality, value-free problem-solving processes, the nature of cognitive functions, or means/end relations. Nevertheless, this view is in agreement with current research on knowing and learning in everyday settings and scientific laboratories.

<u>Students and Scientists</u>

Throughout this chapter, I have highlighted the similarities between students' and scientists' problem solving. However, I do not want to hide the fact that there are also poignant differences between them. First, there are considerable differences in the relative conceptual backgrounds which they bring to their work. Scientists, having worked for many years in their domain, have highly elaborated frameworks for interpreting their experience. Second, there are differences in the embodied laboratory practices, the "vulgar" competencies and familiar efficiency which scientists exhibit on the work bench. These differences arise from differential opportunities for developing experiential meanings in terms of laboratory practice which are so important to the successful rendering of preconceived solution (six months in one course versus years in research laboratories); students have had little time developing canonical physicists' discourse and laboratory practices; and students have had little practice in practical, situated problem solving. Finally, because of my presence as guide, facilitator, and resource,[10] the students' activities were more comparable to physicists-in-training (Ph.D. students, post-docs).

Implications

The present findings are compatible with current research on situated cognition and authentic learning experiences (Brown et al., 1989; Lave, 1988) and have important implications for teaching physics problem solving in high schools (Marton, 1993; Ueno, 1993). If students are to do physics as physicists do, that is, research interesting phenomena of their own interest and choice rather than finding answers to word problems, high school physics has to change. In the following paragraphs, I will point out important changes to the teaching-learning environment and to the evaluation of problem solving.

Teaching-learning environment. To follow the recommendations for problem solving in ill-structured settings requires classrooms in which students can pursue interesting physics problems of their own interest. While it is nearly impossible to give prescriptions that can be turned unequivocally into desired learning environments, the description of my classroom provided earlier illustrates the way I made it work (The interested reader will find extensive illustrations of how we made open-inquiry science work in: Roth, 1991b, in press; Roth & Bowen, 1993, in press). The illustrations provided here were derived from physics classrooms where students had the opportunity to design their own research projects and construct their own solutions. I have outlined some facets of the resulting problem framing and problem solving. It is clear that such classrooms differ from traditional ones. First, as teachers, we no longer have total control of what is happening in our classrooms. Because of our responsibility towards schools, parents, and tax payers, we have to make sure that learning is happening and that we can show it. In my experience in public and private schools (at times under extremely difficult situations), giving students responsibility over their own learning and relinquishing tight control of classrooms and students actually leads to increased rather than decreased learning, to improved student attitudes, and to classroom learning characterized by epistemic interests. In such classrooms, students not only learn the products and processes

10 For one year, a trained physicist worked with me as a laboratory and teaching assistant and served as an additional point of reference, resource, and authentic model for students.

of science, but they also experience the social nature of scientific knowledge and knowledge construction.

Evaluation of student work. When classroom environments change to focus on physics as process, evaluation procedures have to change to take account of the different nature of problem solving. We cannot expect that students, albeit great efforts, will necessarily arrive at the claims of canonical science. With the shift from product to process teaching, we have to change evaluation of knowledge production as it is practiced in science (Latour, 1987; Scardamalia & Bereiter, 1994; Ueno, 1993); rather than studying science's products (ready-made science), we need to study science as a process (science-in-the-making). Process-oriented evaluation focuses on (a) the contributions of individuals to the inquiry, (b) the viability of claims in terms of the data which students collected, (c) the creativity of research questions, and (d) on the skills of constructing problems and their corresponding solutions.

Besides focusing on the process aspects of problem framing and solution finding, evaluation has to account for the (socially and physically) distributed nature of problem solving. That is, in accounts of thinking during problem framing and solution finding, teachers and problem-solving researchers have to account for, and build into their theories, the situated nature of knowing and learning, and the dialectic of goals, activity, and problem-solving processes. My teacher colleagues who took the time to (record and) observe pupils' problem solving on videotape, were amazed with the amount (both in content and process) their students learned above and beyond what they made available in written reports, class discussions, and formal tests and examinations. In the context of most teachers' daily life, detailed analyses of each student's problem solving are prohibitive in terms of time and effort involved. It is noteworthy, however, that these colleagues changed their views about evaluation and began to include significant process components.

Future research. I envision two major areas that need to be clarified by future research in the area of problem solving in physics. First, future research needs to construct a better understanding of the situated nature of intelligent human agency in ill-structured contexts: How does the social context affect problem-solving activities?, What are the relative contributions of tools and materials in problem framing and solution finding?, and How does the classroom community frame, limit, or enhance the activities of individuals and small groups? I would expect that such research will result in models (e.g. graphical) that strike a balance between descriptive complexity and representational simplicity. Second, future research needs to investigate the changes in problem-solving competence as individuals gain experience in dealing with ill-structured settings. The research question in most general terms would then be, What are the domain-specific and domain-independent knowledge and skills acquired in the process of becoming a physicist? and What are the most beneficial settings in acquiring the knowledge and skills of a physicist? I would expect that the answers to these questions will lead to the design of better learning environments that encourage students to develop widely useable and marketable skills.

Acknowledgments

I thank Michael Bowen, Todd Alexander, Guennadi Verechaka, and all my physics students taught during a three-year span for countless hours of conversations and assistance through which I developed the understanding presented in this chapter. I extend my thanks to Sylvie Boutonné for her editorial help in preparing this chapter.

References

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Champagne, A. B., Klopfer, L.E., & Gunstone, R.F. (1982). Cognitive research and the design of science instruction. *Educational Psychologist, 17*, 31-53.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1980). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-151.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*, 66-71.

Collins, H. M. (1982). Tacit knowledge and scientific networks. In B. Barnes & D. Edge (Eds.), *Science in context: Readings in the sociology of science* (pp. 44-64). Cambridge, MA: MIT Press.

Collins, A., Brown, J. S., & Newman, S. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. Resnick (Ed.), *Cognition and instruction: Issues and agendas* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.

diSessa, A. (1993). Responses. *Cognition and Instruction, 10*, 261-280.

Drake, S. (1990). *Galileo: Pioneer scientist.* Toronto: University of Toronto Press.

Ferguson-Hessler, M. G. M., & de Jong, T. (1987). On the quality of knowledge in the field of electricity and magnetism. *American Journal of Physics, 55*, 492-497.

Forsythe, D. E. (1993). The construction of work in artificial intelligence. *Science, Technology, & Human Values, 18*, 460-479.

Gallagher, J.J. & Tobin, K. (1987). Teacher management and student engagement in high school science. *Science Education, 71*, 535-555.

Garfinkel, H., Lynch, M., & Livingston, E. (1981). The work of a discovering science construed with materials from the optically discovered pulsar. *Philosophy of the Social Sciences, 11*, 131-158.

Gooding, D. (1992). Putting agency back into experiment. In A. Pickering (Ed.), *Science as practice and culture* (pp. 65-112). Chicago, IL: The University of Chicago Press.

Heller, J. I., & Reif, F. (1984). Prescribing effective human problem-solving processes: Problem description in physics. *Cognition and Instruction, 1*, 177-216.

Hutchins, E. (1991). The social organization of distributed cognition. In L. Resnick & J. Levine (Eds.), *Perspectives on socially shared cognition* (pp. 283-307). Washington, DC: APA Press.

Knorr-Cetina, K. D. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science.* Oxford: Pergamon Press.

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

Larkin, J. H., & Rainard, B. (1984). A research method for studying how people think. *Journal of Research in Science Teaching, 21*, 235-254.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society.* Milton Keynes: Open University Press.

Latour, B. (1992). *Aramis ou l'amour des techniques* [Aramis or the love of technique]. Paris: Éditions la Découverte.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life.* Cambridge: Cambridge University Press.

Lave, J. (1993). The practice of learning. In S. Chaiklin & J. Lave (Eds.), *Understanding practice: Perspectives on activity and context* (pp. 3-32). Cambridge: Cambridge University Press.

Lestel, D. (1989). L'anthropologie des laboratoires et la pratique d'intelligence artificielle [The anthropology of laboratories and the practice of artificial intelligence]. *Social Science Information, 28*, 685-703.

March, J. G. (1991). How decisions happen in organizations. *Human-Computer Interaction, 6*, 95-117.

Marton, F. (1993). Our experience of the physical world. *Cognition and Instruction, 10*, 105-225.

Pea, R. D. (1988). *Distributed intelligence in learning and reasoning processes.* Paper presented at the meeting of the Cognitive Science Society, Montreal.

Pickering, A. (1985). *Constructing quarks: A sociological history of particle physics.* Chicago: Chicago University Press.

Pickering, A., & Stephanides, A. (1992). Constructing quaternions: On the analysis of conceptual practice. In A. Pickering (Ed.), *Science as practice and culture* (pp. 139-167). Chicago, IL: The University of Chicago Press.

Pinch, T. (1985). Theory-testing in science—The case of solar neutrinos: Do crucial experiments test theories or theorists? *Philosophy of the Social Sciences, 15*, 167-187.

Resnick, L., & Levine, J. (Eds.). (1991). *Perspectives on socially shared cognition.* Washington, DC: APA Press.

Roth, W. -M. (1990a). Short-term memory and problem solving in physical science. *School Science and Mathematics, 90*, 271-282.

Roth, W. -M. (1990b). Neo-Piagetian predictors of achievement in physical science. *Journal of Research in Science Teaching, 27*, 509-521.

Roth, W. -M. (1991a). The development of reasoning on the balance beam. *Journal of Research in Science Teaching, 28*, 631-645.

Roth, W. -M. (1991b). Open-ended inquiry: How to beat the cookbook blahs. *The Science Teacher, 58*, 40-47.

Roth, W. -M. (1992). Bridging the gap between school and real life: Toward an integration of science, mathematics, and technology in the context of authentic practice. *School Science and Mathematics, 92*, 307-317.

Roth, W.-M. (1993). Problem-centered learning or the integration of mathematics and science in a constructivist laboratory: A case study. *School Science and Mathematics, 93*, 113-122.

Roth, W. -M. (1994). Experimenting in a constructivist high school physics laboratory. *Journal of Research in Science Teaching, 31*, 197-223.

Roth, W.-M. (in press). *Authentic school science*. Dordrecht, Netherlands: Kluwer Academic Publishing.

Roth, W. -M., & Bowen, G. M. (1993). An investigation of problem solving in the context of a Grade 8 open-inquiry science program. *Journal for the Learning Sciences, 3*, 165-204.

Roth, W. -M., & Bowen, G. M. (in press). Knowing and interacting: A study of culture, practices, and resources in a Grade 8 open-inquiry science classroom guided by a cognitive apprenticeship metaphor. *Cognition and Instruction.*

Roth, W. -M., Gabel, D., Brown, L., & Rice, D. (1992). A combined method of INDSCAL and cognitive mapping for describing changes in students' cognitive structure. *Educational and Psychological Measurement, 52*, 769-779.

Roth, W. -M., & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching, 30*, 127-152.

Roychoudhury, A., & Roth, W. -M. (in press). Interactions in an open-inquiry physics laboratory. *International Journal of Science Education.*

Scardamalia, M., & Bereiter, C. (1994). Computer support for knowledge-building communities. *The Journal of the Learning Sciences, 3*, 265-283.

Shavelson, R. J. (1974). Some aspects of the correspondence between context structure and cognitive structure in physics instruction. *Journal of Educational Psychology, 63*, 225-234.

Starling, G. (1992). Problem solving as a social process: A theoretical and empirical analysis. *The Social Science Journal, 29,* 211-225.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication.* Cambridge: Cambridge University Press.

Suchman, L. A., & Trigg, R. H. (1993). Artificial intelligence as craftwork. In S. Chaiklin & J. Lave (Eds.), *Understanding practice: Perspectives on activity and context* (pp. 144-178). Cambridge: Cambridge University Press.

Suzuki, D. (1989). *Inventing the future: Reflections on science, technology, and nature.* Toronto: Stoddart.

Thro, M. P. (1978). Relationship between associative and content structure of physics concepts. *Journal of Educational Psychology, 70,* 971-978.

Traweek, S. (1988). *Beamtimes and lifetimes: The world of high energy physicists.* Cambridge, MA: MIT Press.

Turkle, S., & Papert, S. (1991). Epistemological pluralism and the revaluation of the concrete. In I. Harel & S. Papert (Eds.), *Constructionism: Research reports and essays, 1985 - 1990* (pp. 161-191). Norwood, NJ: Ablex.

Ueno, N. (1993). Reconsidering p-prims theory from the viewpoint of situated cognition. *Cognition and Instruction, 10,* 239-248.

Ueno, N., & Arimoto, N. (1993). Learning physics by expanding the metacontext of phenomena. *The Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 15,* 53-63.

Varela, F. J., Thompson, E., & Rosch, E. (1993). *The embodied mind: Cognitive science and human experience.* Cambridge, MA: MIT Press.

Wheatley, G. H. (1991). Constructivist perspectives on science and mathematics learning. *Science Education, 75,* 9-21.

Wiggins, G. (1989, November). The futility of trying to teach everything of importance. *Educational Leadership, 46*(9), 44-48, 57-59.

Woolgar, S. (1990). Time and documents in researcher interaction: Some ways of making out what is happening in experimental science. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 123-152). Cambridge, MA: The MIT Press.

Assessment of Problem Solving in Science

Rodney L. Doran
Graduate School of Education
University at Buffalo
Buffalo, NY 14260

Stanley L. Helgeson
National Center for Science Teaching and Learning
The Ohio State University
Columbus, OH 43210

David D. Kumar
College of Education
Florida Atlantic University
Davie, FL 33314

Introduction

Problem solving is viewed by some as a gap between where the learner is and where he or she wants to be (Hayes, 1981). Others see problem solving as the application of previously learned rules, but with new learning included in the process (Gagne, 1977). Problem solving also has been operationally defined to include the integrated science process skills of interpreting data, controlling variables, defining operationally, and formulating hypotheses (Shaw, 1983). Smith (1991) considers a problem any task that requires analysis and reasoning toward a goal, thus problem solving is performing the required analysis and reasoning for a solution. He notes further that, "This analysis and reasoning must be based on an understanding of the domain from which the task is drawn. A problem cannot be solved by recall, recognition, or reproduction..." (p. 8). Smith includes both declarative knowledge and procedural knowledge as necessary to the problem-solving process (cf. Hauslein and Smith; Lavoie this Monograph).

Brandwein discussed scientists' role in problem solving and included a schematic as a description of the "scientist's method of intelligence." Brandwein wished, "we did not require a schema of the scientific art of investigation; being an art, it defies description: It is idiosyncratic: it is creative" (1971, p. 19).

Brandwein (1971) stated that this

> brief, oversimplified scheme emphasizes the poverty of <u>steps</u> in a 'scientific method;' the essence of the scientific way (call it inquiry if you must) is an art which calls upon the bright luminosity of ... imagination as well

Chapter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for scientific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

> as the sober processes in the library, laboratory, and the conference room. But the sea of science is essentially an individual's mind, not the laboratory. (p. 22)

In addition to the artistic, individualistic portrayal of "doing science" by Brandwein, one is struck by the cyclical, interactive nature of his schema. The methods of science are clearly much more than a series of steps, however dear they are to textbooks. As a matter of fact, there is no "end" to the Brandwein schema. This is quite different from the right answer, verification, and algorithmic approaches to science in many instructional programs.

The Assessment of Performance Unit (APU) "developed a problem-solving model to help make sense of what pupils do when they investigate" (Strang et al., 1991, p. 7). The APU model has a cyclical structure like Brandwein's. The British researchers found this to be a useful framework against which to compare pupil performance. In addition it is helpful for analyzing the cognitive demands of different investigations.

Solomon (1988) cites several perspectives to problem-solving:

> Everyone, it seems, refers to "problem-solving as the aim of the teaching experiment. Some may mean designing a test for the absorbency of paper towels, others mean 'finding out what happens if . . .," some mean taking part in "egg-race" competitions, and still others mean testing out pupils' own theories. (p. 103)

Modes of assessment, not surprisingly, are as varied as the definitions of problem solving. Unfortunately, not all attempts at assessment can be regarded as successful. Shavelson, Carey, and Webb (1990) note that "the current technology produces tests that emphasize recall of facts and performance of isolated skills but tend not to measure students' conceptual understanding and problem-solving skills" (p. 697). Reviewing methods of assessment, Meng and Doran (1990) note that written tests, projects and written work, observations, discussions, and interviews are all useful methods. However, practical tests are the most appropriate when attempting to assess learners' ability to apply their knowledge. Such application, in a laboratory-based problem-solving setting, is the basis for many alternative or performance approaches to assessment in science. A major consideration in performance assessment is the selection of a model with tasks appropriate for the students and for the domain of knowledge being assessed (Doran, Boorman, Chan, and Hejaily, 1992). In England, the APU assessment procedures included paper and pencil assessment, but a major contribution was the implementation of practical testing on a national scale (Johnson, 1989). As educational technology improves, computer applications increasingly add to available assessment techniques. The possibilities range from access to large test banks (Aesche and Parslow, 1988) to computer adaptive testing (Herb, 1992), to the use of Hypermedia (Kumar, in press), and to the use of interactive videodisc (Lomask, Jacobson and Hafner, 1992).

Assessment approaches have also included consideration of other bases of problem solving. The reasoning required in problem solving appears to involve some of the same skills as those required for the integrated science process skills. Attempts to assess these skills led to the development of the Test of Integrated Process Skills (TIPS) (Dillashaw and Okey, 1980) and TIPS II (Burns, Okey and Wise, 1985) along with derivatives of these instruments. In a somewhat similar vein, the skills involved in formal reasoning and logical thinking also appear to bear on problem-solving ability. Concern for assessing logical thinking abilities led to the Test of Logical Thinking (TOLT) (Tobin and Capie, 1981) and the Group Assessment of Logical Thinking (GALT) (Roadrangka, Yeany and Padilla, 1983).

Jung recognized the difficulties in defining the field of cognitive science, then stated "obviously, cognitive science is an interdisciplinary field, centering around a common subject, which is cognition, or more generally, the 'nature of mind," e.g., an astoundingly old-fashioned description of a scientific subject" (Jung, 1993, p. 31). Jung described the work of science educators in the cognitive science field as concentrating on representations of knowledge and problem-solving skills. He cited some of the conceptual tools available for describing and analyzing knowledge as "schemas, frames, scripts, networks, lists, production systems, analogical and symbolic representations (Jung, 1993, p. 48). We found these ideas useful as we described attempts to assess youngster's ability to solve problems in science contexts. The general problem-solving skills, as well as specific content knowledge are critical elements for our assessment framework. While some assessment instruments have focused on a comprehensive approach, most are dwelling on specific skills or abilities.

Assessment of Structural Knowledge

Students' conceptual knowledge of the science domain related to a problem has a bearing on their approach to and success in solving the problem. Stewart (1980) describes five paper-and-pencil techniques for assessing structural (declarative) knowledge. In a concept-map line labeling task students are given a set of gummed labels each with a single concept name printed on it. The students are asked to arrange the labels on a piece of paper, draw lines between related concepts, and describe the relationship that each line represents. In the tree construction line labeling task, students are presented with a set of concept labels and asked to pick the two terms that they think are most similar to each other, write the terms in the middle of the page, and then connect them with a line. Then, the students either select a new term they believe to be most similar to one of the terms already on the page and connect it to the appropriate term, thus beginning a concept tree, or they select a new pair of similar terms and start a new tree. When all the concepts have been connected, the students describe the relationships represented by each of the lines drawn between two terms. In the concept relations task, all possible combinations of two concept labels are provided for the students who are asked to provide definitions for each of the concepts and then to describe how the concepts are related. For the sentence generation task, students are given a list of concept labels and asked to use two or more of the concepts at a time in sentences to describe how the concepts are related to one another. In the essay test, a label is placed at the top of a page and the students are given three minutes to write all they know about the concept label.

The data collected by the techniques described can then be graphically represented by concept maps or by semantic networks. In either case, the graphic constructions contain nodes that represent concepts and labeled lines that explain how the nodes are related. This allows the researcher to compare the students' cognitive structure to instructional or discipline structure for consistency. Stewart suggests that if the assessment techniques he describes are combined with other techniques to assess specific problem-solving strategies in science, it might be possible to identify conceptual variables that influence these strategies.

Novak (1990) approaches concept mapping based on Ausubel's (1968) assimilation theory of cognitive learning which holds that cognitive structure is organized hierarchically and that most new learning occurs through the subsumption of new concept meanings within existing concepts and propositional ideas. Thus, concept maps serve as a means of representing specific concept and propositional meanings in an explicit hierarchical framework. From this derives the idea of a hierarchical representation of concept and propositional frameworks in the form of concept maps. The maps are constructed with the most general or inclusive concept listed at the top of the page and more specific or exclusive concepts arranged in descending order down the page. Lines are then drawn linking related concepts and propositional words or phrases written on each line to indicate the nature of the relationship. The structural knowledge that students bring to a problem can then be assessed according to the hierarchy represented by the maps and the nature of the propositions the students indicate as links between the concepts in the map (Novak & Gowin, 1984; Malone & Dekkers, 1984). The concept maps constructed by students change as the students construct new meanings through instruction and experience. This means that if learning takes place during the problem-solving process, examining the students' concept maps before and after these activities will reflect such change.

Assessment of Process Skills

The integrated science process skills are considered by many science educators to be important in problem solving. Three examples of instruments useful in assessing these skills are presented. The Test of Integrated Science Process Skills (TIPS) was developed by Dillashaw and Okey (1980) as a non-curriculum-specific process skills test. The skills to be assessed include stating hypotheses, operationally defining variables, designing investigations, and interpreting data. This is a 36-item paper and pencil instrument employing a four-choice multiple response format. Based on administration of the instrument to a sample of 709 students, the authors reported that "TIPS appears to a valid and reliable measure of process skill achievement for students in the 7th to 12th grade range" (pp. 606-607).

Tobin and Capie (1984) developed the Test of Integrated Science Processes (TISP) to assess student performance on a set of 12 objectives related to the generic objective of planning and conducting an investigation. The test was designed to be used with students over a wide range of ages. The 24-item test was administered to a total of 396 students from middle school to the college level. Analysis of results indicated that the test differentiated students of differing ability with respect to these science process skills. The authors note that the test includes a set of interrelated, cumulative objectives which reflect autonomous problem solving. Recognizing the need for a larger pool of items to assess the integrated science process skills, Burns, Okey, and Wise (1985) developed TIPS II to measure the same set of process skill objectives assessed by TIPS. This is also a 36-item, multiple choice instrument intended for use with students over a wide range of ages. The test was administered to 459 students in grades 7 through 12 and found to have a total test reliability (Cronbach's alpha) of 0.86. To determine the equivalency of TIPS and TIPS II, tests composed of items from the two tests were administered to 359 students in grades 8 through 12. Results of this comparison indicated that the two are highly equivalent tests, with similar mean scores, and scores of the two test were highly correlated. The authors report that TIPS II is a reliable instrument for assessing students' competence in the integrated science process skills and that "reflects the students' ability to apply the logic required to conduct fair investigations" (p. 175).

Assessment of Logical Thinking Skills

The relationship between logical thinking ability and problem solving is reflected in the next three examples. As a result of his search for reliable and valid measures of the developmental levels of learners, Lawson (1978) constructed the Classroom Test of Formal Reasoning. The items included in this instrument assessed the isolation and control of variables, combinatorial reasoning, proportional reasoning, and probabilistic reasoning. The 15 items in the test are based upon Piagetian tasks and involve demonstrations using physical materials and equipment. The questions require the students to select the best answer from those presented and to write an explanation for their choices. Based on the analysis of data from 513 students in grades 8, 9 and 10, Lawson concluded that the test reliably measured the same psychological parameters as did the Piagetian interviews upon which the items were based, and that the test had face validity, convergent validity, and factorial validity.

Tobin and Capie (1981) selected items that had been used by Lawson as a basis for the development of the Test of Logical Thinking (TOLT). "Two items were selected to measure each of five modes of formal reasoning: controlling variables, proportional reasoning, probabilistic reasoning, correlational reasoning, and combinatorial reasoning" (p. 415). A double multiple choice format is used in which the student is presented a problem and asked to select the best response from among five choices, and then to select a reason for this response from among five more choices. This minimizes the effect of guessing, since the student has to choose both the right answer and the best reason in order to get a correct score for the item. The TOLT was validated with data from secondary and college students by correlating the results on its items with performance demonstrated on traditional Piagetian interviews. A correlation of 0.76 was found between TOLT scores and interview tasks which used formal reasoning. Reliability (Cronbach's alpha) for the TOLT was found to be 0.84. Predictive validity was determined to be 0.74 and factor analysis indicated a strong factor accounting for 33 percent of the variance. The evidence supports the TOLT as a valid and reliable instrument for assessing formal reasoning. In order to avoid some of the limitations exhibited by most measures of logical thinking, Roadrangka, Yeany, and Padilla (1983) developed the Group Assessment of Logical Thinking (GALT). The 21-item GALT measures six logical operations: conservation, proportional reasoning, controlling variables, combinatorial reasoning, probabilistic reasoning, and correlational reasoning. The instrument uses a multiple choice format for presenting options for answers as well as for the justification or reason for that answer. Pictorial representations of real objects are used in all the GALT test items. The test was administered to 450 students in grades six through college. The reliability (coefficient alpha) was 0.89, and factor analysis produced a one-factor solution. The validity of the GALT was supported by a correlation of 0.80 between the classification of an individual through the use of the GALT and by using interview techniques. The authors concluded that the test has sufficient reliability and validity to distinguish among groups of students at the concrete, transitional, and formal levels of development. The evidence suggests that the GALT can provide a means for teachers and researchers to assess the cognitive development of a large number of students within a single class period of testing.

Integrating Problem Solving into Curriculum, Instruction, and Assessment

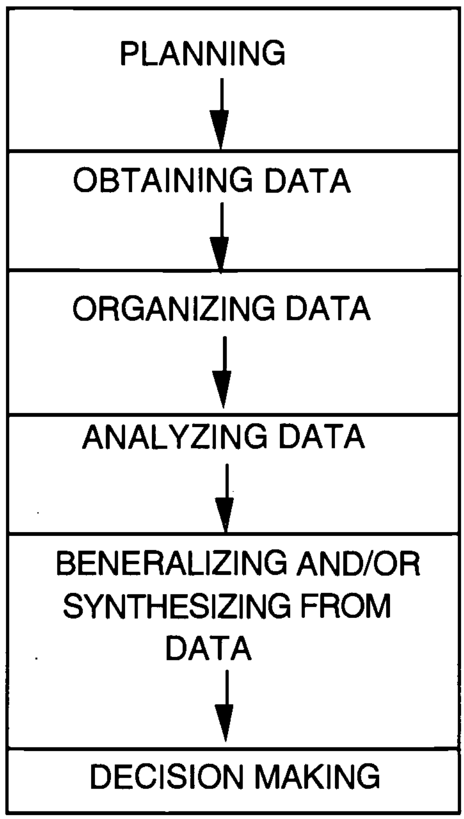
Problem solving has been an important idea in science education for some time, especially at the rhetorical level. Few have attempted to operationalize problem solving at the curriculum, instruction, and assessment levels. This is where the "rubber hits the road."

In the last decade, syllabi for New York State science curricula have begun the long path toward inclusion of problem solving as an overall theme for their courses of study. A scientifically literate person is defined in the New York science syllabi as an effective problem solver. Problem solving should be an integral part of the student's learning experience, rather than being treated as a special topic. "Problem solving is the application of logical and creative thinking to a new and unfamiliar situation requiring resolution" (New York State Education Department, 1987, p. 1).

The New York syllabi go on to link problem solving to science learning experiences, which are based on student recognized problems. The following section for teachers describes the ideal instructional context for facilitating problem solving:

> Before a student can solve a problem the student must recognize that there is a problem to be solved. To help students recognize a problem, use their EXPERIENCES to make them aware of DISCREPANCIES. The discrepancies should lead them to raise QUESTIONS, and the questions will help them to define the problem.
>
> EXPERIENCES can be both spontaneous and teacher-planned activities. The experiences chosen for instruction should be ones that make students aware of discrepancies.
>
> DISCREPANCIES are differences, inconsistencies, disagreements, or disharmonies that we encounter. A discrepancy becomes evident only when we have some prior experience or basis for comparison.
> (New York State Education Department, 1987, p. 1)

Once a student has identified a problem by forming a question, the student can proceed with solving the problem. The New York problem-solving model (New York State Education Department, 1985, p. 8) is a series of steps, most familiar to descriptions of scientific methods. However, the first step (planning) and the last (decision making) may put this model more in a problem-solving mode.



An advantage of the above model is its perceived similarity to existing, traditional science instruction. While that may make it more easily understandable or believable, it can also be accepted as describing what already is going on (likely true in only a small proportion of actual instruction). A major difficulty of this model (and most other models) becomes evident when assessment is planned.

The only statewide assessment program in New York at the elementary school level is the Elementary Science Program Evaluation Test (ESPET), which is a multi-faceted assessment administered to all fourth grade students. It includes a paper and pencil test (45 multiple choice items) and a manipulative skills test (with 22 points possible). This skills test is organized around five tasks (or stations) which primarily assesses individual skills, such as observing, measuring, or inferring. Given the student's limited experiences in science by fourth grade, it was felt that this approach was a useful first step.

As part of the Regents high school physics syllabus (New York State Education Department, 1987), several additional assessment strategies related to problem solving are used. Most of these reflect a belief that the laboratory is a prime vehicle for experiencing problem solving.

Before the final exam can be taken, students must meet these two requirements:

a). Thirty periods (each a minimum of 40 minutes) of laboratory activities. Satisfactory written reports of these experiences must be prepared by the student. Standards for these reports must be established by the local school district at the beginning of the school year.

b). Competency in seven manipulative skills. It is the responsibility of the teacher that these skills have been demonstrated. A sample student skill checklist is included in the physics syllabus. As each skill is satisfactorily demonstrated, the date and initials of both the teacher and student are recorded on this form.

While this approach may be heavily focused on the laboratory phase of problem solving and could be criticized as being single skill oriented, it is more than is done in many school science programs. If some of these laboratory activities are problem oriented, their reports could be excellent assessments of problem solving. The next section will describe some efforts to assess problem solving through performance assessment tasks.

Assessing Problem Solving Through Performance Assessment

Despite a lack of agreement on a definition of problem solving, few would limit its assessment to paper and pencil formats. In fact, some would claim that it is best assessed via practical or performance formats. The APU in Great Britain focused their evaluation of science programs on six categories of science performance.

The APU Categories of Science Performance

1. Use of graphical and symbolic representation
   - reading information from graphs, tables, and charts
   - representing information as graphs, tables, and charts

2. Use of apparatus and measuring instruments
   - using measuring instruments
   - estimating physical quantities
   - following instructions for practical work

3. Observation
   - making and interpreting observations

4. Interpretation and application
   - I interpreting presented information
   - II applying: biology concepts
     physics concepts
     chemistry concepts

5. Planning of investigations
   - planning parts of investigation
   - planning entire investigations

6. Performance of investigations
   - performing entire investigations

(Strang et al., 1991, p. 44)

Their last category - "performance of investigations" seems like a close match to problem solving as conceived by many science educators. The APU researchers used written tasks, a set of practical tasks (short and single skill oriented), and practical investigations to assess science performance overall. The only format of assessment they used for "performance of investigation" was the practical investigation tasks. These tasks involved the student in planning, conducting, and reporting on an investigation designed to solve a given problem. Several of the problems the APU developed for this research have been cited and used in various U.S. projects; survival (choosing fabric to keep you warm, wood lice (finding the right condition for growth), water level (determine if level of water determines speed of water coming out of urn) and swing board (does length of board determine speed of swing?). These tasks were used with students at age 11, 13, and 15 levels in the APU surveys.

Without the practical experiences, students have to "imagine the investigation" in order to write a comprehensive plan. This means that they need a general familiarity with the situations and variables and an "internal model of scientific investigative activity" (Strang et al, 1991, p. 36).

Tamir and Glassman (1971) developed practical investigations as part of the matriculation exam in biology for high school seniors in Israel. Students are given a problem, a set of materials, and a series of questions to guide their work. The student reports of their work are scored at regional centers by science teachers who receive special training for this task. This assessment system in high school biology has been in existence since the early 1970's. Some of the problems given to the students in past years were entitled:

1. Measuring the rate of photosynthesis
2. Daphnia - alternation of activity
3. Measuring the rate of human respiration
4. Grasshopper respiration
5. Yeast fermentation
6. Water relations of a plant tissue

The province of Ontario sponsored the development of Ontario Assessment Instrument Pools (OAIP) for high school courses in biology (1989), chemistry (1981) and physics (1983). Each pool contained many formats of assessment and included some labeled investigations. The Israel and Ontario tasks influenced the work in the US by Doran and colleagues (Doran et al 1992, 1993a, 1993b), who developed a set of six assessment tasks in each of the major science areas. The titles of these tasks are listed below.

	Biology	Chemistry	Physics
Task 1	*Using a Microscope	Acid/Base Titration	Hookes Law
Task 2	*Testing with Indicators	Kinetics/Reaction Rates	Acceleration
Task 3	Model of a Population	The Mole Concept	The Pendulum
Task 4	Diffusion	Solubility	Snell's Law
Task 5	Aerobic Respiration	Melting Point	Electric Circuits
Task 6	Water Holding Capacity	Hydrated Salt	Magnetic Fields

* Did NOT follow the Part A - Part B format

These assessment tasks were constructed in two parts: Part A involved the planning of an investigation for a given problem, and Part B required the performing and reporting of the results of an investigation (with procedures provided to eliminate the double jeopardy problem for those with meager plans). The booklet for Part A was handed in to the test administrator before receiving the booklet for Part B. Both booklets were scored by experienced science teachers (who received special training) and scores were entered on the scoring form located below (Figure 1). The seven skills assessed within each task are among those commonly cited as stages or steps or elements of problem solving.

Novelty of Tasks

A key criteria of tasks categorized as "problem solving" is the context and nature of the problems given to the students. If the task is simply a repeat of the instructional activity, students can respond via recall and will not be stimulated nor reinforced for serious attempts at a more innovative, problem-solving strategies. On the other hand, if the task is too novel (with unfamiliar materials and unrecognized problem) some students will not perceive this to be a fair assessment of the instruction. If this latter situation could be described as "far transfer," Doran and colleagues developed tasks for "near transfer" of science concepts and skills.

Dumas-Carre' and Larcher (1987) described three levels of problems: 1) those in familiar situations requiring a familiar algorithm, 2) those in an unfamiliar situation requiring a familiar algorithm and 3) those which are novel in both respects. Most people attempt to treat all problems by looking for a familiar algorithm to use. Teachers need to assure that the kinds of problems used in assessment are consistent with the instructional experiences.

Polya described four categories of problems in terms of the level of sophistication required for their solution:

- Rule right under your nose - The problem is to be solved by mechanical application of a rule just presented or discussed.

- Application with some choice - The problem can be solved by applying a rule given earlier in class; this involves some judgment or choice on the part of the student.

- Choice of a combination - The student is required to combine two or more rules given in class.

- Approaching research level - The problem's solution requires a novel combination of rules and has many ramifications requiring a high level of independent thought and plausible reasoning (Polya, 1981, p. 139).

School/Student ID No. _____ Reader ID No. _____ Date

Subject Area B C P Task 1 2 3 4 5 6

1. Please circle the NA code if a skill is not assessed in a particular task.
2. The NR code is to be circled when no attempt to respond is apparent.
3. You may check each element present (_) and sum up to determine a student's score for each skill.

Part A: Experiment Design — **Student's Score**

1. STATEMENT OF HYPOTHESIS — NR 0 1 2 3 4 5 NA
   * Effect linked to variable
   * Directionality of effect
   * Expected effect/change
   * Independent variable
   * Dependent variable

2. PROCEDURE FOR INVESTIGATION — NR 0 1 2 3 4 5 NA
   * Resolves experimental problem/feasible
   * Sequenced and detailed plan
   * General strategy
   * Safety procedures
   * Use of equipment/diagram of set-up

3. PLAN TO RECORD AND ORGANIZE OBSERVATIONS/DATA — NR 0 1 2 3 4 5 NA
   * Space for measured/calculated data
   * Matched to plan
   * Organized sequentially
   * Labeled fully (units included)
   * Variables identified

Part B: Experiment Report

4. QUALITY OF OBSERVATIONS/DATA — NR 0 1 2 3 4 5 NA
   * Consistent data
   * Accurate measurements/observations
   * Completed data table
   * Correct units
   * Qualitative description

5. GRAPH — NR 0 1 2 3 4 5 NA
   * Curve is appropriate to data trend
   * Points plotted accurately
   * Appropriate scale (units included)
   * Axes labeled with correct variables
   * Has an appropriate title

6. CALCULATIONS — NR 0 1 2 3 4 5 NA
   * Calculated accurately
   * Substituted correctly into relationship_____
   * Relationship stated or implied ______
   * Units used correctly ________
   * Used all data available ________

7. FORMS A CONCLUSION FROM THE EXPERIMENT — NR 0 1 2 3 4 5 NA
   * Consistent with scientific principle
   * Sources of error __________
   * Consistent with data _________
   * Relationship among variables stated ___
   * Variables stated in conclusion ______

Figure 1: Scoring form for science laboratory test.

Assessing Problem Solving Through Imbedded Formats

A major goal of alternative assessments as described by Baron (1990) is to "blur the distinction" between instruction and assessment. From this point of view, the major difference between instructional and assessment activities is the purpose for their use. Baron wishes to erase that dichotomy between teaching and testing, so sadly illustrated by a teacher saying, "Today we will stop to have a test."

Baron has developed a number of tasks that are incorporated within units of instruction, justifying the label "curriculum imbedded assessment." They have also been called "long term" and "project" assessments, because they take place over several days (even weeks) and are often focused on a problem or project. At times, students work as part of a group and at times they work individually. The titles of some of the Baron projects convey a good match to problem solving:

a) to determine the optimal salinity of water to be used to ship brine shrimp to a friend;

b) to determine which of two liquids is the regular soda pop and which is the diet version;

c) to decide whether a particular site would be appropriate for a nuclear power plant; and

d) to determine the distance that a hot wheels car can jump between ramps when released from a given height.

Other long term investigation tasks were developed and used a part of the Introductory Physical Science (IPS) junior high school science curriculum (Dodge, 1970). Two of the most cited IPS assessment tasks are the "Sludge Test" and "Identifying Gases." Below are the problem statement and materials list for "Identifying Gases."

To the student

Arrange apparatus to produce and collect a gas. Two of the materials you have been given will produce a gas when dropped into water; the third is a liquid which must be heated to evolve gas. Collect samples of the gas produced by each of the materials. Find out as much as you can about the properties of the gas produced in each case, using the tests with which you are now familiar. Describe each gas as carefully and as completely as you can. Identify the gas if you can.

Apparatus and Materials

Peg board
2 Clamps
6 Test tubes
1 One-hole No. 2 stopper
4 No-hole No. 2 stoppers
1 Right-angle glass bend
Rubber tubing
1 Plastic bucket or large pan
1 Burner
Wood splints
Matches
1 Alka-Seltzer tablet
1 Calcium carbide (1.2 g)
Solution of ammonium chloride and sodium nitrite (50 cm^3)
Limewater (45 cm^3)
Water
Paper towel

This example certainly represents scientific problem solving for junior high school students. Wiggins speaks clearly to the need for assessment to relate to "the tasks, context, and 'feel' of real-world challenges - in all their messiness" (Wiggins, 1993, p. 200). Wiggins claims that we cannot be said to understand something unless we can employ our knowledge wisely, fluently, flexibly, and aptly in particular and diverse contexts. Wiggins cited the IPS "sludge" test as a minimally structured task that requires thoughtful performance. The student's specific task is to chemically analyze a mixture of solids and liquids. Though the methods and criteria should be quite clear to all students from the course activities, there are no "pat routines, procedures, or recipes for solving the problem. Thus the test faithfully simulates a wide range of real-world 'tests' of chemical analysis" (Wiggins, 1993, p. 205). Directions for strengthening this approach would be to select problems more within the world of student normal experiences such as "Water in Liquid Soap," "pH of Shampoo," and "Sugar in Soda."

Some British researchers (Strang, Daniels, and Bell, 1991) present a summary of how "performing investigations" changes across a wide age range. The following schematic (Figure 2) shows how the complexity of tasks changes in terms of context, independent variable, dependent variable, apparatus, and conceptual burden (Strang et al., 1991, p. 10). While the examples are from different tasks, it is a potential analytic tool determining student proficiency in planning and conducting investigations that are consistent with a cognitive-science perspective.

Elements of progression	Nature of Progression	
Context	Set in everyday, familiar contexts -------> the home the playground the shops	Set in new and increasingly unfamiliar contexts the laboratory the factory the hospital
Variables to be changed (independent)		
Number	Single -----------------> guitar string length	Multiple guitar string length and diameter
Type	Categoric ---------------> color of car type of material gender	Continuous length of car mass of material age
Variable(s) to be measured (dependent) Nature	Can be appropriately judged without making measurements ---------> floating/sinking pitch bendiness	More appropriately measured length temperature voltage
Apparatus	Simple ----------------> rules kitchen scales pipette	Complex micrometer top pan balance burette
Conceptual burden	Low -------------------> Tasks depending on limited understanding or application of particular scientific concepts, e.g.: *Investigate the extent to which a selection of everyday waste decays naturally.*	High Tasks depending on increasing understanding or application of particular scientific concepts, e.g.: *Investigate the factors limiting the rate of photosynthesis*

Figure 2. Progression in APU task complexity across an age range

The elements labeled context and conceptual burden relate to specific content knowledge related to a particular investigation, while the other elements relate to skills of selecting appropriate variables and measurement tools. One could analyze the necessity of each to student growth or progression in performing investigations over time.

Strang and colleagues used another visual (Figure 3) to present a comprehensive summary of levels of performance for students planning investigations (Strang, et al., 1991, p. 23). Level IV represents a more complete plan while students at Level I are not able to develop a plan that includes these critical elements. It may be that these elements are specific cognitions that could be analyzed from a cognitive-science perspective.

IV	- Appropriate independent and dependent variables identified. - Some control variables identified and not mixed up with other variables. - Independent and dependent variable operationalized quantitatively. - At least three values or plural values given for the independent variable. - Measurements repeated.
III	- Appropriate independent and dependent variable identified. - Some control variables identified but may be mixed up with other variables. - Independent variable operationalized quantitatively. - At least two values given for the independent variables. - Dependent variable operationalized quantitatively.
II	- Appropriate independent and dependent variable identified. - No control variable identified. - Independent variable operationalized qualitatively. - Dependent variable operationalized qualitatively. - At least two values given for the independent variable.
I	- A response which meets none of the above combinations of criteria

Figure 3. Descriptions for Levels of Planning Whole Investigations

Content Dimensions of Assessment Tasks

When the APU started developing investigations for use as assessment items, they attempted to write tasks that would be "concept-free." Their motivation for this was to find out more about student's skills with the procedures or processes of science. Although the tasks they developed had very reduced content demands, it became clear that a "concept-free task" does not exist. The understandings student have of the variables in a specific investigation will have a profound effect on what they can do.

For instance, in a task involving the number and sizes of leaves as an important control variable for drawing water up a celery stem, it is not sufficient to have a concept of control alone.

> They need to know what to control. This means they have to use their conceptual knowledge as well as their procedural knowledge to answer the question (Strang et al., 1991, p. 15).

When looking at APU findings, it becomes clear that carrying out investigations involved more than stringing together a number of scientific processes in the right order. Scientific processes are not independent of content, there is an intimate relationship between the way pupils apply scientific processes and their conceptual knowledge. "There are concepts embedded within the variables of any investigation and the way that pupils understand these concepts plays a fundamental part in the strategies they use" (Strang et al., 1991, p. 15). At every stage of the investigating, pupils' existing concepts are being brought to bear on their actions.

Hodson (1982) was very critical of approaches to teach so-called process skills in isolation:

> In essence, I believe it to be philosophically unsound (because it is not science), educationally worthless (because it trivializes learning), and pedagogically dangerous (because it encourages bad teaching). (p. 115)

Emphasis must be placed on an holistic approach to assessing skills in the context of investigation.

When pupils are doing the APU investigations, they were told that they could repeat any part of the investigation at any stage they wished. The APU researchers "found that pupils often changed their minds as they were going along" (Strang et al., 1991, p. 9). They cited that:

> The practical nature of scientific investigations means that pupils are able to try out their ideas in a concrete situation and this can lead them to reflect on and change their ideas. In other words, as the process of going through an investigation, pupils develop their understanding of scientific processes and concepts. (Strang et al, 1991, p. 9).

When youngsters do an investigation, "we can think of pupils developing their own internal 'model' of the investigation. This model of the task must be dependent on their knowledge and understanding of scientific processes and concepts" (Strang, et al, 1991, p. 16). Many features of a task may affect pupils' ability to access their knowledge base.

> The content, method of presentation, apparatus, and illustrations may help pupils to recall particular experiences which influence the way they perceive and understand the task (Strang et al., p. 16).

The pupil's perception of a task (their model) determines how they will approach the investigation.

Context Dimensions of Tasks

The APU also examined student performance in tasks couched within "everyday" and within "scientific" contexts. Surprisingly, they found that students were more successful with tasks having a scientific context. It was thought that the scientific context provided cues by reminding them of activities they had done previously. It also appears that "scientific" responses were expected by the adult developers and fit the scoring model better. However, the APU found that tasks in everyday contexts can encourage pupils to get involved in investigations.

The APU researchers concluded that we must use a wide variety of contexts for assessment tasks.

> This is particularly important given the varied nature of pupil's experiences and when the possibility of introducing gender and cultural bias into assessment is considered . . . If we only assess investigations in a few restricted contexts we may tend to favor one group (Strang et al., 1991, p. 20).

Millar and Driver argue that the performance of an individual on observation and pattern recognition activities "depends on the particular character of the context in which the activity is undertaken and the mental representations used. This is also seen to be the case for 'processes' in more complex activities such as experimentation and problem solving" (Millar and Driver, 1987, p. 48-49). In a review of research on formal reasoning, Lawson (1985) concludes that "although general patterns of reasoning do seem to exist and do influence performance across domains, context effects do occur" (Lawson, 1985, p. 590).

Perkins and Salomon (1989) reviewed the literature surrounding the context dependence of cognitive skills. They cited evidence for the "strong specialist" position and the "strong generalist" position and the historical shifts between these camps. They believe that each camp has oversimplified the interaction between general strategic knowledge and specialized domain knowledge. They summarized by saying:

> General cognitive skills do not function by somehow taking the place of domain-specific knowledge, nor by operating exactly the same way from domain to domain. Rather, cognitive skills are general tools as in much the way the human hand is. Your hands alone are not enough; you need objects to grasp. Moreover, as you reach for an object, whether a pen or ball, you shape your hand to assure a good grip. And you need to learn to handle different objects appropriately -- you don't pick up a baby in the same way you pick up a basket of laundry.
>
> Likewise, general cognitive skills can be thought of as general gripping devices for retrieving and wielding domain-specific knowledge, as hands that need pieces of knowledge to grip and wield and that need to configure to the kind of knowledge in question" (Perkins and Salomon, 1989, p. 23).

The processes of science are at a high level of generality. The transfer or application of these skills from one context to another is not as easy as may be assumed. The crucial factor appears to be the "distance" between the learned context and the applied context. If these high level skills are at too abstract a level or the real examples and practical experiences are too few, the transfer is difficult to attain.

Millar and Driver illustrate the dependence of student performance with process/problem-solving tasks on content knowledge and familiarity of context with the following incident:

> Thus the fourth year low-achiever in science may be able, while out fishing the local river on weekends, to

> infer, hypothesize, and predict with great skill, in a multi-variable situation involving weather conditions, wind direction, light levels, type of fly, species of fish, position on the boat, and so on (Millar and Driver, 1987, pp 53-54).

It is clear that choice of content and context will greatly affect our assessment of pupil skills.

Another perspective to choosing topics for classroom (and we would argue, assessment) experiences is:

> . . . not to 'tidy up' syllabi by cutting out 'difficult' topics . . . where the knowledge base has unexpectedly become 'messy.' Instead these may provide a means of helping pupils towards some awareness of the sheer difficulty of wresting knowledge from recalcitrant nature" (Millar, 1987, p. 116).

Millar suggests finding ways to include "messy" topics, which reflect real world problems to "help students develop a more realistic appreciation of scientific knowledge, its possibilities, and its limitations" (Millar, 1987, p. 116; cf. Roth, this monograph).

Solomon (1988) cites that most children stick to their old ideas, but only in the old context. Everyday conversation about familiar situations will thus carry its burden of non-scientific meaning throughout most of their lives, despite scientific learning. This implies a difficult situation for the unfortunate learner of science. We should not then be surprised if pupils' fluctuate between different domains of explanation using the non-scientific ones more readily when the context seems familiar (Solomon, 1988, p. 106). While this contradicts the popular view that everyday happenings are easy to understand, but that science is difficult, Solomon cited research which supported her statement.

Problem Solving and Cognitive Science

Some researchers have questioned the view of children's understanding as interconnected frameworks of ideas. Some contrast the structure of scientific knowledge with the unstructured and piecemeal character of life-world knowledge, and hence of children's alternative ideas. Ogborn suggested that it is implausible to think of students' reasoning as inferences made from formal propositions, but rather as "memories of an episodic kind" (Ogborn, 1985, p. 146). Similarly, Kuhn (1977) argued that we learn science, not be acquiring complete inter-related structures of theory, but by becoming familiar with and accepting as valid, a collection of concrete problem-solutions, each of which is seen as an appropriate way of going about problems of this type. According to Millar, "These exemplary problem solutions are paradigms. Paradigms in this sense may be theoretical, or experimental, or even be embedded in a piece of equipment" (Millar, 1989, p. 593). Millar describes this "piecemeal model of science knowledge" as consistent with the research on expert and novice problem-solving in physics. Millar summarized that:

> . . . an essential step in reaching a solution is the ability to see one's 'bank' of paradigms and to see the new problem as an example of a particular type. In general, there is increasing support from work in cognitive psychology for the view that knowledge is held in memory as a collection of discrete instances, and that new problems (whether involving the application of existing knowledge as in the research cited above, or the generation of new knowledge) are tackled by analytic, rather than algorithmic methods (Millar, 1989, p. 593).

When a youngster offers a prediction or explanation about a specific phenomena, their response is not seen as being derived from applying some basic 'laws' (or alternative laws) about the branch of science in question to this particular instance; rather she or he is seen as noting similarities or making analogies between the current problem situation and another recalled situation for which the answer is known. The basis of prediction and theoretical understanding is a set of base-level or "paradigm" situations, which the child really "knows" about. Having heard a child's prediction or explanation, we would want to inquire why they think this will happen or why they put forward this explanation.

Millar and Driver (1987), after reviewing a number of philosophies of science, summarize that:

> While we might not wish to deny that scientists have characteristic ways of working, and of reporting their results, we would argue that there is no warrant for portraying the "scientific method" as a series of specifiable stages, or as anything which remotely approaches an algorithm or a set of rules of procedure (Millar and Driver, 1987, p. 41).

They further contend that the "processes of science" are not unique to science, but are characteristic of logical thought in general. These elements or methods of inquiry are used by historians and scholars in many disciplines as well as scientists. Koertge (1969) commented that, "many elements of what is often called scientific method, such as observing carefully, keeping records, and reasoning in an orderly fashion, are as typical of the supermarket manager as of the scientist" (Koertge, 1969, p. 39).

Millar and Driver argue that these "commonly cited 'processes of science' cannot be divorced from content and context, and that it is the content and context which actually give meaning to the so-called 'process-based' activities in science" (Millar and Driver, 1987, p. 42).

Learning is recognized by most educators as an active/constructive process in which the learner brings:

> . . . prior sets of ideas, schemas, or internal mental representations to any interaction with the environment. These enable predictions and interpretations about features in the environment to be made, expectations generated and assessed and, as a result, the mental representation may be changed. Implicit in this view is that learning depends on both the representations a learner brings to a situation and the characteristics of the learning situation itself (Millar and Driver, 1987, pp. 45-46).

This interactionist view means that when children observe or predict about phenomena:

> . . . the approaches they take in problem solving or experimenting, depend crucially on the way they construe their world. Furthermore, since people may differ in their conceptions, they may respond differently to the same task (Millar and Driver, 1987, p. 46).

Further, students (and adults) attempt to retain their old ideas as long as possible. Millar and Driver summarized research results citing that

> . . . students do identify patterns in data, but these may reflect their prior conceptions rather than the pattern intended. Changing students' conceptions requires more than providing practical experiences; it involves, among other things, bringing a different "way of seeing" to bear on the situation (Millar and Driver, 1987, p. 48).

The situation is further complicated by the observation that most of this research has been undertaken with tasks devised and presented by adult researchers. Pupils are being asked to respond to and be evaluated with questions posed from the researcher's knowledge framework. A different perspective on children's problem-solving ability may be seen when they are investigating questions they have posed about the natural world. Christofides-Henrique (1984, cited from Millar and Driver, 1981) described children investigating materials and then pursuing problems they chose themselves. She indicates that they conducted the investigations in logical ways, did control variables and made valid inferences from data. However, it was the variables they chose to investigate and the hypotheses they used that differed from those of "accepted science." According to Christofides-Henrique:

> Some of the hypotheses which they emit and try to verify seem to us . . . strange. They are problems and hypotheses which are of evident actuality to children and seem to make sense only in a physical world which is no longer ours . . . Thus, to hear children speak of the strength of liquids or to see a sixth-grader, among the best pupils, do an experiment in order to prove to us that the oxygen which is in the water is the cause of the flame of an oil lamp, incite us to think that scientific truth, even the most simple . . . belong to a physical world which is not yet his" (Christofides-Henrique, 1984, p. 7).

This led her to infer that the way youngsters commonly planned investigations, chose control variables, and manipulated variables derive from the experimenter's mental representations of the situation in question.

There also have been descriptions of differences in how experts and novices solve problems. While novices tend to focus on the more superficial features of the problem, experts draw on their knowledge base in:

> setting up an initial redescription of the problem often using analogies with other well understood systems and introducing qualities which may not be mentioned explicitly in the question . . . improvement in performance in such activities depends not on the exercise of a general skill but on the development of the learner's content specific knowledge (Millar and Driver, 1987, p. 50).

Jung discussed how critical one's "frame of mind" is to recognizing, describing, and solving problems. According to Jung, "Learning physics means more than learning formulas and schemas, it means learning a different approach, a different frame of mind" (Jung, 1993, p. 40). In interviewing high school students studying optics, Jung found that they switch from one frame to another; a commonsense frame and a physics frame. We should be cautious in interpreting responses and not take for granted that physics says "what nature really is like." Many misunderstandings or misconceptions in teacher-student discussions can be explained as their using different frames to explain phenomena. Most students move slowly to abstract levels of understanding but remain in the context-dependent episodically organized arena of everyday life (Jung, 1993, p. 46).

Future Research Questions

Research in this area needs to follow the tenets for all scientific inquiry; a careful description of hypotheses or questions, a detailed delineation of the linkage between construct and assessment, sufficient evidence for the validity and reliability of the assessment, appropriate selection and adequate sample size for planned statistical comparisons, and conclusions consistent with procedures and inherent limitations. As critics are wont to describe work in new fields as "old wine in new bottles," it is especially important to address these procedural tenets.

As with any new field, initial research may become concentrated in a particular content area or level of schooling. For the findings of research to be widely used and acceptable, the research base should represent that breadth and depth. Specifically, research conducted with physics concepts at the high school level should be replicated with other science concepts at that level of schooling. Similarly, a few key concepts in one area, such as optics, should be examined with youngsters across the entire range of elementary and secondary schooling. Once such a comprehensive catalog of findings is assembled, application by curriculum developers and teachers will be facilitated.

As this work proceeds, agreement will occur between researchers as to appropriate assessments for key concepts and skills relevant to problem solving. Once this happens, action research projects with interested teachers can be a most effective way to "fill the cells" of this research design.

As these research studies are planned and conducted, a wide variety of control or context variables, must be monitored. Some of these variables might include: gender of student and teacher, opportunity to learn, learning environment of school and community, experience and characteristics of teachers, and district and state expectations. Each of these might have an impact on the findings obtained.

One specific area of research which needs attention is the use of new educational technologies such as hypermedia. With some of the newer equipment, many aspects of this research can be investigated quicker, be modified easily, and can collect much information about the details of a student's path to a solution.

Concluding Comments

It is evident that the assessment of problem solving is "unfinished business." This chapter has attempted to present some of the perspectives and techniques that have been researched as well as the myriad of difficulties with this topic. The definitions of problem solving, as well as, the curriculum approaches and instructional strategies are quite diverse. Most of the research cited is being "retro-fitted" into a cognitive-science perspective. This may be unfair to the researchers and their findings. It would be helpful if each cognitive-science researcher clarifies how the elements of their model are being assessed in their research and the evidence for the assessment's validity and reliability. This chapter involved many examples of assessment studies and ways the researchers have supported their claims for the qualities of their assessments. The diversity of the assessment formats cited is considerable; from interviews and concept maps to multiple choice tests and performance assessment. It is hoped that this discussion of assessment issues will assist in our work to incorporate problem-solving outcomes into school science programs.

References

Aesche, D.W. & Parslow, G.R. (1988). Use of questions from the medical biochemistry question bank with the 'Q' instruction package. *Biochemical Education,16*(1), *24-28*.

Ausubel, D.P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart & Winston.

Baron, J. (1990). Performance assessment: blurring the edges among assessment, curriculum and instruction. In Champagne, A., Levitts, B. & Collinger, B. (Eds). *Assessment in the service of instruction*. Washington, D.C.: American Association for the Advancement of Science.

Brandwein, P. (1971). *Substance, structure and style in the teaching of science*. New York: Harcourt, Brace, Jovanovich, Inc.

Burns, J.C., Okey, J.R. & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching, 22*, 169-177.

Christofides-Henriques, A. (1984). *On the experimental behavior of primary school children* (Mienc). University at Geneva, Cited from Millar and Driver (1987).

Dillashaw, F.G. & Okey, J.R. (1980). Test of the integrated science process skills for secondary science students. *Science Education, 6*, 601-608.

Dodge, J. (1970). Testing - Measuring student achievement. In *IPS-PS II - A progress report* (pp 11-14). Newton, MA: Educational Development Center.

Doran, R. L., Boorman, J., Chan, A. & Hejaily, N. (1992). Successful laboratory assessment. *The Science Teacher, 59*, 22-27.

Doran, R.L., Boorman, J., Chan, F., & Hejaily, N. (1993a). Authentic assessment: an instrument for consistency. *The Science Teacher, 60*, 37-41.

Doran, R.L., Boorman, J., Chan, F., & Hejaily, N. (1993b). Authentic assessment of high school laboratory skills. *Journal of Research in Science Teaching, 30* , 1121 - 1131.

Dumas-Carre, A. & Larcher, C. (1987). The stepping stones of learning and evaluation. *International Journal of Science Education, 9*(1), 93-104.

Gagne, R. M. (1977). *The conditions of learning.* Third Edition New York: Holt, Rinehart and Winston, Inc.

Hayes, J. R. (1981). *The complete problem solver.* Philadelphia: The Franklin Institute Press.

Herb. A. (1992). Computer adaptive testing. *ADVANCE*, (Newsletter of the Medical Laboratory Professionals, February 17.

Hodson, D. (1982). Is there a scientific method? *Education in Chemistry, 19*, 112-116.

Johnson, S. (1989). *National assessment: The APU science approach.* London: Department of Science and Education.

Jung. W. (1993). Uses of cognitive science to science education. *Science and Education, 2*, 31-56.

Koertge, N. (1969). Toward an integration of content and methods in the science curriculum. *Curriculum Theory Network, 4*(1). 26-44.

Kuhn, T. (1977). *The essential tension.* Chicago: University of Chicago Press.

Kumar, D.D. (in press). *Hypermedia: A tool for alternative assessment?* Education and Training Technology International.

Lawson, A. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching, 15*(1), 11-24.

Lawson, A. (1985). A review of research on formal reasoning and science education. *Journal of Research in Science Teaching, 22*, 569-617.

Lomask, M., Jacobson, L.S., & Hafner, L.P. (1992). Interactive videodisc as a tool for assessing science teachers' knowledge of safety regulations in school laboratories. Paper presented at the annual meeting of the *National Association for Research in Science Teaching*, Boston, March 21-25.

Malone, J., & Dekkers, J. (1984). The concept map as an aid to instruction in science and mathematics. *School Science and Mathematics, 84*, 220-223.

Meng, E., & Doran, R. L. (1990). What research says. . . About appropriate methods of assessment. *Science and Children, 28*(1), 42-45.

Millar, R. (1989). Constructive criticisms. *International Journal of Science Education 11*, 587-596.

Millar, R. (1987). Research notes - Toward a role for experiment in the science teaching laboratory. *Studies in Science Education, 14*, 109-118.

Millar, R., & Driver, R. (1987). Beyond processes. *Studies in Science Education, 14*, 33-62.

New York State Education Department. (1985). *Elementary science syllabus.* Albany: Author.

New York State Education Department. (1987). *Regents physics syllabus.* Albany: Author.

Novak, J.D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*, 937-949.

Novak, J.D. & Gowin, D.B. (1984). *Learning how to learn.* New York: Cambridge University Press.

Ogborn, J. (1985). Understanding students' understandings: an example from dynamics. *European Journal of Science Education, 7*. 141-150.

Ontario Ministry of Education. (1981). *Ontario assessment instrument pool - Chemistry*. Toronto: Author.

Ontario Ministry of Education. (1983). *Ontario assessment instrument pool - Physics*. Toronto: Author.

Ontario Ministry of Education. (1989). *Ontario assessment instrument pool - Biology* - Draft. Toronto: Author.

Perkins, D., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 19*: 16-25.

Polya, G. (1981). *Mathematical discovery: On understanding learning and teaching problem solving Vol. 2*. New York : John Wiley.

Roadrangka, V., Yeany, R.H., & Padilla, M.J. (1983). *The construction of a group assessment of logical thinking (GALT)*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Boston, April 11-13. (ERIC Document Reproduction Service No. ED 192 024).

Shavelson, R.J., Carey, N.B., & Webb, N.M. (1990). Indicators of science achievement: Options for a powerful policy instrument. *Phi Delta Kappan 71*, 692-697.

Shaw, T. J. (1983). The effect of a process-oriented science curriculum upon problem-solving ability. *Science Education 67*, 615-623.

Smith, M. U. (1991). A view from biology. In Smith, M. U. (Ed.). *Toward a unified theory of problem solving*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 1-9.

Solomon, J. (1988). Research notes - Learning through experiment. *Studies in Science Education, 15*, 103-108.

Stewart, J. (1980). Techniques for assessing and representing information in cognitive structure. *Science Education 64*, 223-235.

Strang, J., Daniels, S., & Bell, J. (1991). *Assessment Matters, No 6: Planning and carrying out investigations*. London: School Examinations and Assessment Council.

Tamir, P., & Glassman, S. (1971). Laboratory tests for BSCS students. *BSCS Newsletter, 42*, 9-13.

Tobin, K.G., & Capie, W. (1984). The development and validation of a group test of integrated science processes. *Journal of Research in Science Teaching, 19*, 131-141.

Tobin, K.G., & Capie, W. (1981). The development and validation of a group test of logical thinking. *Educational and Psychological Measurement, 41*(2), 413-423.

Wiggins, G. (1993). Assessment: authenticity, context and validity. *Phi Delta Kappan, 75*, 200-214.

How Can Assessment Practices Foster Problem Solving?

Marcia C. Linn
Education in Mathematics, Science,
and Technology Tolman Hall,
University of California, Berkeley, CA 94720

Helen C. Clark
Education in Mathematics, Science,
and Technology Tolman Hall,
University of California, Berkeley, CA 94720

Abstract

Problem solving and assessment have played complementary, supplementary, and even disparate roles in science education. Historically, in science education, problem-solving researchers have often distinguished what might be called "science information" from what might be called "problem-solving strategies." Researchers disagree about the definition of these two forms of knowledge and their connections. In this paper we discuss how views of assessment and problem solving have shifted and converged historically. We illustrate current interactions between assessment and problem solving by discussing the Computer as Learning Partner project and examine an investigation of performance evaluation in this project. We close with a discussion of trends and recommendations. We suggest that assessment and problem solving should be closely coupled. We question the value of standardized tests that primarily measure recall of science information. We recommend "authentic" problems that require sustained reasoning for valid assessment.

Introduction

Problem solving and assessment have played complementary, supplementary, and even disparate roles in science education. We argue here for assessments that not only determine how well students solve problems but also help students learn to solve new complex problems. In essence, we argue for integrating assessment into the instructional program rather than isolating assessment. This perspective has implications for classroom, state, and national testing programs.

We define assessment broadly to include any systematic information about students that informs educational decision making. We define problem solving broadly as well, including any dilemma, question, or situation that requires linking and combining information to

chapter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for* *ientific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

reach a conclusion. We discuss problem solving and assessment in science and draw examples from a range of historical periods. This is not a review but a position paper, thus our examples are selective rather than exhaustive.

We take a cognitive perspective, examining the way students represent and integrate information. By examining how students solve problems we can characterize the information students draw on as well as the new ideas that students form.

Historically, in science education, problem-solving researchers have often distinguished what might be called "science information" from what might be called "problem-solving strategies." However, researchers disagree about the definition of these two forms of knowledge and their connections. Historically, these definitions have shifted and are starting to converge. Recent contributions to the cognitive perspective make these connections more apparent.

We identify three periods based on collaborative patterns in science education that have contributed to these shifts and convergences (Linn, Songer, & Eylon, in press). Prior to 1950, groups studying science education tended to work separately. Those devising assessments, those studying problem solving, and those developing the curriculum often worked in isolation. As a result, few research programs connected assessment and instruction. We refer to this as the *separation* period. The reforms of the 1960s increased interactions between natural scientists and precollege science instructors as well as other groups. In curriculum reform, the natural scientists took the lead (e.g., Welch, 1979). Science educators were called in to design assessments and perform evaluations of the new materials but were not consulted during the development. We refer to this as the *interaction* period. Starting around 1975, partnerships formed. These groups came together in a context of mutual respect and included individuals with training in natural science, cognitive science, classroom teaching, technology, and other fields. Partnership projects linked assessment and problem solving. We refer to this as the *partnership* period. The first section of this paper discusses historical shifts and convergences in these periods.

We illustrate partnership work in the second section of this paper describing assessment and problem solving for the Computer as Learning Partner (CLP) project.

Recently, many are calling for "performance evaluation" to link problem solving and assessment (Resnick, 1987; Shavelson, Baxter, & Pine, 1992). Performance evaluation refers to measurement that taps how students perform on authentic, problem-solving tasks. In contrast, most evaluations are "predictive," providing an indicator of how students will perform on future classroom or workplace tasks. In the third section of this paper we discuss how performance evaluation in the form of portfolio assessment contributes to understanding in the CLP project. In the same section we also report a pilot project using portfolio assessment in middle school science classrooms.

We conclude this paper with a discussion of trends and recommendations for the future. As the boundaries between science information and science problem-solving strategies become blurred and as we seek to measure performance rather than to predict future success, how should we conceptualize problem solving and assessment?

Trends in Assessment and Problem Solving

We discuss shifts and convergences in assessment and problem solving in three historical periods: separation, interaction, and partnership. We identify the role of science information and science problem-solving strategies in each period.

We analyze three aspects of assessment. One aspect of assessment is classification. Tests classify students both for selection into advanced programs and for comparison between states, school districts, and cultural groups. In this paper we highlight comparisons between males and females.

Another aspect of assessment is course evaluation. Evaluation can determine the fate of an instructional innovation. Matching the assessment to the innovation can challenge educators and policy makers alike.

A third aspect of assessment is diagnosis and learning. Assessments can help students recognize their weaknesses and instructors refine their practices. Views of learning and instruction determine both what information is used for diagnosis and how this information is interpreted.

Separation Period

In the separation period psychologists studying mental abilities investigated general problem-solving strategies. Science educators used tests of science information to assess student progress in science classes. Psychologists studying achievement designed standardized tests of science knowledge. By separating problem-solving strategies and science information researchers encountered confusion and sometimes reached misleading conclusions.

Problem-solving strategies. Psychologists following the mental abilities perspective viewed problem-solving strategies as either (a) innate and unteachable, or (b) acquired from studying logic, Latin, or geometry. These strategies were measured using tests of *general ability*. Around the turn of the century, Binet developed mental tests to distinguish retarded students from those who would benefit from education (Binet, 1908). By World War I, the armed services had discovered the advantages of measures of general ability for distinguishing among recruits and assigning individuals to training programs (Yerkes, 1921).

In science, using mental tests for classification reinforced the view that students succeed in science if they have the necessary ability. Research revealed that students of higher ability were more likely to succeed in advanced science courses (Curtis, 1926; 1931). These results justified the tracking of students in advanced courses.

Science information. During this period, science learning consisted primarily of gaining discipline-specific knowledge. Instruction was designed to tell students about science and students were viewed as absorbing this information. Evaluation of course success emphasized recall of scientific information often measured by standard achievement tests. Many national and international assessments in use today also primarily measure recall of scientific information.

Science achievement tests, when used for course evaluation, often convinced science educators that innovative programs were unnecessary or unsuccessful. For example, early science educators compared student-conducted experiments to teacher demonstrations and found that teacher demonstrations resulted in either the same or better performance on standardized tests (Curtis, 1926; 1931; 1939). Students recalled the results of the experiment at least as well when it was demonstrated as when they conducted it themselves. This data led some researchers to recommended that laboratories be abandoned. Others, finding the studies contradictory, argued that "no very great benefit can be gained by more group studies" (Cunningham, 1924; Curtis, 1926, p. 103).

When science information tests were used for diagnosis of student understanding two main findings arose. First, researchers identified difficult science topics and often argued that these be postponed to advanced courses. Second, researchers reported that students did not always learn what they were told. One researcher, analyzing a test of science information recall noted, for example, that students (a) confused the terms cyclone and tornado, (b) did not know that "wrigglers" are a stage in the life history of mosquitoes, and (c) thought that CO2 stands for company number 2 (Boenig, 1969, p. 153; Matteson & Kambley, 1940). Diagnosis of student difficulties foreshadows the analyses of students' intuitions about science in the interaction period.

Participation patterns. Advanced courses were chosen by males far more than females (Whipple, 1932). Yet, data from mental tests and course grades indicated that females had the ability necessary to succeed in science. Most psychologists constructed mental tests to equalize performance of males and females. Binet modified his test when initial results favored females (Maccoby & Jacklin, 1974). Thorndike noted that his test "perhaps slightly penalizes girls in comparison with boys, having been designed primarily for the latter" (quoted in Gambrill, 1922, p. 232). In science courses, instructors used tests featuring both problem solving and recall of science information and females earned higher grades (Whipple, 1932).

Why did females avoid science courses, given their ability to succeed? First, college courses and careers in science were open primarily to white males. Second, since males predominated in science careers many concluded that males were more interested in science than females. Many commentators, especially women, disagreed (Rossiter, 1982).

Summary. In summary, during the separation period psychologists and science educators devised separate assessments because they believed that problem solving consisted of two distinct components. The mental testing movement produced classification tests that justified tracking of students. Tests of science information were used for course evaluation, student evaluation, and diagnosis of difficulties. Course activities such as student conducted laboratories were rejected when they were less successful than teacher demonstrations in imparting science knowledge. To explain why students succeeded but did not persist, researchers pointed to interest in science. Since most believed that students would learn the information they were told these conclusions made sense.

However, not everyone agreed. Dewey (1938) believed that students learned from sustained projects rather than by absorbing information. He argued against over-reliance on tests of recall of science information. He worried that quantitative research investigations might yield false and irrelevant comparisons and mislead thoughtful educators.

Interaction Period

In the interaction period, at least four lines of research relevant to assessment of problem solving emerged (Eylon & Linn, 1988). Those interested in mental testing subdivided the general measure of problem-solving strategies into more specific abilities including verbal ability, mathematical ability, and spatial ability (Cattell, 1963; Guilford &

Hoepfner, 1971). Evidence that topics in science were difficult motivated some to study developmental constraints on problem solving (e.g., Inhelder & Piaget, 1958). Those specifically interested in how scientists solve problems distinguished between strategies used by experts and those used by novices, emphasizing how experts handled information-processing demands of problem solving (Larkin, McDermott, Simon, & Simon, 1980). And, those interested in science information studied performance on difficult scientific concepts (Driver, 1981; Driver & Easley, 1978).

In general, interactions between groups concerned with science education such as between natural scientists and science teachers convinced researchers that problem solving in science had unique characteristics. Reforms of the science curriculum emphasized science inquiry skills rather than general problem-solving skills (Welch, 1979). Texts described steps in the scientific method. This view of problem solving suggested assessment practices that measured problem solving in science rather than general problem solving. The distinction between science strategies and science information remained but the nature of both changed and research groups regularly redefined the distinction.

Individual differences perspective and assessment. Reasoners expanded the mental testing view of general ability to include many abilities during this period. For example, Cattell (1963) distinguished fluid and crystallized ability. Fluid ability refers to general problem-solving skills similar to the skills that had been labeled general ability. Fluid ability is measured by such tasks as Raven's progressive matrices. Crystallized ability refers to knowledge that can be used to solve problems. Crystallized ability is measured by such tasks as vocabulary tests or science information tests. Thus, ability to learn about the discipline and ability to reason about the discipline were measured separately.

Researchers also expanded the general trait of interest in science to include anxiety about science, interest in specific disciplines, and attribution of success to luck or effort. Some believed interest and anxiety changed little as a result of instruction; others examined how courses influenced interest and anxiety.

Group measures of ability and aptitude proliferated during this period. In science, the National Assessment of Education Progress (NAEP) included measures of science knowledge and science inquiry skills, as well as measures of science interest (National Assessment of Educational Progress (NAEP), 1978a; 1978b; 1979a; 1979b; 1988). These tests used multiple choice items to classify groups of students such as males and females. Women earned lower scores on recall items, but not on inquiry items, attributable, in part, to differences in course experience. These differences were almost completely attributable to females choosing the "I don't know" option more than males (Linn, de Benedictis, Delucchi, Harris, & Stage, 1987). By choosing "I don't know," females avoided guessing among the other options, motivating some to argue that females were more realistic about their abilities than males. Others thought this indicated a lack of confidence on the part of females, perhaps motivated by the societal perception of women as less likely than men to succeed in science.

Developmental perspective on assessment. Another group, inspired by developmental theory, sought to design assessment and curriculum following the theory of Piaget (Inhelder & Piaget, 1958). Bruner (1960) spurred American science curriculum reformers to add scientific reasoning or scientific inquiry skills to the goals of the curriculum. Piagetian theory postulated that scientific inquiry skills developed as the result of a variety of concrete experiences and, once acquired, applied across scientific disciplines (Inhelder & Piaget, 1958).

These ideas resonated with the intuitions of natural scientists who believed that they had developed general skills in scientific research that students should learn. For example, Piagetian studies showing how students developed the ability to separate out the effect of individual variables made sense to natural scientists as did the view that these skills were best learned by doing science rather than by studying geometry, Latin, or logic.

Researchers devised a variety of standard tests of scientific inquiry and used them widely (Lawson, 1985; Shayer, Adey, & Wylam, 1981.). These tests asked students to solve scientific problems, and assumed that the primary contributor to success was science inquiry skill, not science discipline knowledge. Efforts to foster scientific problem solving were often unsuccessful, reinforcing the belief that reasoning tests should be used for selection of those "ready" to succeed (Eylon & Linn, 1988). Consistent with the views of the separation period, some believed that problem-solving skill developed independent of instruction.

Evidence accumulated showing that inquiry assessments failed to measure the contribution of discipline-specific problem-solving skills students learned in science courses (Linn, Clement, & Pulos, 1983). Those convinced that inquiry skills predominated in problem solving relegated evidence of discipline-specific influences to "context" effects. This perspective made comparisons among curricula difficult. Most approaches contributed slightly to inquiry skills while varying in their impact on disciplinary knowledge. Students tended to succeed on problems similar to those emphasized in the curriculum (Shymansky, Kyle, & Alport, 1983).

Information-processing perspective on assessment. Another group, inspired by information-processing theory, work on expert systems, and comparisons of experts and novices, sought to design assessment and curriculum based on information-processing demands (Gagné, 1968). Science information, referred to as declarative knowledge, was distinguished from procedural skills such as planning and evaluating (Larkin et al., 1980). Several important findings from this period, including the limitations on processing capacity identified by Miller (1956), had profound implications for science education. Researchers viewed processing capacity as a major constraint on problem solving and a factor in selecting those ready for instruction (Case, 1974; Siegler, 1976).

From the information-processing perspective, students need to learn how to combine information to solve problems without overloading their processing capacity. Thus, they need to compile knowledge into patterns and procedures that use processing capacity efficiently. Or, they need instruction designed to reduce processing demands.

Designing instruction to reduce processing demands arose in the separation period. Thorndike (1910; 1927) analyzed the vocabulary in textbooks and demonstrated that textbooks often overwhelmed students with advanced words that stood in the way of understanding. For example, Powers (Curtis, 1931, p. 348; Powers, 1925) found that about 5% of the words in early science texts were not in Thorndike's list of the 10,000 most common words. Powers reported that over half of the uncommon words appeared only once, and speculated that texts were treating science concepts too superficially. To improve science textbooks, some sought to reduce the vocabulary demands while others sought to limit the topics covered.

Researchers conducted more complex analyses of behavior when computers became available for modeling student learning. Information-processing analyses were applied to complex tasks such as solving algebra or geometry problems (Anderson, 1983). Problems such as Tower of Hanoi (that required limited discipline-specific knowledge) were studied (Greeno & Simon, 1986). In this research, declarative knowledge defined as discipline-specific information was represented in propositions or production rules. Computer programs were written to simulate the behavior of students acquiring procedural and declarative knowledge (Larkin et al., 1980).

Instruction based on information processing diverged from instruction based on fostering development of inquiry skills. For example, tutors were designed to teach the production rules students needed for geometry, algebra, and LISP (Anderson, Boyle, & Reiser, 1985; McArthur, Stasz, & Zmuidzinas, 1990). These tutors limited exploration, quickly connecting students' errors. Assessments were designed to diagnose the productions that students lacked and to tutor them on these productions.

Assessment from the information-processing perspective emphasized processing limits and diagnosis of missing production rules. These studies provided a more detailed understanding of problem solving, but were limited to relatively straightforward problems in algebra, programming, and a few other domains.

Science concept learning and assessment. Researchers from the concept-learning perspective asked students to solve complex problems and explain their approaches. These studies demonstrated that the development of understanding of scientific concepts such as acceleration required more than just absorbing the information and using procedural skills to apply it (e.g., Champagne, Klopfer, & Gunstone, 1982). Studies of scientific concept understanding revealed that student views were often based on observation and description rather than underlying principles. This work motivated interactions between adherents to the information-processing perspective and adherents to the individual differences perspective (Eylon & Linn, 1988).

Students report coherent scientific ideas (e.g., McCloskey, Caramazza, & Green, 1980), many of which are consistent with notions once held by scientists (Clement, 1991; 1992; Clement, Brown, & Zietsman, 1989; Viennot, 1979; Wiser & Carey, 1983). Thus, students believe that sound dies out, heat and temperature are the same, and the earth is round like a pancake.

At first many labeled these as misconceptions because they differed from the views held by scientists. Labeling them as misconceptions criticizes both the reasoning used to achieve these ideas and the ideas themselves (Nersessian, 1991). Viewing intuitions as struggles to make sense of the world emphasizes the problem-solving aspect of these ideas.

Participation patterns. The continued underrepresentation of women and minorities in science was attributed to specific abilities by some (especially spatial ability), and aptitudes such as interest, confidence, and attribution of success by others (Hyde & Linn, 1986). Controversy arose, however, since women had assumed responsible positions during the wars. Many felt that established members of the community were setting policies that precluded the continued participation of women (Rossiter, 1982; Solomon, 1985). They argued that those who were in positions of power in science preferred to work with others like themselves. As a result, men were more encouraged than women to participate in science and women had difficulty retaining self-esteem and confidence.

Summary. In science during the interaction period, several groups expanded and reconceptualized problem solving and assessment. Individual differences focused on a repertoire of abilities and aptitudes. Developmental psychologists elaborated a role for scientific inquiry skills. Information-processing theorists emphasized processing capacity. Concept-learning researchers refined our understanding of discipline-specific knowledge. Together, these varied research programs raised several issues for problem solving and assessment.

Scientific problem-solving strategies were distinguished from general problem-solving strategies during this period. Assessments were designed to measure science inquiry skills. In general, inquiry skills replaced general ability as a constraint on learners. Science disciplinary knowledge continued to be separate from inquiry skill.

Individual differences in crystallized ability were used to explain why some students were more successful than others on measures of recall. The developmental perspective of Piaget, as well as exposure to concrete scientific experience, was used to explain why some students were better at scientific reasoning than others.

Work in the concept-learning tradition, revealed that students had "intuitions" or "frameworks" about scientific phenomena (Pfundt & Duit, 1991). These frameworks entwine disciplinary knowledge and problem-solving strategies. They call into question the distinction between science information and science inquiry skill setting the stage for the partnership period.

Partnership Period

In the partnership period, groups including cognitive researchers, natural scientists, classroom teachers, and technology experts were formed to take a more comprehensive view of problem solving and assessment. These partnerships addressed complex educational questions such as how to teach a topic like thermodynamics to all middle school students. Partnerships found that the organization of knowledge and metacognitive problem-solving strategies influence performance (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). For example, students who develop metacognitive skills such as reflection, self-monitoring, and self-regulation are better learners perhaps because they take responsibility for their own learning (Champagne et al., 1982; Chi et al., 1989; Gunstone, Gray, & Searle, 1992; White, in press). These themes converged on a new view of problem solving that has been called "situated cognition," or "knowledge integration," among other names (Brown, Collins, & Duguid, 1989; Linn et al., in press).

The shift to viewing problem solving as linked to disciplinary knowledge and metacognition is consistent with evidence that even the best students have views of science that differ from those of scientific experts (Linn & Songer, 1993). This shift has motivated partnerships to investigate (a) conceptual change and (b) knowledge organization.

Conceptual change. The conceptual change approach to instruction emphasizes the factors that lead to new perspectives on scientific phenomena. Telling students about science and expecting them to absorb the answer was unsuccessful. Models of change stemmed from detailed observations of learning. For example, course evaluators devised much more finely tuned assessments than had been common in the past (Cronbach, 1982). Rather than relying on standardized measures of scientific inquiry or scientific knowledge, evaluators observed students as they solved problems, often diagnosing the problem-solving practices of individuals (e.g., Clement, 1992). These studies suggested the advantages of scaffolding students as they link and connect information and organize their ideas.

Several mechanisms for conceptual change have resulted from these investigations (Linn & Songer, 1991a). Some draw on development, others on information processing, but all emphasize that conceptual change occurs in a disciplinary context rather than more generally. All these views also point out that students are responsible for change, drawing on metacognitive skill to reformulate their ideas.

For example, one compelling approach to conceptual change emphasizes that students develop problem-solving patterns for specific situations and then select from this repertoire of patterns when attempting a new problem (Linn, diSessa, Pea, & Songer, 1994). Thus, students might develop a descriptive pattern for solving thermodynamics problems based on observations of temperature changes and a microscopic pattern based on reading a text about molecular kinetic theory. Students develop ability to solve problems by expanding their repertoire of patterns and by learning how to select the relevant pattern. To help students learn following this perspective, partnerships designed materials that provided richer feedback to students, and encouraged more autonomous problem solving. This was often accomplished by using simulations or microworlds (diSessa & Minstrell, in press; Smith, Carey, & Wiser, 1985; White & Frederiksen, 1990).

Knowledge organization. Some partnerships built on studies that compared students and experts solving complex problems. Accomplished problem solvers drew on organized, cohesive views of a science topic and had the capability of reformulating their information to apply it to novel problems. The development of discipline-based problem solving, while influenced by the abilities and aptitudes that had been studied in the past, had many unique features. In particular, general inquiry skills were replaced by metacognitive skills as the student characteristic likely to predict success.

Several projects found that discipline-specific knowledge was most reusable when organized in "patterns" or "templates" or "plans" (Linn & Clancy, 1992; Soloway, Pinto, Letovsky, Littman, & Lampert, 1988). A template or plan is a linked set of activities and considerations that, when applied jointly, generate a problem solution. For example, a template for solving a thermal equilibrium problem would include information about rates of cooling and about objects that are heat sources. Thus, in developing the ability to solve problems in thermodynamics, instruction should help students develop templates that link the concepts to the problem-solving activities. The progression of understanding from this perspective is more discipline dependent than had been anticipated (e.g., Brown et al., 1989; Resnick, 1981).

Consider an example. Suppose that a student is asked to solve the problem of determining which of five different materials, aluminum foil, Saran Wrap, paper, wool, and Styrofoam, would be most effective for maintaining the temperature of a cold beverage that has just been removed the refrigerator, and for maintaining the temperature of a loaf of bread that had just been removed from the oven. Examining how students respond to this question, one quickly recognizes the crucial importance of the knowledge linked to the question by the student. For example, one student might say, "Aluminum foil and other metals feel cold, so I think aluminum would be good for keeping the drink cold. Sweaters made out of wool keep me warm, so I suspect a sweater would be good for warming something up, so I'd wrap the bread in wool." This response involves a number of accurate observations. For example, the learner knows that metals feel cold and that sweaters keep people warm. The learner combines these pieces of information in an apparently logical fashion. The learners' conjectures differ substantially from those of experts. How should this solution be assessed? Should one conclude that the student cannot solve problems because the answer is incorrect? Should one conclude that the student can solve problems, but lacks crucial pieces of information? Should one conclude that in this disciplinary area, the student has difficulties solving problems? Or, should one diagnose the kind of information that the student needs: a template for thermal equilibrium, links between heating and cooling, and templates organizing ideas about metals and other materials. Each of these approaches sends a different message about learning and instruction. All require introduction that guides students as they solve authentic problems. General skills are not sufficient.

Persistence and participation. During the partnership period the participation of women has increased, but women are still less than 20% of the employed scientists and engineers. Women continue to earn salaries that are 60 - 70% of salaries for men with comparable experience (Selvin, 1992). Explanations for these disparities rely more and more on the societal perception that science is a male domain and women face extra challenges (Keller, 1985).

Summary. In summary, during the partnership period research conducted in classrooms has blurred the lines between problem-solving strategies and discipline knowledge. Evidence that conceptual change in science is concept–specific strengthens the links between problem solving and science knowledge. In addition, viewing students as holding a repertoire of ideas and selecting among them places emphasis on metacognition. During this period metacognition has gained importance as an aspect of problem solving and the strategies of scientific inquiry have been less emphasized. These changes motivated substantial alternations to the assessment of student ideas.

This richer view of problem solving results from examining how students solve problems. This requires more authentic, situated, linked, and comprehensive assessments. Standardized measures of inquiry or recall provide only hints about the sort of science understanding that students have acquired. In fact, such tests send the wrong message to students, suggesting that memorizing is more useful than gaining an integrated view of a scientific domain, or that practice of general problem solving is sufficient for developing skill in solving discipline-specific problems. They also fail to emphasize the metacognitive skills that are central to effective knowledge integration.

Yet, individuals and groups wishing to classify students find that knowledge integration assessments are often too difficult to administer and too complex to reduce to a simple score. And, those designing curricula wonder how to measure whether their programs are successful. Answering these questions while incorporating a knowledge integration perspective on learning and instruction requires innovation in assessment. We address these issues in the section on the Computer as Learning Partner curriculum.

Implications

These periods in science education highlight a number of issues. First, the methods available for assessment have changed dramatically. For more discussion of this topic, see Sheppard (1993). In the separation period, the statistics of correlation were just emerging. The studies of aptitudes and abilities accompanied the development of powerful regression models and computational procedures, as well as computer technologies that enabled rapid analysis of complex sets of data. The interaction period accompanied an explosion of social science research comparing learning in contextually-narrow areas, such as the learning of nonsense syllables or pattern recognition as well as rapid expansion of computer modeling techniques including the design of expert systems. The partnership period accompanied the development of a broad range of cognitive methods including protocol analysis and video analysis as well as a growing realization that assessment is best achieved by triangulating information from multiple sources.

Second, throughout these periods, researchers have been motivated to modify and expand their assessment procedures based on changes in the audience and the goals of science education. Initially, during the separation period, science was assumed to be primarily studied by an elite group of individuals who might pursue scientific careers. The interaction period accompanied a desire to reform the science curriculum and add modern scientific topics. Alternative teaching techniques such as discovery learning gained acceptance. Studies of aptitudes and abilities helped to distinguish among those who might profit from one approach or another. Widespread national testing movements began to reveal weaknesses in student understanding of scientific ideas. Researchers sought explanations for these weaknesses and commented that typical instructional materials provided fleeting coverage of a broad range topics, overloading the cognitive-processing capacity of most students. Finally, the partnership period accompanied a broadening of the audience for science instruction. Courses in "science for citizens" as well as "science, technology, and values" have been designed to meet the needs of more citizens.

Third, the demands for assessment information have expanded as accountability requirements have increased. Policy makers and admissions committees want to know which programs, schools, and students are the best. New views of learning and instruction have established that these are complex questions. Often, however, decision makers have the intuition that there should be a simple, straightforward way to measure scientific understanding and see the achievement tests used in the past as established criteria.

To explore the partnership perspective on assessment we analyze a case study of curriculum refinement in the next section. This case study elaborates the cognitive-science perspective we mentioned at the onset. We look closely at how students solve problems and at how they use information and procedures available in the classroom. We then describe an investigation of portfolio assessment techniques. We conclude with some recommendations and implications.

The Computer as Learning Partner: Assessment Practices

To illustrate the relationship between assessment practices and curriculum reform, we look at the Computer as Learning Partner (CLP) curriculum project. Over the past eight years, this project has defined, refined, and reformulated materials for middle school science. We discuss how goals for the curriculum and assessment practices have changed by characterizing four main versions of the curriculum: (a) teaching molecular kinetic explanations, (b) explaining thermal events using heat flow instead of kinetic theory, (c) solving relevant, everyday problems using heat flow, and (d) integrating ideas about thermal phenomena.

Teaching Molecular Kinetic Explanations

In the first version of the CLP curriculum we augmented traditional instruction in thermodynamics with real-time data collection (Linn & Songer, 1991b; Songer & Linn, 1991). The curriculum was delivered using computer-based experiments and worksheets rather than the textbook, the duration of the instruction was dramatically increased to 13 weeks from the usual single week, and the traditional learning goals were retained. Students were taught molecular kinetic theory to explain thermal events.

The assessment primarily informed revision of the curriculum, but was also used to classify students and award grades. Assessment in this version consisted of graphing activities, short answer questions, and an essay on the question: "In general, are heat energy and temperature the same or different? What is the main reason for their similarity or difference?" We also interviewed selected students before and after instruction.

Results were disappointing. Students could complete graphs of phenomena involving familiar variables such as mass, but had difficulty with abstract ideas, such as distinguishing heat and temperature. The most common answer to the heat energy and temperature question was, "Heat is number of calories, and temperature is number of degrees." This memorized response did not help students solve problems.

The traditional goal of solving thermal problems with molecular kinetic theory proved too difficult (see also Wiser & Carey, 1983). The scientifically-accurate molecular kinetic theory was not understood by students, even after 13 weeks of instruction. Over half the students resorted to memorizing and parroting information when the curriculum emphasized molecular kinetic theory. Since it was not feasible to increase the time spent on the topic we sought more attainable goals (Linn & Songer, 1991a).

Explaining Thermal Events Using Heat Flow

In the second version of the CLP curriculum we modified the goals, the activities, and the assessment to improve knowledge integration and foster metacognition. We chose new goals using principles of heat flow.

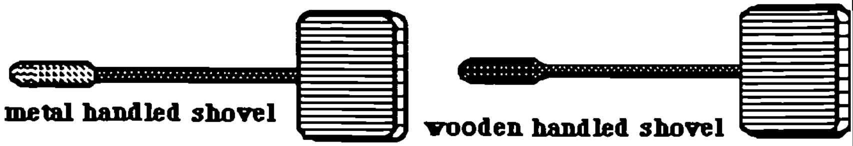
Goals of the heat flow version. We sought accessible goals and added authentic problems. We added pragmatic explanations for thermal events based on heat flow and problems from everyday experience. The CLP heat flow model is descriptive or phenomenological and qualitative. This is one of the repertoire of models used by scientists to explain such everyday events as why a metal spoon gets warmer than a wooden spoon when both are used for stirring a pot of soup as it heats on the stove (Lewis & Linn, 1994). Historically, scientists developed a descriptive model before they determined the microscopic view (e.g., Wiser & Carey, 1983).

We use the term heat energy to refer to the total kinetic energy available for transfer in a substance. We devised heat flow principles such as: (a) Direction of Heat Flow Principle: Heat energy flows only from objects at higher temperature to objects at lower temperature, and (b) Temperature Difference Principle: The greater the temperature difference between objects and their surround, the faster heat energy flows.

Activities. This version of CLP featured an electronic notebook where students could conduct experimental, real-time investigations with classroom apparatus and record their results on a computer. This second version of the curriculum helped students integrate their understanding of thermodynamics by asking for predictions and explanations. Students used evidence from early experiments to predict results of later investigations and to construct principles. The curriculum assessment techniques were expanded to include a broader range of problems, including more everyday problems such as keeping drinks cold in a lunch, and principle construction problems as shown in Figure 1. The second version of the curriculum came closer to the project's goals of helping students distinguish heat energy and temperature, but revealed severe weaknesses in students' abilities to apply these ideas to everyday problems.

The findings from the CLP project concerning equity and assessment are part of a broader set of results. In general, essays and short answer questions are more gender-neutral than multiple-choice questions. International studies, especially those in England where certification tests at the end of secondary school involve both essays and multiple-choice questions, confirm this general trend (see Cresswell, 1990; Linn, 1992; Stobart, Elwood, & Quinlan, 1992 for further discussion). Why do females tend to outperform males on open-ended essays, short answer questions, and course grades? These achievements require linked understanding of scientific events as well as verbal communication skills. Recall of a specific detail is less important than ability to frame an argument relevant to the topic. What disposes students to develop understanding suited to these questions?

A. You are shoveling snow without wearing gloves.



a. Which shovel handle would you rather use? **(circle one)**

metal-handled shovel	**wooden-handled shovel**	**either shovel handle would be OK**

b. Fill in the blanks to make a principle that applies to these shovels.

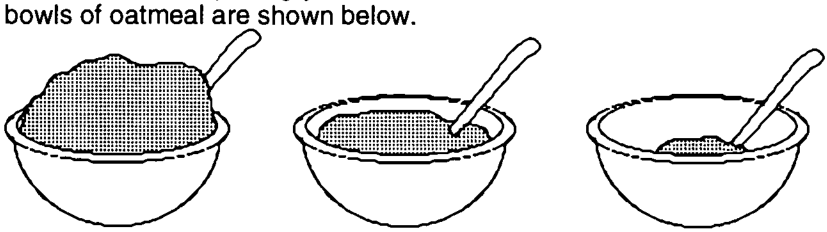
A good_________________________ **allows heat energy to flow faster than**

insulator / conductor

a poor ____________________.

insulator / conductor

B. Three friends were going through a cafeteria line together. One was very hungry and asked for a very, very large bowl of oatmeal. Another was moderately hungry and asked for a large bowl of oatmeal. The third person wasn't very hungry and asked for a small bowl of oatmeal. These bowls of oatmeal are shown below.



a. **Which oatmeal will cool the fastest** so that it could be eaten first?
b. **Why** do you think so?
c. What evidence do you have for your answer?
d. Write a **scientific rule or principle** that applies to this situation.
e. Write **a completely different scientific principle or rule** (not the reverse of the one you just wrote) that also applies to this situation.

Figure 1. Questions that require principle construction

Assessment and equity. On the CLP assessments males and females performed equally (see Figure 2). Furthermore, consistent with national and international results, females earned higher grades than males when the cumulative work of the semester was considered. These results suggest that females and males have equal potential in science.

In contrast to the CLP results, a broad array of multiple choice tests of science achievement such as the National Assessment of Educational Progress (NAEP) report males scoring higher than females. To compare our assessments to NAEP, we administered 25 multiple-choice NAEP items relevant to thermodynamics (Songer & Linn, 1989 March). We found that the eighth grade students in CLP classes outperformed students from similar schools and achieved at a level close to that of 11th graders. Thus, CLP is effective compared to traditional curricula.

We then compared males and females in CLP (see Figure 2). We found that CLP males outperformed females on general NAEP items but not on thermodynamics and graphing items. Thus, males were better at multiple-choice questions about science in general, probably reflecting more informal exposure to science. When the CLP course experience is provided, males and females perform equally on multiple-choice items relevant to the curriculum.

One answer comes from examining the metacognitive goals students set for their own learning. Metacognitive skills are defined as skills in reflection and analyzing one's own learning. A variety of studies indicate that females, compared to males, are more likely to define their objectives in learning complex material as involving gaining an integrated and cohesive view of the material, checking their ideas against those of others, and understanding the main points raised by the instructors (e.g., Linn, 1992; Mann, 1992 October 24). Thus, students' dispositions towards linked understanding may contribute to gender differences in performance on tasks requiring the connecting of ideas compared to measures of recall of specific details.

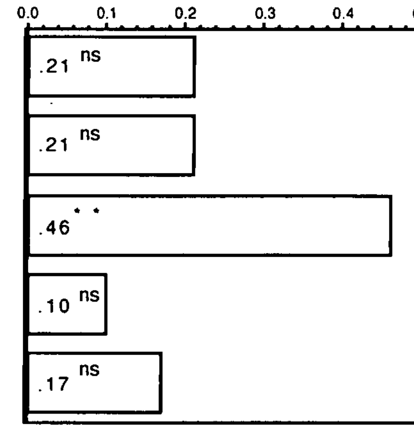
Reactions to the heat-flow version. Linking ideas in science typically involves connecting everyday experiences and views of research scientists. Metacognitive skills contribute to this process. Indeed, difficulties students had in the heat-flow version with everyday problems gave insight into future curricular revision (Songer, 1989). Subsequently, we increased emphasis on knowledge integration and metacognition by scaffolding students as they solve everyday problems.

Solving Relevant, Everyday Problems Using Heat Flow

The third version of the curriculum drew on collaborative discussion of conceptual change informed by assessments, interviews,

Figure 2: Gender Influences in comparitive assessments

	Duration	# of items	Male μ (SD)	Female μ (SD)	Gender Effect Size
CLP Pretest[1]	2 hrs	114	61 (11)	58 (8)	.21 ns
CLP Posttest[1]	2 hrs	114	85 (15)	82 (12)	.21 ns
NAEP General[2]	25 min	25	15 (5)	13 (4)	.46 **
NAEP Thermodynamics[2]	(items)	3	1.8 (.8)	1.7 (.8)	.10 ns
NAEP Graphing[2]	(items)	3	1.2 (.8)	1.1 (.7)	.17 ns



(1) Note that pretest to posttest gains for CLP are about two standard deviation units (effect size 2.0). See Lewis (1991) for details.

(2) Songer, N. B. & Linn, M. C. (1989 March).

(ns) Tested and found not significant.

(**) $p < .05$

and observations. We noted that students lacked knowledge relevant to everyday problems (Songer, 1989). We broadened the activities of the curriculum to allow experimental investigation of everyday problems involving heating and cooling, in direct response to difficulties students had generalizing their ideas. These revisions also featured enhanced opportunities for self-monitoring. In this version, students could conduct experimental, simulated investigations of everyday problems. They were encouraged to link their findings to everyday experience. Assessment activities involving more complex and ambiguous situations were added to assess knowledge integration.

To support knowledge integration, the instructional materials incorporated "prototypes." Prototypes serve as well-understood examples for thermal events. They feature problems for which students have accurate predictions as shown in Figure 3, such as whether soup will cool more rapidly if left in a large tureen or if distributed into small bowls. Prototypes appear in the electronic notebook and students are encouraged to link their experimental investigations to these examples. These opportunities to link ideas provide additional scaffolding for students as they integrate their knowledge.

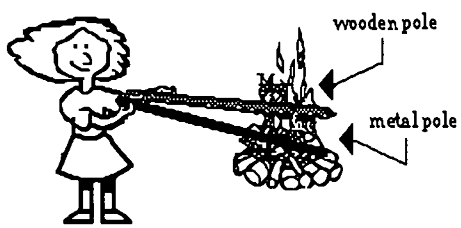
DESCRIPTION

Sam has two poles of equal length. One is made of wood and one is made of metal.

QUESTION

If she holds the ends of both poles and sticks the other ends into the campfire, which pole will heat up first (and burn her hand first!)?

PICTURE



EXPLANATION

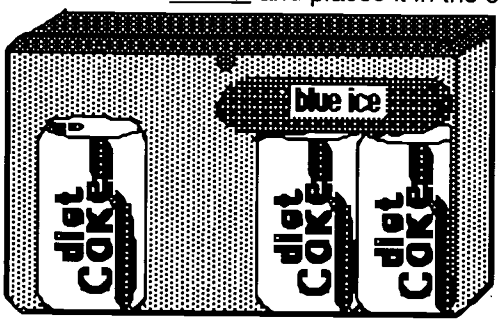
The metal pole would get hotter faster and burn her hand first.

Figure 3. Example of CLP prototype

Assessment practices. Assessment in this revision was augmented to tap explanations of complex, ambiguous, everyday phenomena and metacognitive skills. Students learned the principles and prototypes taught in the CLP curriculum and about one-third formed cohesive views that they could apply to a broad range of problems (see

Figure 4). This was an improvement over the earlier version without prototypes; however, students continued to struggle with applying their understanding to real life events, and with making connections between various parts of the curriculum.

Jessica takes a very good cooler that is at room temperature and places a few room temperature cokes in it. She takes a package of "travel blue ice" out of the freezer and places it in the cooler.



(Circle as many as appropriate)

a. Which will **lose** heat energy?
blue ice cokes air none of these

b. Which will **gain** heat energy?
blue ice cokes air none of these

c. For which will the **temperature increase**?
blue ice cokes air none of these

d. For which will the **temperature decrease?**
blue ice cokes air none of these

e. For which will the **temperature remain the same?**
blue ice cokes air none of these

Figure 4. Example of everyday problem from CLP assessment.

Students' metacognitive skills improved as a result of CLP (Linn & Songer, 1993). In addition, students with robust understanding of the nature of science learned more about thermodynamics than others. Thus, the assessment provided information for both classification of students and for curriculum evaluation. Knowledge integration remained an important and elusive goal.

Evaluating the Integration of Ideas About Thermal Events

In the fourth version of the curriculum we sought to motivate more students to develop an integrated view of thermal phenomena. The students who lacked integrated understanding seemed to believe that such understanding was unnecessary (Linn, Songer, & Lewis, 1991).

To instill more responsibility for knowledge integration and to provide self-diagnosis opportunities, we added portfolio assessment to the curriculum.

Portfolio activities foster integrated understanding in several ways. First, they make students more responsible for their own learning and therefore more likely to use metacognitive skills. Second, they provide multiple opportunities for students to synthesize their understandings. Third, they depend on linking and connecting ideas to solve problems. For example, they engage students as critics, as teachers, and as synthesizers. We describe one investigation of the portfolio approach in the next section of this paper.

These four versions of the Computer as Learning Partner curriculum illustrate both how assessment results can help diagnose student progress and how they help improve the curriculum. Portfolio activities have the potential to help students diagnose their own weaknesses and to motivate students to connect their ideas.

Portfolio Assessment In CLP: An Investigation

Portfolio assessment was added to CLP to stress reflection, synthesis, explanation, and critique. The goal was to increase knowledge integration and metacognitive skill. Portfolios are performance-based assessments in the sense that they measure students proficiency while engaged in *authentic* activities. Authentic activities require sustained thought and effort, integrate ideas across topics, and challenge students to apply their knowledge to novel problems. These new assessments provide opportunity to channel student interest and enthusiasm for projects and increase student control and responsibility. Additionally, portfolios help students gain skill in metacognition by encouraging planning and self-monitoring.

We introduced portfolios to CLP concurrently with the addition of a curriculum unit on light. The addition of concepts of light to a curriculum that had previously addressed concepts of heat and temperature increased the opportunity for knowledge integration. We hoped that the CLP portfolio activities would foster integration across science topics as well as deepen understanding within each topic. CLP portfolio activities encouraged students to combine insights from several experiments and observations. Students reflected on their observations, synthesized and made sense of multiple experiments, and linked their conjectures to everyday experiences.

Portfolios increased student choice of activities in CLP. Students could choose complex, demanding experiments. They could explore a topic in depth. Methods of reporting include oral presentations, video presentations, and artistic interpretations. Portfolios allow serious choices yet provide sufficient feedback and scaffolding to insure that projects are substantive rather than decorative.

Designing of Portfolio Activities

We created portfolio assessment activities to (a) encourage explanation, (b) stimulate reflection, (c) support synthesis, and (d) develop critical evaluation skills.

Encourage explanation. Portfolio activities encourage explanation in two ways. By asking the student to become the teacher and instruct another about the topic, and by asking them to explain a real world situation. For example, one activity asks the student to "tape record or video tape yourself teaching a younger child about light sources and vision. Be sure to find ways to communicate these ideas so that they can be understood by the child. Choose words and images that make sense to that age group. Include the child in your tape with lots of discussion back and forth. Be sure to ask questions to see if the child understands. Role play an entire lesson." To successfully complete this activity students need to recast their own understanding in words and images that a younger student would understand.

Stimulate reflection. Portfolio activities can stimulate reflection by asking the student to use the experience gained from the classroom to predict an outcome of a novel situation. For example a question might be: "Derek is having twenty guests over for dinner. He is going to make spaghetti. He got stuck in traffic on the way home from school and now he only has twenty minutes to cook the noodles. He has one big pan and three smaller pans. He can either cook all the noodles in the one big pan or in the three small pans. His stove has two large burners and two small burners. What can Derek do with these pans and this stove to cook the noodles in the shortest amount of time? What is the main reason for your answer? What principles from this class helped you to make you decision?" This question challenges the student to reflect on which principles apply to making the prediction, and, when more than one are possible (here both surface area and mass) to struggle with how to weigh their effects.

Support synthesis. Portfolio activities support synthesis by drawing attention to possible connections and interrelationships and by encouraging students to seek common themes and principles. For example: "You have conducted three laboratory experiments (named coke, potato, and equilibrium) regarding heat and temperature. What do they have in common? Write a paragraph describing how and in what ways you think these experiments relate to each other." Another example might be: "Are there any similarities between heat energy and light energy? Describe in your own words how you think these may be similar and different. Give several examples." These tasks can be quite difficult for students. To help students succeed, a networked coaching facility encourages students to seek deeper insights into the material rather than rely on superficial ideas. This facility is discussed in a subsequent section.

Develop evaluation skills. Portfolio activities help students develop the ability to evaluate both their own work, and that of others. An example is an activity in which students are told to "Identify another person in class who has done the same activity as you. It can be in any topic area. Exchange copies of your responses. Make a copy of both responses and compare your explanation to the other person's. The emphasis of the comparison should be on identifying where you were similar and different. Write a written summary of your findings, and present it to the other student." By critiquing an activity authored by another student, participants gain insights into their own work. Comparing approaches helps students see the strengths of alternative solutions.

Introducing Portfolio Activities

Portfolio activities increase student responsibility for their own learning by providing choices. To foster autonomous choice, portfolio activities were described in a hypermedia stack available on the classroom computers. The stack created by the research team described each activity and also indicated (a) the rationale for the activities, (b) the expected benefits of the activities, and (c) the methods for seeking help with the activities, including information about the networked coaching facility. Students browsed through the stack according to their own interests.

To encourage knowledge integration and to scaffold use of principles and prototypes in portfolio solutions, we added a curriculum review to the stack a few weeks after making the assignments. The review section of the stack included (a) an index of experiments conducted in the CLP class, and (b) the principles and prototypes used to explain experimental results.

Activities varied in size and complexity. Students were required to select a variety in both complexity and content. When we asked students how they selected activities the most common answer was, "I chose the easiest ones." Some students, however, reported selecting the most interesting or the most challenging activities. In the future we plan to expand the study of factors governing activity selection. Prior research indicates that students need metacognitive skill to make choices (Wainer & Thissen, 1994).

Coaching Students using Portfolio Activities

We investigated several approaches to providing scaffolding and guidance to students as they perform portfolio activities. Without some guidance, students may flounder unnecessarily or produce superficial solutions, especially with the more complex questions. As mentioned, the review section of the hypermedia stack provided memory support for students. We also experimented with a networked coaching facility, hoping to gain some of the advantages of tutoring (Bloom, 1984).

Networked coaching built on the successful intervention conducted by Lewis (1991). Essentially coaches were instructed to foster self-monitoring and reflection (e.g., Chi & Bassok, 1989; Chi et al., 1989). Several members of the research team (four graduate students) agreed to be coaches. Students communicated with their coach via computer electronic mail. Student use of coaches was optional. The coaches never gave direct answers to student assignments, but instead scaffolded students by (a) asking questions, (b) suggesting previous experiments or class discussions that might be relevant to the project, (c) criticizing ideas, (d) encouraging linking of ideas, and (e) cajoling students into thinking about whether their approach made sense. For example, when responding to an item such as "Explain the difference between a blind-folded person in a lighted room, and a sighted person in darkness," many students would stop at a trivial response, "They both can't see." One coach responded: "You might want to spend some time explaining what is happening with the light....Here are some things to think about: ...Does the blindfold matter? Would it matter to a blind person if the lights were out?".

Besides getting input from coaches, students also received feedback from their peers. On larger activities students were encouraged to work in groups of two or three and could monitor each other's ideas as well as their own. At the end of the term, each group presented its large activity to the class and responded to questions and comments from peers.

Results and Implications

Students were required to turn in their completed activities, and to have their large activity presentation evaluated as part of their term grade (Portfolio evaluations comprised approximately 20% of their term grade). All activities were evaluated and categorized according to the level of knowledge integration demonstrated.

In reviewing the impact of these new assessments we looked at the accomplishments of the student, the impact on CLP, and the role of portfolios as an indicator of student performance. Responses were rated on a scale of 1 – 9 with higher score indicating better integrated and more principled understanding of the subject (see Figure 5).

0	No response
1-2	Inaccurate understanding of principle[s], poor effort
3-4	Inaccurate understanding of principle[s], good effort
5-6	Accurately restates principle. No elaboration
7	Clear and accurate understanding of principle and adds elaboration and or context. [i.e.: diagrams, paraphrasing in their own words, or adding "what–if" type comments such as, What would happen if too much light, etc.]
8	Clear and accurate understanding of principle and also ties in one or more additional principles from the same topic area. [i.e.: light sources and reflection]
9	Clear and accurate understanding of principle and ties in one or more principles or examples from another topic area. [i.e.: light principles and heat, or light and biology of retina]

Figure 5. CLP scoring guidelines for portfolio activities.

We analyzed the impact of coaches and the performance on large activities. In the pilot study, 60 CLP students participated. Approximately two-thirds of the students took advantage of coaching. Of course, not all the students who communicated with the coach followed the advice obtained. About 25% of students improved their projects after feedback from the coaches. No students changed their projects for the worse.

The group average gain after coaching was significant. The group average pre-coached submission mean was 4.96 while the group average post-coached mean was 5.66 ($p=.001$). When only students who incorporated coaching feedback are compared, the pre score of 5.5 increases to a post score of 8.25 (see the pattern in Figure 6).

Analysis of class presentations of large activities revealed many examples of knowledge integration and self-diagnosis. The most

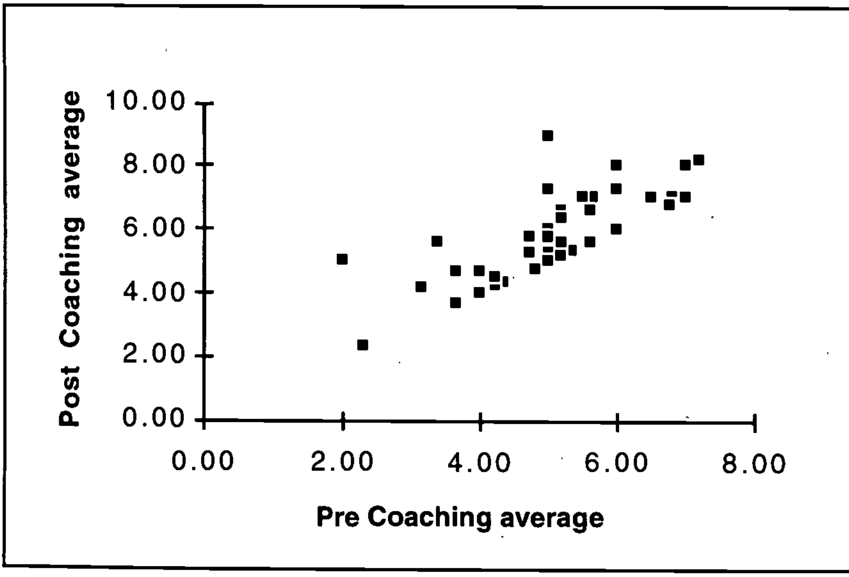


Figure 6. Results of coaching intervention on CLP portfolios.

revealing large activities were intended to develop evaluation skills by asking students to explain complex phenomena and to defend their views against another participant's questioning. In one example, a student teaches a younger sibling about light. The CLP student first identified an analogy using the CLP strategy of teaching with prototypes. When the younger child questions the explanation, the student is forced into self-diagnosis and must either retrieve another explanation or admit that the question is puzzling. One student used a tennis ball as a model for reflecting light. After explaining how light "bounces" off objects and reflects into your eye to see, she then extended the analogy to a discussion of how "hard" the ball is thrown as a parallel for light intensity. This CLP student demonstrated a robust grasp of the principle of reflection. In responding to questions, the CLP student showed both creativity and ability to recognize limits of the model.

In another example, two students are asked to critique each other's work on a previous portfolio assignment. The students' responses are challenged and they are forced to defend or modify their responses in order to reconcile differences. One group of CLP students were trying to solve the question of cooking pasta quickly by either dividing it into three small pots, or using one large. One student emphasized mass as the most important variable, concluding three small was better, while the other focused on surface area opting for one large pot. This pair of students chose to reconcile their differences by ignoring them and concluding "We're both right." Meanwhile, another pair was

1988). A template or plan is a linked set of activities and considerations that, when applied jointly, generate a problem solution. For example, a template for solving a thermal equilibrium problem would include information about rates of cooling and about objects that are heat sources. Thus, in developing the ability to solve problems in thermodynamics, instruction should help students develop templates that link the concepts to the problem-solving activities. The progression of understanding from this perspective is more discipline dependent than had been anticipated (e.g., Brown et al., 1989; Resnick, 1981).

Consider an example. Suppose that a student is asked to solve the problem of determining which of five different materials, aluminum foil, Saran Wrap, paper, wool, and Styrofoam, would be most effective for maintaining the temperature of a cold beverage that has just been removed the refrigerator, and for maintaining the temperature of a loaf of bread that had just been removed from the oven. Examining how students respond to this question, one quickly recognizes the crucial importance of the knowledge linked to the question by the student. For example, one student might say, "Aluminum foil and other metals feel cold, so I think aluminum would be good for keeping the drink cold. Sweaters made out of wool keep me warm, so I suspect a sweater would be good for warming something up, so I'd wrap the bread in wool." This response involves a number of accurate observations. For example, the learner knows that metals feel cold and that sweaters keep people warm. The learner combines these pieces of information in an apparently logical fashion. The learners' conjectures differ substantially from those of experts. How should this solution be assessed? Should one conclude that the student cannot solve problems because the answer is incorrect? Should one conclude that the student can solve problems, but lacks crucial pieces of information? Should one conclude that in this disciplinary area, the student has difficulties solving problems? Or, should one diagnose the kind of information that the student needs: a template for thermal equilibrium, links between heating and cooling, and templates organizing ideas about metals and other materials. Each of these approaches sends a different message about learning and instruction. All require introduction that guides students as they solve authentic problems. General skills are not sufficient.

Persistence and participation. During the partnership period the participation of women has increased, but women are still less than 20% of the employed scientists and engineers. Women continue to earn salaries that are 60 - 70% of salaries for men with comparable experience (Selvin, 1992). Explanations for these disparities rely more and more on the societal perception that science is a male domain and women face extra challenges (Keller, 1985).

Summary. In summary, during the partnership period research conducted in classrooms has blurred the lines between problem-solving strategies and discipline knowledge. Evidence that conceptual change in science is concept–specific strengthens the links between problem solving and science knowledge. In addition, viewing students as holding a repertoire of ideas and selecting among them places emphasis on metacognition. During this period metacognition has gained importance as an aspect of problem solving and the strategies of scientific inquiry have been less emphasized. These changes motivated substantial alternations to the assessment of student ideas.

This richer view of problem solving results from examining how students solve problems. This requires more authentic, situated, linked, and comprehensive assessments. Standardized measures of inquiry or recall provide only hints about the sort of science understanding that students have acquired. In fact, such tests send the wrong message to students, suggesting that memorizing is more useful than gaining an integrated view of a scientific domain, or that practice of general problem solving is sufficient for developing skill in solving discipline-specific problems. They also fail to emphasize the metacognitive skills that are central to effective knowledge integration.

Yet, individuals and groups wishing to classify students find that knowledge integration assessments are often too difficult to administer and too complex to reduce to a simple score. And, those designing curricula wonder how to measure whether their programs are successful. Answering these questions while incorporating a knowledge integration perspective on learning and instruction requires innovation in assessment. We address these issues in the section on the Computer as Learning Partner curriculum.

Implications

These periods in science education highlight a number of issues. First, the methods available for assessment have changed dramatically. For more discussion of this topic, see Sheppard (1993). In the separation period, the statistics of correlation were just emerging. The studies of aptitudes and abilities accompanied the development of powerful regression models and computational procedures, as well as computer technologies that enabled rapid analysis of complex sets of data. The interaction period accompanied an explosion of social science research comparing learning in contextually-narrow areas, such as the learning of nonsense syllables or pattern recognition as well as rapid expansion of computer modeling techniques including the design of expert systems. The partnership period accompanied the development of a broad range of cognitive methods including protocol analysis and video analysis as well as a growing realization that assessment is best achieved by triangulating information from multiple sources.

Second, throughout these periods, researchers have been motivated to modify and expand their assessment procedures based on changes in the audience and the goals of science education. Initially, during the separation period, science was assumed to be primarily studied by an elite group of individuals who might pursue scientific careers. The interaction period accompanied a desire to reform the science curriculum and add modern scientific topics. Alternative teaching techniques such as discovery learning gained acceptance. Studies of aptitudes and abilities helped to distinguish among those who might profit from one approach or another. Widespread national testing movements began to reveal weaknesses in student understanding of scientific ideas. Researchers sought explanations for these weaknesses and commented that typical instructional materials provided fleeting coverage of a broad range topics, overloading the cognitive-processing capacity of most students. Finally, the partnership period accompanied a broadening of the audience for science instruction. Courses in "science for citizens" as well as "science, technology, and values" have been designed to meet the needs of more citizens.

Third, the demands for assessment information have expanded as accountability requirements have increased. Policy makers and admissions committees want to know which programs, schools, and students are the best. New views of learning and instruction have established that these are complex questions. Often, however, decision makers have the intuition that there should be a simple, straightforward way to measure scientific understanding and see the achievement tests used in the past as established criteria.

To explore the partnership perspective on assessment we analyze a case study of curriculum refinement in the next section. This case study elaborates the cognitive-science perspective we mentioned at the onset. We look closely at how students solve problems and at how they use information and procedures available in the classroom. We then describe an investigation of portfolio assessment techniques. We conclude with some recommendations and implications.

The Computer as Learning Partner: Assessment Practices

To illustrate the relationship between assessment practices and curriculum reform, we look at the Computer as Learning Partner (CLP) curriculum project. Over the past eight years, this project has defined, refined, and reformulated materials for middle school science. We discuss how goals for the curriculum and assessment practices have changed by characterizing four main versions of the curriculum: (a) teaching molecular kinetic explanations, (b) explaining thermal events using heat flow instead of kinetic theory, (c) solving relevant, everyday problems using heat flow, and (d) integrating ideas about thermal phenomena.

Teaching Molecular Kinetic Explanations

In the first version of the CLP curriculum we augmented traditional instruction in thermodynamics with real-time data collection (Linn & Songer, 1991b; Songer & Linn, 1991). The curriculum was delivered using computer-based experiments and worksheets rather than the textbook, the duration of the instruction was dramatically increased to 13 weeks from the usual single week, and the traditional learning goals were retained. Students were taught molecular kinetic theory to explain thermal events.

The assessment primarily informed revision of the curriculum, but was also used to classify students and award grades. Assessment in this version consisted of graphing activities, short answer questions, and an essay on the question: "In general, are heat energy and temperature the same or different? What is the main reason for their similarity or difference?" We also interviewed selected students before and after instruction.

Results were disappointing. Students could complete graphs of phenomena involving familiar variables such as mass, but had difficulty with abstract ideas, such as distinguishing heat and temperature. The most common answer to the heat energy and temperature question was, "Heat is number of calories, and temperature is number of degrees." This memorized response did not help students solve problems.

The traditional goal of solving thermal problems with molecular kinetic theory proved too difficult (see also Wiser & Carey, 1983). The scientifically-accurate molecular kinetic theory was not understood by students, even after 13 weeks of instruction. Over half the students resorted to memorizing and parroting information when the curriculum emphasized molecular kinetic theory. Since it was not feasible to increase the time spent on the topic we sought more attainable goals (Linn & Songer, 1991a).

Explaining Thermal Events Using Heat Flow

In the second version of the CLP curriculum we modified the goals, the activities, and the assessment to improve knowledge integration and foster metacognition. We chose new goals using principles of heat flow.

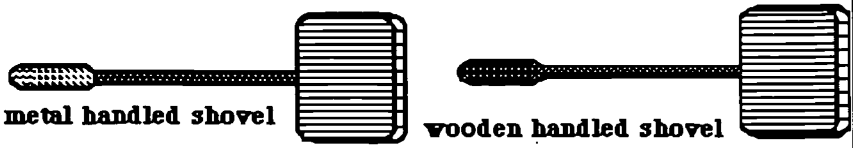
Goals of the heat flow version. We sought accessible goals and added authentic problems. We added pragmatic explanations for thermal events based on heat flow and problems from everyday experience. The CLP heat flow model is descriptive or phenomenological and qualitative. This is one of the repertoire of models used by scientists to explain such everyday events as why a metal spoon gets warmer than a wooden spoon when both are used for stirring a pot of soup as it heats on the stove (Lewis & Linn, 1994). Historically, scientists developed a descriptive model before they determined the microscopic view (e.g., Wiser & Carey, 1983).

We use the term heat energy to refer to the total kinetic energy available for transfer in a substance. We devised heat flow principles such as: (a) Direction of Heat Flow Principle: Heat energy flows only from objects at higher temperature to objects at lower temperature, and (b) Temperature Difference Principle: The greater the temperature difference between objects and their surround, the faster heat energy flows.

Activities. This version of CLP featured an electronic notebook where students could conduct experimental, real-time investigations with classroom apparatus and record their results on a computer. This second version of the curriculum helped students integrate their understanding of thermodynamics by asking for predictions and explanations. Students used evidence from early experiments to predict results of later investigations and to construct principles. The curriculum assessment techniques were expanded to include a broader range of problems, including more everyday problems such as keeping drinks cold in a lunch, and principle construction problems as shown in Figure 1. The second version of the curriculum came closer to the project's goals of helping students distinguish heat energy and temperature, but revealed severe weaknesses in students' abilities to apply these ideas to everyday problems.

The findings from the CLP project concerning equity and assessment are part of a broader set of results. In general, essays and short answer questions are more gender-neutral than multiple-choice questions. International studies, especially those in England where certification tests at the end of secondary school involve both essays and multiple-choice questions, confirm this general trend (see Cresswell, 1990; Linn, 1992; Stobart, Elwood, & Quinlan, 1992 for further discussion). Why do females tend to outperform males on open-ended essays, short answer questions, and course grades? These achievements require linked understanding of scientific events as well as verbal communication skills. Recall of a specific detail is less important than ability to frame an argument relevant to the topic. What disposes students to develop understanding suited to these questions?

A. You are shoveling snow without wearing gloves.



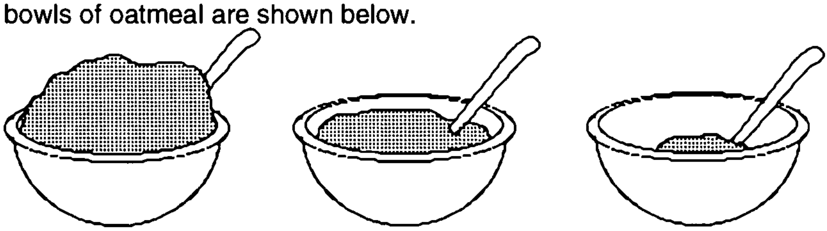
a. Which shovel handle would you rather use? **(circle one)**

metal-handled shovel	**wooden-handled shovel**	**either shovel handle would be OK**

b. Fill in the blanks to make a principle that applies to these shovels.

A good____________________________ allows heat energy to flow faster than
insulator / conductor
a poor ______________________.
insulator / conductor

B. Three friends were going through a cafeteria line together. One was very hungry and asked for a very, very large bowl of oatmeal. Another was moderately hungry and asked for a large bowl of oatmeal. The third person wasn't very hungry and asked for a small bowl of oatmeal. These bowls of oatmeal are shown below.



a. **Which oatmeal will cool the fastest** so that it could be eaten first?
b. **Why** do you think so?
c. What evidence do you have for your answer?
d. Write a **scientific rule or principle** that applies to this situation.
e. Write **a completely different scientific principle or rule** (<u>not the reverse of the one you just wrote</u>) that also applies to this situation.

Figure 1. Questions that require principle construction

Assessment and equity. On the CLP assessments males and females performed equally (see Figure 2). Furthermore, consistent with national and international results, females earned higher grades than males when the cumulative work of the semester was considered. These results suggest that females and males have equal potential in science.

In contrast to the CLP results, a broad array of multiple choice tests of science achievement such as the National Assessment of Educational Progress (NAEP) report males scoring higher than females. To compare our assessments to NAEP, we administered 25 multiple-choice NAEP items relevant to thermodynamics (Songer & Linn, 1989 March). We found that the eighth grade students in CLP classes outperformed students from similar schools and achieved at a level close to that of 11th graders. Thus, CLP is effective compared to traditional curricula.

We then compared males and females in CLP (see Figure 2). We found that CLP males outperformed females on general NAEP items but not on thermodynamics and graphing items. Thus, males were better at multiple-choice questions about science in general, probably reflecting more informal exposure to science. When the CLP course experience is provided, males and females perform equally on multiple-choice items relevant to the curriculum.

One answer comes from examining the metacognitive goals students set for their own learning. Metacognitive skills are defined as skills in reflection and analyzing one's own learning. A variety of studies indicate that females, compared to males, are more likely to define their objectives in learning complex material as involving gaining an integrated and cohesive view of the material, checking their ideas against those of others, and understanding the main points raised by the instructors (e.g., Linn, 1992; Mann, 1992 October 24). Thus, students' dispositions towards linked understanding may contribute to gender differences in performance on tasks requiring the connecting of ideas compared to measures of recall of specific details.

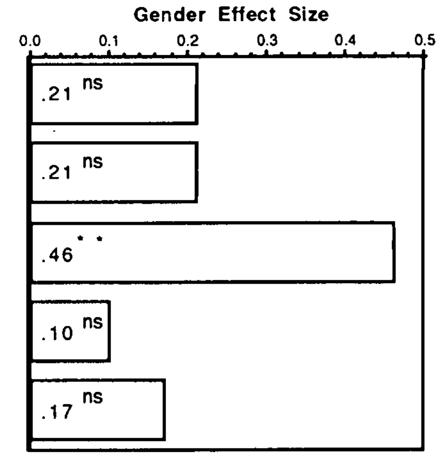
Reactions to the heat-flow version. Linking ideas in science typically involves connecting everyday experiences and views of research scientists. Metacognitive skills contribute to this process. Indeed, difficulties students had in the heat-flow version with everyday problems gave insight into future curricular revision (Songer, 1989). Subsequently, we increased emphasis on knowledge integration and metacognition by scaffolding students as they solve everyday problems.

Solving Relevant, Everyday Problems Using Heat Flow

The third version of the curriculum drew on collaborative discussion of conceptual change informed by assessments, interviews,

Figure 2: Gender Influences in comparitive assessments

	Duration	# of items	Male μ (SD)	Female μ (SD)
CLP Pretest[1]	2 hrs	114	61 (11)	58 (8)
CLP Posttest[1]	2 hrs	114	85 (15)	82 (12)
NAEP General[2]	25 min	25	15 (5)	13 (4)
NAEP Thermodynamics[2]	(items)	3	1.8 (.8)	1.7 (.8)
NAEP Graphing[2]	(items)	3	1.2 (.8)	1.1 (.7)



(1) Note that pretest to posttest gains for CLP are about two standard deviation units (effect size 2.0). See Lewis (1991) for details.

(2) Songer, N. B. & Linn, M. C. (1989 March).

(ns) Tested and found not significant.

(**) $p < .05$

and observations. We noted that students lacked knowledge relevant to everyday problems (Songer, 1989). We broadened the activities of the curriculum to allow experimental investigation of everyday problems involving heating and cooling, in direct response to difficulties students had generalizing their ideas. These revisions also featured enhanced opportunities for self-monitoring. In this version, students could conduct experimental, simulated investigations of everyday problems. They were encouraged to link their findings to everyday experience. Assessment activities involving more complex and ambiguous situations were added to assess knowledge integration.

To support knowledge integration, the instructional materials incorporated "prototypes." Prototypes serve as well-understood examples for thermal events. They feature problems for which students have accurate predictions as shown in Figure 3, such as whether soup will cool more rapidly if left in a large tureen or if distributed into small bowls. Prototypes appear in the electronic notebook and students are encouraged to link their experimental investigations to these examples. These opportunities to link ideas provide additional scaffolding for students as they integrate their knowledge.

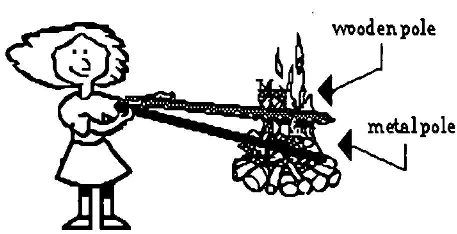
DESCRIPTION

Sam has two poles of equal length. One is made of wood and one is made of metal.

QUESTION

If she holds the ends of both poles and sticks the other ends into the campfire, which pole will heat up first (and burn her hand first!)?

PICTURE



EXPLANATION

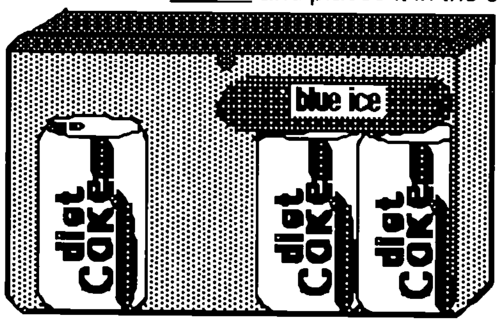
The metal pole would get hotter faster and burn her hand first.

Figure 3. Example of CLP prototype

Assessment practices. Assessment in this revision was augmented to tap explanations of complex, ambiguous, everyday phenomena and metacognitive skills. Students learned the principles and prototypes taught in the CLP curriculum and about one-third formed cohesive views that they could apply to a broad range of problems (see

Figure 4). This was an improvement over the earlier version without prototypes; however, students continued to struggle with applying their understanding to real life events, and with making connections between various parts of the curriculum.

Jessica takes a very good cooler that is at room temperature and places a few room temperature cokes in it. She takes a package of "travel blue ice" out of the freezer and places it in the cooler.



(Circle as many as appropriate)

a. Which will **lose** heat energy?
   blue ice   cokes   air   none of these
b. Which will **gain** heat energy?
   blue ice   cokes   air   none of these
c. For which will the **temperature increase**?
   blue ice   cokes   air   none of these
d. For which will the **temperature decrease?**
   blue ice   cokes   air   none of these
e. For which will the **temperature remain the same?**
   blue ice   cokes   air   none of these

Figure 4. Example of everyday problem from CLP assessment.

Students' metacognitive skills improved as a result of CLP (Linn & Songer, 1993). In addition, students with robust understanding of the nature of science learned more about thermodynamics than others. Thus, the assessment provided information for both classification of students and for curriculum evaluation. Knowledge integration remained an important and elusive goal.

Evaluating the Integration of Ideas About Thermal Events

In the fourth version of the curriculum we sought to motivate more students to develop an integrated view of thermal phenomena. The students who lacked integrated understanding seemed to believe that such understanding was unnecessary (Linn, Songer, & Lewis, 1991).

To instill more responsibility for knowledge integration and to provide self-diagnosis opportunities, we added portfolio assessment to the curriculum.

Portfolio activities foster integrated understanding in several ways. First, they make students more responsible for their own learning and therefore more likely to use metacognitive skills. Second, they provide multiple opportunities for students to synthesize their understandings. Third, they depend on linking and connecting ideas to solve problems. For example, they engage students as critics, as teachers, and as synthesizers. We describe one investigation of the portfolio approach in the next section of this paper.

These four versions of the Computer as Learning Partner curriculum illustrate both how assessment results can help diagnose student progress and how they help improve the curriculum. Portfolio activities have the potential to help students diagnose their own weaknesses and to motivate students to connect their ideas.

Portfolio Assessment In CLP: An Investigation

Portfolio assessment was added to CLP to stress reflection, synthesis, explanation, and critique. The goal was to increase knowledge integration and metacognitive skill. Portfolios are performance-based assessments in the sense that they measure students proficiency while engaged in *authentic* activities. Authentic activities require sustained thought and effort, integrate ideas across topics, and challenge students to apply their knowledge to novel problems. These new assessments provide opportunity to channel student interest and enthusiasm for projects and increase student control and responsibility. Additionally, portfolios help students gain skill in metacognition by encouraging planning and self-monitoring.

We introduced portfolios to CLP concurrently with the addition of a curriculum unit on light. The addition of concepts of light to a curriculum that had previously addressed concepts of heat and temperature increased the opportunity for knowledge integration. We hoped that the CLP portfolio activities would foster integration across science topics as well as deepen understanding within each topic. CLP portfolio activities encouraged students to combine insights from several experiments and observations. Students reflected on their observations, synthesized and made sense of multiple experiments, and linked their conjectures to everyday experiences.

Portfolios increased student choice of activities in CLP. Students could choose complex, demanding experiments. They could explore a topic in depth. Methods of reporting include oral presentations, video presentations, and artistic interpretations. Portfolios allow serious choices yet provide sufficient feedback and scaffolding to insure that projects are substantive rather than decorative.

Designing of Portfolio Activities

We created portfolio assessment activities to (a) encourage explanation, (b) stimulate reflection, (c) support synthesis, and (d) develop critical evaluation skills.

Encourage explanation. Portfolio activities encourage explanation in two ways. By asking the student to become the teacher and instruct another about the topic, and by asking them to explain a real world situation. For example, one activity asks the student to "tape record or video tape yourself teaching a younger child about light sources and vision. Be sure to find ways to communicate these ideas so that they can be understood by the child. Choose words and images that make sense to that age group. Include the child in your tape with lots of discussion back and forth. Be sure to ask questions to see if the child understands. Role play an entire lesson." To successfully complete this activity students need to recast their own understanding in words and images that a younger student would understand.

Stimulate reflection. Portfolio activities can stimulate reflection by asking the student to use the experience gained from the classroom to predict an outcome of a novel situation. For example a question might be: "Derek is having twenty guests over for dinner. He is going to make spaghetti. He got stuck in traffic on the way home from school and now he only has twenty minutes to cook the noodles. He has one big pan and three smaller pans. He can either cook all the noodles in the one big pan or in the three small pans. His stove has two large burners and two small burners. What can Derek do with these pans and this stove to cook the noodles in the shortest amount of time? What is the main reason for your answer? What principles from this class helped you to make you decision?" This question challenges the student to reflect on which principles apply to making the prediction, and, when more than one are possible (here both surface area and mass) to struggle with how to weigh their effects.

Support synthesis. Portfolio activities support synthesis by drawing attention to possible connections and interrelationships and by encouraging students to seek common themes and principles. For example: "You have conducted three laboratory experiments (named coke, potato, and equilibrium) regarding heat and temperature. What do they have in common? Write a paragraph describing how and in what ways you think these experiments relate to each other." Another example might be: "Are there any similarities between heat energy and light energy? Describe in your own words how you think these may be similar and different. Give several examples." These tasks can be quite difficult for students. To help students succeed, a networked coaching facility encourages students to seek deeper insights into the material rather than rely on superficial ideas. This facility is discussed in a subsequent section.

Develop evaluation skills. Portfolio activities help students develop the ability to evaluate both their own work, and that of others. An example is an activity in which students are told to "Identify another person in class who has done the same activity as you. It can be in any topic area. Exchange copies of your responses. Make a copy of both responses and compare your explanation to the other person's. The emphasis of the comparison should be on identifying where you were similar and different. Write a written summary of your findings, and present it to the other student." By critiquing an activity authored by another student, participants gain insights into their own work. Comparing approaches helps students see the strengths of alternative solutions.

Introducing Portfolio Activities

Portfolio activities increase student responsibility for their own learning by providing choices. To foster autonomous choice, portfolio activities were described in a hypermedia stack available on the classroom computers. The stack created by the research team described each activity and also indicated (a) the rationale for the activities, (b) the expected benefits of the activities, and (c) the methods for seeking help with the activities, including information about the networked coaching facility. Students browsed through the stack according to their own interests.

To encourage knowledge integration and to scaffold use of principles and prototypes in portfolio solutions, we added a curriculum review to the stack a few weeks after making the assignments. The review section of the stack included (a) an index of experiments conducted in the CLP class, and (b) the principles and prototypes used to explain experimental results.

Activities varied in size and complexity. Students were required to select a variety in both complexity and content. When we asked students how they selected activities the most common answer was, "I chose the easiest ones." Some students, however, reported selecting the most interesting or the most challenging activities. In the future we plan to expand the study of factors governing activity selection. Prior research indicates that students need metacognitive skill to make choices (Wainer & Thissen, 1994).

Coaching Students using Portfolio Activities

We investigated several approaches to providing scaffolding and guidance to students as they perform portfolio activities. Without some guidance, students may flounder unnecessarily or produce superficial solutions, especially with the more complex questions. As mentioned, the review section of the hypermedia stack provided memory support for students. We also experimented with a networked coaching facility, hoping to gain some of the advantages of tutoring (Bloom, 1984).

Networked coaching built on the successful intervention conducted by Lewis (1991). Essentially coaches were instructed to foster self-monitoring and reflection (e.g., Chi & Bassok, 1989; Chi et al., 1989). Several members of the research team (four graduate students) agreed to be coaches. Students communicated with their coach via computer electronic mail. Student use of coaches was optional. The coaches never gave direct answers to student assignments, but instead scaffolded students by (a) asking questions, (b) suggesting previous experiments or class discussions that might be relevant to the project, (c) criticizing ideas, (d) encouraging linking of ideas, and (e) cajoling students into thinking about whether their approach made sense. For example, when responding to an item such as "Explain the difference between a blind-folded person in a lighted room, and a sighted person in darkness," many students would stop at a trivial response, "They both can't see." One coach responded: "You might want to spend some time explaining what is happening with the light....Here are some things to think about: ...Does the blindfold matter? Would it matter to a blind person if the lights were out?".

Besides getting input from coaches, students also received feedback from their peers. On larger activities students were encouraged to work in groups of two or three and could monitor each other's ideas as well as their own. At the end of the term, each group presented its large activity to the class and responded to questions and comments from peers.

Results and Implications

Students were required to turn in their completed activities, and to have their large activity presentation evaluated as part of their term grade (Portfolio evaluations comprised approximately 20% of their term grade). All activities were evaluated and categorized according to the level of knowledge integration demonstrated.

In reviewing the impact of these new assessments we looked at the accomplishments of the student, the impact on CLP, and the role of portfolios as an indicator of student performance. Responses were rated on a scale of 1 – 9 with higher score indicating better integrated and more principled understanding of the subject (see Figure 5).

0	No response
1-2	Inaccurate understanding of principle[s], poor effort
3-4	Inaccurate understanding of principle[s], good effort
5-6	Accurately restates principle. No elaboration
7	Clear and accurate understanding of principle and adds elaboration and or context. [i.e.: diagrams, paraphrasing in their own words, or adding "what–if" type comments such as, What would happen if too much light, etc.]
8	Clear and accurate understanding of principle and also ties in one or more additional principles from the same topic area. [i.e.: light sources and reflection]
9	Clear and accurate understanding of principle and ties in one or more principles or examples from another topic area. [i.e.: light principles and heat, or light and biology of retina]

Figure 5. CLP scoring guidelines for portfolio activities.

We analyzed the impact of coaches and the performance on large activities. In the pilot study, 60 CLP students participated. Approximately two-thirds of the students took advantage of coaching. Of course, not all the students who communicated with the coach followed the advice obtained. About 25% of students improved their projects after feedback from the coaches. No students changed their projects for the worse.

The group average gain after coaching was significant. The group average pre-coached submission mean was 4.96 while the group average post-coached mean was 5.66 (p=.001). When only students who incorporated coaching feedback are compared, the pre score of 5.5 increases to a post score of 8.25 (see the pattern in Figure 6).

Analysis of class presentations of large activities revealed many examples of knowledge integration and self-diagnosis. The most

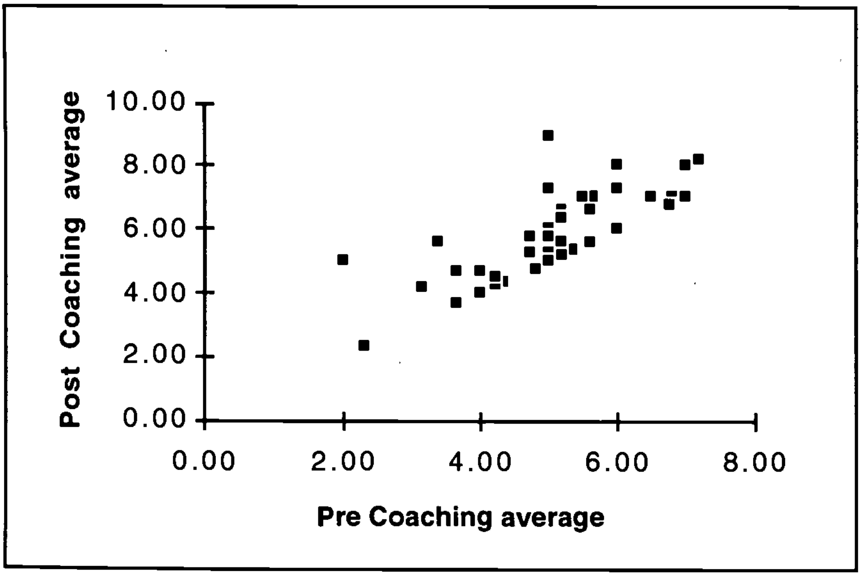


Figure 6. Results of coaching intervention on CLP portfolios.

revealing large activities were intended to develop evaluation skills by asking students to explain complex phenomena and to defend their views against another participant's questioning. In one example, a student teaches a younger sibling about light. The CLP student first identified an analogy using the CLP strategy of teaching with prototypes. When the younger child questions the explanation, the student is forced into self-diagnosis and must either retrieve another explanation or admit that the question is puzzling. One student used a tennis ball as a model for reflecting light. After explaining how light "bounces" off objects and reflects into your eye to see, she then extended the analogy to a discussion of how "hard" the ball is thrown as a parallel for light intensity. This CLP student demonstrated a robust grasp of the principle of reflection. In responding to questions, the CLP student showed both creativity and ability to recognize limits of the model.

In another example, two students are asked to critique each other's work on a previous portfolio assignment. The students' responses are challenged and they are forced to defend or modify their responses in order to reconcile differences. One group of CLP students were trying to solve the question of cooking pasta quickly by either dividing it into three small pots, or using one large. One student emphasized mass as the most important variable, concluding three small was better, while the other focused on surface area opting for one large pot. This pair of students chose to reconcile their differences by ignoring them and concluding "We're both right." Meanwhile, another pair was faced with the dilemma of leaving different size coke bottles in the freezer or refrigerator and then predicting which would be coldest. This group also disagreed. They resolved the dispute by conducting an experiment to "prove" who was right by empirical evidence. These students described multiple scientific principles at work in a single situation and also illustrated their preference for empirical evidence rather than conjuncture.

Summary

The introduction of portfolio activities into CLP supports knowledge integration and provides opportunities for self-diagnosis. These activities afford both learning and assessment opportunities. They reveal more about the students' knowledge than traditional questions. However, if these activities are to be used for primary assessment, more investigation is needed regarding how students make selections and how those selections impact the equity of the assessment. Networked coaching illustrates how seemingly small interventions can have significant impact.

Conclusions

Understanding of the nature of problem solving has led to new methods of assessment and new expectations about the role of assessment. Initially science problem solving was viewed as a combination of two distinct skills: knowledge of the science discipline and general ability. As a result, the first response was to design assessments to measure (a) science inquiry skills and (b) knowledge of science information. However, evidence has accumulated to suggest that science inquiry skills differ from general reasoning skills. When those studying problem solving and those expert in science formed partnerships they identified patterns and templates used by expert scientists that combined inquiry skills and science information. In addition, they noted the contribution to problem solving of what were called metacognitive skills. These insights led to reflection on assessment and calls for more authentic measures of performance.

The assessments that resulted from these insights investigate what we call "knowledge integration" or the ability to link and connect ideas to solve novel problems. Performance-based assessments such as portfolio assessment engage students in solving novel, relevant problems. Portfolio assessment involves offering students a choice of novel problems and requiring students to compile a set of solutions to the problems they choose in the form of a portfolio. In the course of solving novel problems students also extend their learning activities. The learning aspect of portfolio assessment is enhanced with coaching. Rather than viewing coaching as "cheating" those employing this perspective see students' ability to utilize coaching as part of the problem-solving ability that should be measured.

Insights into scientific problem solving have changed both the goals of assessment and the design of assessment activities. As researchers documented the links between reasoning skills and disciplinary knowledge the distinction between general ability to solve problems and performance on specific, complex problems has blurred. Partnership projects have shown that problem solving in science results from learning to recognize discipline-specific patterns and to select from among a repertoire of models. The emphasis on general ability and science information has given way to an emphasis on metacognition and knowledge integration. Shifts in the importance of discipline-specific, problem-solving processes, and metareasoning skills have led to more authentic assessments. These authentic assessments serve two roles: They help students learn science, and they help both teachers and learners to diagnose how well their knowledge is linked and integrated.

A Knowledge Integration Approach

A knowledge integration view of problem solving emphasizes that students develop discipline-specific patterns and reasoning strategies in tandem. Thus, measuring general skills, such as Piagetian logical reasoning or general ability, gives an incomplete and even skewed picture of student understanding. In contrast, measuring problem solving using complex tasks similar to those the student should perform in everyday life or future science investigations offers considerable promise. These new assessment practices are more authentic, in the sense that they are closer to activities that students will be expected to perform in the future. Authentic assessments measure knowledge integration by revealing the models, heuristics, and strategies students use when confronted with novel problems. In designing authentic assessments, it is possible to tap metareasoning skills, as well as abilities to work collaboratively with others. Assessment practices that require sustained problem solving and group work can be costly and complicated to administer. However, many argue that assessment is more accurate when conducted in the context of learning.

This approach is consistent with Vygotsky's proposal that the Zone of Proximal Development serve as an indicator of a student's ability to learn. In Vygotsky's view measuring learning from prompts taps problem-solving potential. The portfolios with coaching described in this paper have the potential of assessing the Zone of Proximal Development or ability to solve problems in a scientific discipline.

The knowledge integration view of students' understanding has expanded the goals of assessment to examine acquisition of patterns, discipline-specific reasoning strategies, and metareasoning.

Furthermore, these studies reinforce the repertoire of models view of conceptual change, and illustrate how learning activities help students refine their repertoire.

Goals of Assessment: Classification, Evaluation, Diagnosis

A knowledge integration perspective on assessment provides rich information about classification, course evaluation, and individual diagnosis. Assessments have served these three different purposes historically and often quite different techniques were used, depending on the purpose of the assessment. Classification initially depended on general ability and individual motivation, and tests of general ability were used. Today, problem solving in the discipline is performed as a measure of classification. Course evaluation initially drew on recall but now involves problem solving as well as diagnosis of course weakness. Diagnosis of student learning difficulties has recently extended to discipline-specific investigations. Researchers have discovered that students often intuit information based on their own prior experiences. Today, assessment practices commonly address all three of these concerns. The primary difference between assessments is not their goal, but the detail with which they assess individuals or groups.

Thus, for evaluating a curriculum, information about group progress is essential, however, detailed classification of each individual may be less necessary. The drawbacks of assessments such as portfolios that include student choice which might influence scores for individuals are less serious when the purpose is course evaluation.

More details on individuals are needed when assessment is used for diagnosis, compared to course evaluation. For students to self-diagnose, complex, authentic assessment activities work well, as long as the students pay attention to the information they provide. Instructors can help students identify relevant information. For instructors, performance assessment requires careful analysis when used for student diagnosis. When knowledge integration assessments like portfolios are used for classifying or grading students, instructors should pay attention to their experiences interacting with the students and to the process students used for solving problems, in addition to examining completed work.

Knowledge integration assessment is least effective for high-stakes classification, where fine distinctions among students are needed. For example, many complain that students who succeed with science fair projects do so because they have special advantages at home or at school. For classroom purposes we recommend that students be rewarded for sustained, creative work, rather than for glitzy presentations or dramatic findings. High-stakes assessment, where it is necessary to very accurately characterize the problem-solving skills of individuals, is difficult and costly when performance measures are used.

Equitable assessment is far more complex than first anticipated. Students could be excluded by invalid assessment practices. Selecting an equitable, valid test requires attention to access to instruction, the nature of the assessment, and the criteria for expected performance. To select students for advanced science courses, ability to solve familiar science problems may be a better predictor than (a) a test of general science knowledge such as a NAEP test or (b) a test of general problem solving in science such as a Piagetian reasoning test.

Implications and Future Directions

A knowledge integration approach to assessment has implications not only for assessment characterizing the abilities of students, but also for the design of instructional materials. The interaction between assessment practices and curriculum reform was illustrated in the Computer as Learning Partner example. Careful analysis of students' problem-solving skills and knowledge organization helps in redesign of a curriculum. Future researchers are encouraged to seek similar connections between student performance and curricular reform. Performance assessment has tremendous potential when used in conjunction with curriculum reform.

Assessment reform raises concerns about equity. Performance assessment offers greater insight into student performance that may also require more observer judgment than traditional approaches. Observer bias and stereotyped expectations can influence decisions. Researchers can help by investigating equity in assessment.

Assessment practices and problem-solving perspectives have shifted and converged over the last century. At the same time, the views of Dewey, expressed at the turn of the century, concerning the need to teach, learn, and assess understanding in the context of projects are now being realized. Portfolio assessments typically involve students carrying out a series of projects, guided by their teachers, much as Dewey might have recommended. As these approaches are refined and as new ideas develop, research with a cognitive-science perspective can make substantial contributions.

Acknowledgements

This material is based upon research supported by the National Science Foundation under grant MDR-9155744. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

This work draws on the research of the Computer as Learning Partner project. We appreciate comments from group members including Philip Bell, Elizabeth Davis, Brian Foley, Doug Kirkpatrick, Eileen Lewis, Jacquie Madhock, Lawrence Muilenburg, Dawn Rickey, and Judith Stern. Thanks also to Dawn Davidson, Anna Chang, and Tiffany Davis for help in the production of this manuscript.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R., Boyle, C. F., & Reiser, B. J. (1985). Intelligent tutoring systems. *Science*, *228*, 456-467.

Binet, A. (1908). The development of intelligence in the child. *L'Annee Psychologique*, *14*, 1-94.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*, 4-16.

Boenig, R. W. (1969). *Research in science education: 1938 through 1947*. New York: Teachers College Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32-41.

Bruner, J. S. (1960). *The process of education*. Cambridge, MA: Harvard University Press.

Case, R. (1974). Structures and strictures: Some functional limitations on the course of cognitive growth. *Cognitive Psychology*, *6*, 544-573.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1-22.

Champagne, A. B., Klopfer, L. E., & Gunstone, R. F. (1982). Cognitive research and the design of science instruction. *Educational Psychologist*, *17*(1), 31-53.

Chi, M. T. H. & Bassok, M. (1989). Learning from examples via self-explanations. In L. B. Resnick (Ed.), *Knowing, learning, and* *instruction: Essays in honor of Robert Glaser* (pp. 251-282). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145-182.

Clement, J. (1991). Non-formal reasoning in science: The use of analogies, extreme cases, and physical intuition. In J. F. Voss, D. N. Perkins, & J. Siegel (Eds.), *Informal reasoning and education*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Clement, J. (1992). Students' preconceptions in introductory mechanics. *American Journal of Physics*, *50*, 66-71.

Clement, J., Brown, D., & Zietsman, A. (1989). Not all preconceptions are misconceptions: Finding "anchoring conceptions" for grounding instruction on students' intuitions. *International Journal of Science Education*, *11*, 554-565.

Cresswell, M. J. (1990). *Gender effects in GCSE: Some initial analyses*. Guildford, England: University of London Institute of Education.

Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco, CA: Jossey-Bass.

Cunningham, H. A. (1924). Laboratory methods in natural science teaching. *School Science and Mathematics*, *24*, 709-715, 848-851.

Curtis, F. D. (1926). *A digest of investigations in the teaching of science in the elementary and secondary schools*. York, PA: Maple Press Co.

Curtis, F. D. (1931). *Second digest of investigations in the teaching of science*. York, PA: Maple Press Co.

Curtis, F. D. (1939). *Third digest of investigations in the teaching of science*. York, PA: Maple Press Co.

Dewey, J. (1938). *Experience and education*. New York: Collier books/Macmillan.

diSessa, A. A. & Minstrell, J. (in press). Cultivating conceptual change with benchmark lessons. In J. G. Greeno (Ed.), *Thinking practices*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Driver, R. (1981). Pupils' alternative frameworks in science. *European Journal of Science Education*, *3*, 93-101.

Driver, R. & Easley, J. (1978). Pupils and paradigms: A review of literature related to concept development in adolescent science. *Studies in Science Education*, *5*, 61-84.

Eylon, B. & Linn, M. C. (1988). Learning and instruction: An examination of four research perspectives in science education. *Review of Educational Research*, *58*(3), 251-301.

Gagné, R. (1968). *Instructional variables and learning outcomes.* Los Angeles: University of California Press.

Gambrill, B. L. (1922). Some administrative uses of intelligence tests in the normal school. In G. M. Whipple (Ed.), *The twenty-first yearbook of the National Society for the Study of Education: Intelligence tests and their use* (pp. 223-244). Bloomington, IL: Public School Publishing Company.

Greeno, J. G. & Simon, H. A. (1986). Problem solving and reasoning. In R. C. Atkinson, R. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Steven's handbook of experimental psychology (Revised ed.).* New York: John Wiley & Sons, Inc.

Guilford, J. P. & Hoepfner, R. (1971). *The analysis of intelligence.* New York: McGraw-Hill.

Gunstone, R. F., Gray, C. M. R., & Searle, P. (1992). Some long-term effects of uninformed conceptual change. *Science Education*, *76*(2), 175-197.

Hyde, J. S. & Linn, M. C. (Eds.). (1986). *The psychology of gender: Advances through meta-analysis.* Baltimore, MD: Johns Hopkins University Press.

Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence; an essay on the construction of formal operational structures.* New York: Basic Books.

Keller, E. F. (1985). *Reflections on gender and science.* New Haven, CT: Yale University Press.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(20), 1335-1342.

Lawson, A. E. (1985). A review of research on formal reasoning and science teaching. *Journal of Research in Science Teaching*, *22*, 569-617.

Lewis, E. L. (1991). *The process of scientific knowledge acquisition among middle school students learning thermodynamics.* Unpublished doctoral dissertation, University of California, Berkeley, CA.

Lewis, E. L. & Linn, M. C. (1994). Heat energy and temperature concepts of adolescents, adults, and experts: Implications for curricular improvements. *Journal of Research in Science Teaching, 31*(6), 657-677.

Linn, M. C. (1992). Gender differences in educational achievement. In J. Pfleiderer (Ed.), *Sex equity in educational opportunity, achievement, and testing [Proceedings of 1991 Educational Testing Service Invitational Conference]* (pp. 11-50). Princeton, NJ: Educational Testing Service.

Linn, M. C. & Clancy, M. J. (1992). Can experts' explanations help students develop program design skills? *International Journal of Man-Machine Studies*, *36*(4), 511-551.

Linn, M. C., Clement, C., & Pulos, S. (1983). Is it formal if it's not physics? *Journal of Research in Science Teaching*, *20*(8), 755-770.

Linn, M. C., de Benedictis, T., Delucchi, K., Harris, A., & Stage, E. (1987). Gender differences in National Assessment of Educational Progress science items: What does "I don't know" really mean? *Journal of Research in Science Teaching, 24*(3), 267-278.

Linn, M. C., diSessa, A., Pea, R. D., & Songer, N. B. (1994). Can research on science learning and instruction inform standards for science education? *Journal of Science Education and Technology*, *3*(1), 7-15.

Linn, M. C. & Songer, N. B. (1991a). Cognitive and conceptual change in adolescence. *American Journal of Education, 99*(4), 379-417.

Linn, M. C. & Songer, N. B. (1991b). Teaching thermodynamics to middle school students: What are appropriate cognitive demands? *Journal of Research in Science Teaching, 28*(10), 885-918.

Linn, M. C. & Songer, N. B. (1993). How do students make sense of science? *Merrill-Palmer Quarterly, 39*(1), 47-73.

Linn, M. C., Songer, N. B., & Eylon, B. (in press). Shifts and convergences in science learning and instruction: Alternative views. In R. Calfee & D. Berliner (Eds.), *Handbook of educational psychology.* Riverside, NJ: Macmillan.

Linn, M. C., Songer, N. B., & Lewis, E. L. (Eds.). (1991). Students' models and epistemologies of science [Special issue]. *Journal of Research in Science Teaching, 28*(9).

Maccoby, E. E. & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Mann, L. (1992 October 24). The relationship between study techniques and success of women and men. Paper presented at the *Taking the Lead: Balancing the Educational Equation*, "Symposium, Gender, Mathematics, Science in College," Oakland, California: Mills College.

Matteson, H. D. & Kambley, P. E. (1940). Knowledge of science possessed by pupils entering the seventh grade. *School Science and Mathematics, 40*, 244-247.

McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction, 7*(3), 197-244.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science, 210*, 1139-1141.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

National Assessment of Educational Progress (NAEP) (1978a). *Science achievement in schools: A summary of results from the 1976-77 national assessment of science*. Denver, CO: Education Commission of the States.

National Assessment of Educational Progress (NAEP) (1978b). *Three national assessments of science: Changes in achievement 1969-77*. Denver, CO: Education Commission of the States.

National Assessment of Educational Progress (NAEP) (1979a). *Attitudes toward science: A summary of results from the 1976-77 national assessment of science*. Denver, CO: Education Commission of the States.

National Assessment of Educational Progress (NAEP) (1979b). *Three national assessments of science, 1969-77: Technical summary*. Denver, CO: Education Commission of the States.

National Assessment of Educational Progress (NAEP) (1988). *The science report card: Elements of risk and recovery: Trends and* *achievement based on the 1986 national assessment*. Princeton, NJ: Educational Testing Service.

Nersessian, N. J. (1991). Conceptual change in science and in science education. In M. R. Matthews (Ed.), *History, philosophy, and science teaching* (pp. 133-148). New York: Teacher's College Press.

Pfundt, H. & Duit, R. (1991). *Students' alternative frameworks* (3rd ed.). Federal Republic of Germany: Institute for Science Education at the University of Kiel/Institut für die Pädagogik der Naturwissenschaften.

Powers, S. R. (1925). The vocabularies of high school science text-books. *Teachers College Record, 26*, 368-392.

Resnick, L. (1981). Instructional psychology. *Annual Review of Psychology, 32*, 659-704.

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher, 16*(9), 13–20.

Rossiter, M. W. (1982). *Women scientists in America: Struggles and strategies to 1940.* Baltimore, MD: Johns Hopkins University Press.

Selvin, P. (1992). Profile of a field: Mathematics—Heroism is still the norm. [Special Section: Women in Science]. *Science, 255*(5050), pp. 1382-1383.

Shavelson, R., Baxter, G. P., & Pine, J. (1992). Performance assessments, political rhetoric and measurement reality. *Educational Researcher*(May).

Shayer, M., Adey, P., & Wylam, H. (1981). Group tests of cognitive development ideals and a realization. *Journal of Research in Science Teaching, 18*, 157-168.

Shepard, L. (1993). Evaluating test validity. In L. Darling-Hammond (Ed.), *Review of Research in Education* (pp. 405-450). Washington DC: American Educational Research Association.

Shymansky, J. A., Kyle, W. C., Jr., & Alport, J. M. (1983). The effects of new science curricula on student performance. *Journal of Research in Science Teaching, 20*, 387-404.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8*, 481-520.

Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: A case study of development of the concepts of size, weight, and density. *Cognition, 21*, 177-237.

Solomon, B. M. (1985). *In the company of educated women.* New Haven, CT: Yale University Press.

Soloway, E., Pinto, J., Letovsky, S., Littman, D., & Lampert, R. (1988). Designing documentation to compensate for delocalized plans. *Communications of the ACM, 31*(11), 1259-1267.

Songer, N. B. (1989). *Promoting integration of instructed and natural world knowledge in thermodynamics.* Unpublished doctoral dissertation, University of California, Berkeley, CA.

Songer, N. B. & Linn, M. C. (1989 March). Assessing depth of knowledge: Performance of students using microcomputer-based labs on NAEP thermodynamics. Paper presented at the *National Association for Research in Science Teaching Annual Meeting*, San Francisco, California.

Songer, N. B. & Linn, M. C. (1991). How do students' views of science influence knowledge integration? *Journal of Research in Science Teaching, 28*(9), 761-784.

Stobart, G., Elwood, J., & Quinlan, M. (1992). Gender bias in examinations: How equal are the opportunities? *British Educational Research Journal, 18*(3), 261-276.

Thorndike, E. L. (1910). The contribution of psychology to education. *Journal of Educational Psychology, 1*, 1-14.

Thorndike, E. L. (1927). *The teacher's word book.* New York: Bureau of Publications, Teachers College, Columbia University.

Viennot, L. (1979). Spontaneous reasoning in elementary dynamics. *European Journal of Science Education, 1*(2), 205-221.

Wainer, H. & Thissen, D. (1994). On examinee choice in educational testing. *Review of Educational Research, 64*(Winter).

Welch, W. W. (1979). Twenty years of science curriculum development: A look back. In D. C. Berliner (Ed.), *Review of research in education* (pp. 282-308). Washington, DC: American Educational Research Association.

Whipple, G. M. (Ed.). (1932). *The thirty-first yearbook of the National Society for the Study of Education: A program for teaching science (Pt. I).* Bloomington, IL: Public School Publishing Company.

White, B. Y. & Frederiksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, *24*(1), 99-157.

White, R. T. (in press). Improving the teaching and learning of science in the post–compulsory years. In P. J. Fensham (Ed.), *Science and technological education in Australia*. Melbourne: Australian Council for Educational Research.

Wiser, M. & Carey, S. (1983). When heat and temperature were one. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 267-298). Hillsdale, NJ: Lawrence Erlbaum Associates.

Yerkes, R. M. (1921). Psychological examining in the United States army. *Memoirs of the National Academy of Science*, *15*.

Computer Applications In Assessing Science Problem Solving

David D. Kumar
Florida Atlantic University
College of Education
2912 College Avenue
Davie, Florida 33314

Stanley L. Helgeson
National Center for Science Teaching and Learning
The Ohio State University
1929 Kenny Road
Columbus, Ohio 43210

Abstract

This chapter examines computer applications in assessing science problem solving from a cognitive science perspective. The computer enables us to record and store information as a student works through a solution pathway, providing us with an indirect view of the cognitive processes in problem solving. Features of computer use in assessment include its convenience as a medium for recording the problem solver's thought processes, the dynamics permitted through hypermedia applications and simulations, and immediate feedback to the user. Selected examples of assessing science problem solving using computers are discussed. The available evidence suggests that computer applications in assessing science problem solving have promise and deserve further study.

Introduction

The emerging role of computer technology in assessing science problem solving will be explored in this chapter. Reform calls in science education have emphasized the need for developing problem-solving skills among students with a long range goal of producing functionally literate citizens ready to exploit the opportunities of life and face the challenges of the information age dominated by scientific and technological applications. For example, the availability of scientific information through computer technology has not only escalated the complex nature of science as a classroom subject but also challenged the focus of science assessment from a mere task of knowledge recall to a process of analyzing and understanding problem-solving strategies in science. Considering the process nature of modern science, it is difficult

apter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for ntific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

to overlook the role of cognitive science without exploring the role of computer technology in scientific problem solving.

Innovations in technological applications in education and developments in cognitive theories of learning are rapidly occurring. Helgeson and Kumar (1993) in a review found that computer applications to science assessment are on the rise. Computer applications to study the processes of problem solving are being developed. Cognitive theories and speculations that relate human thinking to information processing by computers, and non-linear problem solving to hypermedia computer applications have given some basic clues to explore the role of computers in scientific problem solving (Young & Kulikowich, 1992; Larkin & Chabay, 1989). The following discussion will address how computer technology applications in problem solving can be rationalized based on cognitive science, followed by descriptions of a few examples in science education.

Understanding the role of computers in assessing problem solving

While addressing the role of computer technology in assessing problem solving, one should bear in mind that it plays cognitive and motivational roles similar to its role in instructional applications as presented in Larkin and Chabay (1989). Because assessment should not be viewed separate from instruction, assessment is an instructional vehicle (Gronlund & Linn, 1990). Further discussion will address the cognitive aspect of assessing problem solving with computers followed by the motivational aspect.

Computer-based assessment may be viewed as an indirect way of understanding the cognitive processes involved in problem solving. The processes may include decisions made, steps taken, and the application of prior knowledge involved in solving a problem. In a classical sense, problem solving can be defined as a "goal directed sequence of cognitive actions" to find a "solution pathway" from the question state to the final answer state (Newell & Simon, 1972; Chi, Feltovich & Glaser, 1981).

The solution pathway could be considered an attribute of mental "space" described in Larkin and Chabay (1989). For example expert-novice studies on problem solving indicate that experts solve problems in a "mental space of scientific reasoning" while novices "work in a mental "space" of equations" (Larkin & Chabay, 1989, p. 151; Chi, Feltovich & Glaser, 1981; Simon & Simon, 1978). Thus, the processes involved in a solution pathway may be considered indicators of the processes taking place in one's mental space.

So far, there are no direct ways of measuring mental processes in a solution pathway of one engaged in a problem-solving task (Weinstein

& Meyer, 1991). Researches involving brain waves and brain chemistry may one day come up with techniques for the direct measurement of human mental processes. Until then, there are only indirect ways stemming from research in cognitive science such as protocol analysis and concept mapping to understand the processes of problem solving.

When considering indirect ways of measuring human thought processes, computer technology plays a key role in recording and storing information as one interacts with a computer in solving a science problem, and provides a valuable tool for managing such information for assessment (Shavelson, Baxter, Pine, Yure, Goldman, & Smith, 1990; Kumar, 1993). Computers can function as an extension of the human mind and provide a tool for understanding human knowledge structure (Collins, 1990; Young & Kulikowich, 1992; Kumar, 1994). In addition, with the inclusion of hypermedia, computers can interact with humans in a non-linear way, and consequently can help us understand how one might proceed through a solution pathway.

The motivational aspect of computer use in assessing problem solving lies in the fact that it provides a convenient medium for expressing one's thought processes through interaction with a key board or a mouse. Second, computer interaction is not only visual but dynamic when hypermedia applications are involved. Therefore, it is even possible to program near hands-on simulations of science experiments for assessment purposes (Shavelson, Baxter, Pine, Yure, Goldman, & Smith, 1990). Third, using computers it is possible to provide immediate feedback to a problem solver so that he/she can make on-the-spot decisions concerning the direction he/she should take in a solution pathway.

In summary, the following features of computer technology make it a viable tool for assessing problem solving in science.

- Computer technology with hypermedia can interact with human thinking which often involves non-linear mental processes.
- Computer technology can record one's moves in the form of signals as one proceeds through a solution pathway.
- Computer technology can store a vast amount of information about problem solving involving many problem solvers.
- Computer technology can provide a tool for organizing and managing problem-solving information for assessment purposes.

- Computer technology can be developed to provide motivational techniques such as immediate feed-back.

Further discussion will provide selected examples of how computer technology has been used for assessing problem solving from the cognitive science point of view discussed earlier in this paper.

Assessment of Science Problem Solving With Computers

In a study involving Macintosh computer platforms, an assessment software called HyperEquation was developed in HyperCard and implemented for chemistry problem solving (Kumar, White & Helgeson, in press). In the HyperEquation, students log in first using a dialogue box. Then, from a program menu of chemical equations the students select the stoichiometric equation of their choice to balance. In order to get students familiarized with the software it is provided with a help menu. By clicking on a button, just below the place where a coefficient is to be placed, a menu of coefficients is opened. The student chooses from this menu the appropriate number for the coefficient and then clicks on another button (called "DONE") to have the balanced equation checked and to receive immediate feedback. The student can move in a non-linear fashion between the equations and go back to an equation and redo it if he/she desires so. The HyperCard medium is responsible for this non-linear nature of the HyperEquation software.

In addition to providing immediate feedback, HyperEquation can also record students' responses. It also keeps track of the number attempts made by each student on every problem. Above all, HyperEquation keeps track of the total time on task for each student. In an updated version of HyperEquation it is even possible to keep track of the order in which coefficients are placed while the student proceeds through a solution pathway. This version is also programmed to keep time on task for each attempt made by the problem solver. Outcome studies involving this updated software will be available in the near future.

As an assessment software, HyperEquation provides the teacher with an overall and individual record of student performance. The overall record helps the teacher estimate the overall effect of his/her teaching practice on a group of students and modify instructional strategies if necessary. The individual record helps the classroom teacher to provide individualized feedback regarding a student's performance in solving stoichiometric problems.

A study of expert and novice chemistry students solving stoichiometric problems using HyperEquation on a computer platform and traditional paper-and-pen method is reported in Kumar, White and

Helgeson (in press). The study presented five stoichiometric equations with an increasing difficulty level based on the number of steps to balance each equation as determined by Niaz and Lawson (1985). The dependent variables were Performance Score, Number of Attempts, Rate of Attempts, Correctness and Time on Task. For more details on the procedure please refer to Kumar et al. (in press). A few significant findings were as follows.

The expert and novice groups performed significantly better on the computer than with the paper-and-pen method. The Number of Attempts for novices was greater while that for the experts was lower when the computer was used. The Rate of Attempt for the experts in the computer group and for the paper-and-pen group was virtually the same while there was a non-significant difference in the Rate of Attempt between the novice computer group and the novice paper-and-pen group. The novices made fewer attempts with paper and pen than with the computer over a longer time span using HyperEquation.

The Time on Task showed a significant interaction across problem-solving methods and student expertise levels. Student experts spent less time on task with the computer than with paper and pen. Novices spent more time on task with the computer than with paper and pen. Correctness for experts and novices using the computer was greater than for those using paper and pen. Some possible reasons for these effects are discussed below.

The non-linear nature of the computer-based assessment method in HyperCard may be one of the reasons for these differences. For example, the novice group performed better on the problem-solving task with the computer than without. Unlike the paper-and-pen task, in the computer task, the students can select any of the five equation to solve, and move freely to another equation without completing the one they are at, and come back and complete it later.

The mouse input device made the student-computer interface less obtrusive (Schneiderman, 1987) than the pen input at the student-paper interface. The computer-based method provided immediate feedback. In a study involving computer-based biology testing, Collins (1984) noted immediate feedback was a major factor which influenced student achievement. Immediate feedback motivated students to stay on-task until a satisfactory solution was reached. Another reason may be that the computer environment provided an external memory for the problem solvers. This external memory might have reduced the cognitive demand on students' working memory. The amount of working memory space required and the number of variables involved have been found to be related to student performance in problem-solving tasks (Pascual-Leone & Goodman, 1979; Staver, 1986). For example, in a Bending Rod problem Staver (1986) found that the performance of eighth graders improved when the number of independent variables was reduced.

Computerized adaptive testing is emerging as an efficient way to assess student knowledge (Helgeson & Kumar, 1993). Computer-based adaptive testing has been "producing a revolution in educational testing" (Jacobson, 1993, p. A22). According to Jacobson (1993),

> ...in adaptive testing big pools of questions are used to customize a test for each examinee. The first question is generally drawn from items of moderate difficulty. From then on, correct answers bring harder questions and incorrect answers bring easier questions. (p. A23)

Basically, "the computer progressively fine-tunes its assessment of the examinee's ability" (p. A22) by constantly reexamining the ability of the examinee resulting in a test that is tailored to each individual student. The tests can be of various lengths depending upon the pass/fail threshold of the examination.

Adaptive testing with computers is a promising approach to assessing problem solving. Potentially it provides opportunities for research on how students proceed through problems of varying difficulty levels. For example, adaptive testing can provide valuable information on whether there is any trend (progressive, recessive, or random) in the way students attempt to solve problems in tasks that are multiple in nature. Such information will help to strengthen the cognitive nature of computer applications in assessing scientific problem solving.

Herb (1992) reported a pilot study of computerized adaptive testing for certification in medical technology. His study revealed that 50 to 100 adaptive test questions served to provide the necessary pass/fail information as compared to 109 traditional written questions. Also, each examinee took about 30-50% time less to complete the computerized test than the written test, which took four hours to complete. Other advantages include easy access to student scores and ability to manage large pools of data using computers.

Browning and Lehman (1988) described a more useful approach to assessment using computers for identifying student misconceptions and difficulties in genetics problem solving. They have used computers for presenting genetic problems and recording student responses. Their study identified three problem areas. The areas are difficulties with computational skills, difficulties in the determination of gametes, and inappropriate application of previous learning to new problems. Such approaches to assessing problem solving using computers show promise for cognitive studies in science education.

In a study involving computerized assessment Jackson (1988) determined whether computers could accelerate student motivation on task by providing instant feedback with a long range view of improving understanding and enhancing scores in future tests. He found out that science students who underwent the computer-based assessment and received immediate feedback, scored significantly higher in a later test on the same material compared to their counterparts who participated in a paper-and-pencil test. Collins (1984) in another science assessment study involving computers, attributed immediate feedback to enhanced learning.

Implications for Teachers

Assessing science problem solving with computers has several implications for teachers. First, using computers for assessing problem solving should in no way replace the instructional leadership role of classroom teachers. Second, teachers should be able to decide whether to use computers for assessing student performance, managing performance data, or both. Third, teachers should be a significant part of the decision making concerning the kinds of processes to look for using computer-based assessment systems. However, teachers should bear in mind that computer platforms have limitations. A certain computer may not facilitate programming for assessing all the processes involved in a particular problem-solving task. For example, linear computer programs may not be suitable for studying non-linear problem-solving tasks. Fourth, in order to use computers for assessing problem solving in science, computers must also be used as a part of regular classroom instruction. Otherwise, assessment will be far removed from instructional objectives. Therefore, teachers should consider assessment as a part of instruction, and computers as a part of both assessment and instruction. Fifth, teachers should interpret student performance measures derived via computers with caution, and consider such measures only as a part of a set of outcome measures that must include observations and evaluations. Finally, teachers should not substitute computer-based problem solving tasks with hands-on tasks, because computers can never provide the psychomotor and affective benefits of real hands-on experiences.

Discussion

Success in problem solving is determined in part by what knowledge the student brings to the problem. Concept mapping techniques enable us to indirectly assess the student's structural or declarative knowledge (Stewart, 1980; Novak, 1990). With the aid of computers as students construct maps of their knowledge, we can record the sequence of events, the moves made, and the time elapsed. In doing so, we can perhaps begin to delineate and understand the mental processes involved as students bring their knowledge to bear on a problem. A complete understanding of how students structure their knowledge domain for solving problems might provide teachers with the insights for improving science instruction.

Much of the current use of computers in assessing problem solving tends to rely on fixed test instruments administered to all students. The emergence of computerized adaptive assessment allows for greater flexibility, with tests adjusted to the individual student (Herb, 1992). At the present time, such tests are limited to measuring declarative knowledge; however, possibilities exist for further exploration. If we can find ways to assess declarative knowledge while reducing the time and number of items required, we can devote more attention to other considerations in problem solving such as identifying the optimum knowledge needed by a student to solve a given problem.

Hypermedia applications offer the possibility of gaining some measure, again indirectly, of students' procedural knowledge as they engage in problem solving. Programs such as HyperEquation allow us to track the solution pathways as students attempt to balance chemical equations. Because the program is non-linear, students are permitted to move freely from one problem to another without sequential constraint, skip from one level of difficulty to another, and loop back to modify earlier solutions. This capability approximates recursive learning in which students gain understanding from the problem-solving process which can then be applied to the process itself. Research in this area is still emerging, yet available findings suggest that computer applications can enhance our understanding of the nature of problem solving in science.

References

Browning, E., & Lehman, J. D. (1988). Identification of student misconceptions in genetics problem solving via computer program. *Journal of Research in Science Teaching, 25* (9), 747-761.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121-152.

Collins, A. (1990). Cognitive apprenticeship and instructional technology. In Idol, L., & Jones, F. B. (Eds.), *Educational values and cognitive instruction: Implications for reform.* Hillsdale, NJ: Erlbaum.

Collins, M. A. J. (1984). Improved learning with computerized tests. *The American Biology Teacher, 46* (3), 20-24.

Gronlund, N., & Linn, R. (1990). *Measurement and evaluation in teaching.* NY: MacMillan Publishing Co.

Helgeson, S. L., & Kumar, D. D. (1993). A review of educational technology in science assessment. *Journal of Computers in Mathematics and Science Teaching, 12* (3/4), 227-243.

Herb, A. (1992). Computer adaptive testing. *ADVANCE*, (Newsletter of the Medical Laboratory Professional). February 17, 1992.

Jackson, B. (1988). A comparison between computer-based and traditional assessment tests, and their effects on pupil learning and scoring. *School Science Review, 69* (249), 809-815.

Jacobson, R. L. (1993). New computer technique seen producing a revolution in educational testing. *The Chronicle of Higher Education*, September 15, A22-23, 26

Kumar, D. D., White, A. L., & Helgeson, S. L. (in press). A study of the effect of HyperCard and Pen-paper performance assessment methods on expert-novice chemistry problem solving. *Journal of Science Education and Technology.*

Kumar, D. D. (1993). Toward a technology for assessment. *Cognosos, 2* (2), 5.

Kumar, D. D. (1994). Hypermedia: A tool for alternative assessment? *Education and Training Technology International, 31* (1), 59-66.

Larkin, J. H., & Chabay, R. W. (1989). Research on teaching scientific thinking: Implications for computer-based instruction. In Resnick, L. B., & Klopfer, L. E. (Eds.), *Toward the thinking curriculum: Current cognitive research.* Association for Supervision and Curriculum Development.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Niaz, M., & Lawson, A. E. (1985). Balancing chemical equations: The role of developmental level and mental capacity. *Journal of Research in Science Teaching, 22* (1), 41-51.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27* (10), 937-949.

Pascual-Leone, J., & Goodman, D. (1979). Intelligence and experience: A neo-Piagetian approach. *Instructional Science, 8*, 301-367.

Schneiderman, B. (1987). *Designing the user interface.* New York: Addison Wesley.

Shavelson, R. J., Baxter, G. P., Pine, J., Yure, J., Goldman, S., & Smith, B. (1990, April). *Performance indicators for large-scale science assessment.* Paper presented at the annual meeting of the American Educational Research Association, Boston.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In Siegler, R. S. (Ed.), *Children thinking: What develops?* Hillsdale, NJ: Lawrence Erlbaum Associates, Boston.

Staver, J. R. (1986). The effects of problem format, number of independent variables, and their interaction on student performance on a control of variables reasoning problem. *Journal of Research in Science Teaching, 23* (6), 533-542.

Stewart, J. (1980). Techniques for assessing and representing information in cognitive structure. *Science Education, 64* (2), 223-235.

Weinstein, C. E., & Meyer, D. K. (1991). Implication of cognitive psychology for testing: Contributions from work in learning strategies. In Wittrock, M. C., & Baker, E. L. (Eds.), *Testing and cognition.* Englewood Cliffs, NJ: Prentice Hall.

Young, M. F., & Kulikowich, J. M. (1992, April). *Anchored instruction and anchored assessment: An ecological approach to measuring situated learning.* A paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Problem Based Macro Contexts in Science Instruction: Theoretical Basis, Design Issues, and the Development of Applications

Robert D. Sherwood, Anthony J. Petrosino, Xiaodong Lin, Mary Lamon & The Cognition and Technology Group at Vanderbilt[1]

Introduction

Our goal in this chapter is to outline a number of issues related to the development and application of a constructivist (cognitive-science based) model of instruction which we call *Anchored Instruction* and the implications this model may have on science education. We will do this by discussing some of the issues that have lead the group to this still evolving theory of instruction, provide examples of projects that have used the design principles, briefly summarize some of the research studies undertaken to support our work, and indicate some of the opportunities for teachers as they implement these designs.

Anchoring Instruction in Meaningful Contexts

In a paper appearing in the *American Psychologist* (Bransford, Sherwood, Vye, & Rieser, 1986) we proposed that a major goal of instruction is to allow students and teachers to experience the kinds of problems and opportunities that experts in various areas encounter (See also, Perfetto, Bransford, & Franks, 1983). Experts in an area have been immersed in phenomena and are often more familiar with how they have been thinking about them than are novices (e.g., Chi, Glaser, & Farr,

---

[1] The Cognition and Technology Group at Vanderbilt (CTGV) is an interdisciplinary team associated with the Vanderbilt Learning Technology Center, Box 45, Peabody College, Vanderbilt University, Nashville, TN 37203. Members of the CTGV who contributed to this article are John Bransford, Susan Goldman, Dan Hickey, Clifford Hofwolt, Teresa Secules and Dan Schwartz. Partial funding for various aspects of the research and development described in this paper were provided by The James S. McDonnell Foundation, NASA, and The National Science Foundation (Grant # ESI-9350510). Opinions expressed are those of the authors and not necessarily those of the funding agencies.

pter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for ific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

1991). When introduced to new theories, concepts and principles that are relevant to their areas of interest, the experts can experience the changes in their own thinking that these ideas afford (e.g., Dewey, 1933; Hanson, 1970; Schwab, 1960). For novices, however, the introduction of concepts and theories often seem like the mere introduction of new facts or mechanical procedures to be memorized. Because the novices have not been immersed in the phenomena being investigated, they are unable to experience the effects of the new information on their own noticing and understanding.

The anchored instruction approach represents an attempt to help students become actively engaged in learning by situating or anchoring instruction in interesting and realistic problem-solving environments. These environments are designed to invite the kinds of thinking that help students develop <u>general</u> skills and attitudes that contribute to effective problem solving, plus acquire <u>specific</u> concepts and principles that allow them to think effectively about particular domains (CTGV, 1992a; see also CTGV, 1990).

Anchored instruction environments share some of the characteristics of inquiry environments. These characteristics have been suggested as a model for science instruction since Schwab (1978). They are similar in that anchored instruction environments, as well as inquiry environments, do not propose to "directly" instruct students but provide a situation where learning can take place, often including many types of inquiry activities.

Anchored instruction also shares commonalities with a class of theories known as constructivist theories (e.g., Bransford & Vye, 1989; Clement, 1982; Duffy & Bednar, 1991; Minstrell, 1989; Perkins, 1991; Scardamalia & Bereiter, 1991; Schoenfeld, 1989). According to a constructivist perspective, knowledge is actively constructed by learners through interaction with their physical and social environments and through the reorganization of their own mental structures (Cobb, Yackel, & Wood, 1992; Wheatley, 1991; Brown, Collings, & Duguid, 1989). Similarly, anchored instruction emphasizes students' active engagement in realistic problem-solving environments that allow them to make modifications to their current understanding.

Instead of having teachers "transmit" information that students then "receive," these theorists place a great amount of importance on having students become actively involved in the construction of knowledge. For instance, constructivist theorists want to assist students to construct and coordinate effective problem representations through the use of physical and symbolic models, through reasoning and argumentation, and through deliberate application of problem-solving strategies (e.g., Bransford & Stein, 1993; Brown, Collins, & Duguid, 1989; Clement, 1982; Minstrell, 1989; Palincsar & Brown, 1989; Resnick

& Klopfer, 1989; Scardamalia & Bereiter, 1991; Schoenfeld, 1989). A fundamental assumption of the constructivist position is that students cannot learn to engage in successful knowledge building activities simply by passively being told new information (Bransford, Franks, Vye, & Sherwood, 1989). Rather, students need repeated opportunities to engage in sustained exploration, assessment, and revision of their ideas over extended periods of time and through a series of reiterations (CTGV, 1992b).

Initial discussions of anchored instruction can be found in CTGV, 1990 and CTGV, 1992c. There we note that the general idea of anchored instruction has a long history, and is related to ideas about project-based learning (Dewey, 1933), case-based learning (Gragg, 1940) and problem-based learning (e.g., Barrows, 1985; Williams, 1992). We also connect our use of anchors to the "situated cognition" arguments of Brown, Collins and Duguid (1989) (see Moore, Lin, Schwartz, Petrosino, Hickey, Campbell, Hmelo & CTGV, in press). We have studied the uses of anchored instruction in domains that focus on literacy (Kinzer, Risko, Goodman, McLarty & Carson, 1990; Kinzer, Risko, Vye & Sherwood, 1988; Bransford, Kinzer, Risko, Rowe & Vye, 1989); mathematics (CTGV, 1992a; CTGV, 1993a,b,c) and science (CTGV, 1992c; Goldman, Petrosino, Sherwood, Garrison, Hickey, Bransford & Pellegrino, in press; Sherwood, Petrosino, Goldman, Garrison, Hickey, Bransford, & Pellegrino, 1993). In several more recent chapters, we discuss some of the lessons learned about ways to increase the power of anchored instruction to produce flexible transfer (Barron, Vye, Zech, Schwartz, Bransford, Goldman, Pellegrino & Kantor, 1994; CTGV, 1994; CTGV, 1993a).

Design Principles for Anchored Instruction

We have developed a set of design principles for anchored instruction and have used these, with slight modifications, in a variety of domains, including science, mathematics, and literacy. Descriptions of the general set of design principles appear in several previously published papers (CTGV, 1991; 1992a; 1993; McLarty, Goodman, Risko, Kinzer, Vye, Rowe, & Carson, 1990). In the present context we provide a brief summary of these design principles. They are instantiated in the architecture of the *Scientists in Action* series, described in the next section.

There are seven design principles that guide our work in a variety of domains including mathematics and literacy (CTGV,1993a; CTGV, 1991). These design principles mutually influence one another and operate as a Gestalt rather than as a set of independent features of the materials. For example, the *narrative format*, the *generative design of the stories* and the fact that the adventures include *embedded data* make it possible for students to learn to generate problem-solving goals, find relevant information, and engage in reasoned decision making. The *complexity of the problems* helps students deal with this important aspect of problem solving and the *use of video* helps make the complexity manageable. The video format also makes it easier to embed the kinds of information that provide opportunities for *links across the curricula*. It is important for *pairs of episodes* to be developed to afford discussions about transfer of problem solving skills.

The design principles of anchored instruction are also consistent with a constructivist cognitive-science perspective. If students are going to actively make changes in their cognitive structures, we need to create environments that afford them the opportunities to do this. Using the design principles of anchored instruction, we have begun to provide these types of environments. Students need interesting and challenging problem situations that are embedded in familiar and engaging scenarios in order to facilitate the utilization of the knowledge they already have available. We believe this assists the learner in the creation of new and modified knowledge structures.

Early Empirical Work on Anchored Instruction in Science

The empirical roots of anchored instruction in science lie in earlier research that we did with a variety of scientific concepts, including density (Sherwood, Kinzer, Bransford, & Franks, 1987). This earlier work involved the use of commercially available movies to create problem solving environments that helped students understand the importance of science information. Students who learned the information in a problem solving context (Indiana Jones' trip to the South American jungle in *Raiders of the Lost Ark*) were better able to remember and to spontaneously use the information in new situations than were students who had learned the information with the intent to remember it. For example, in a transfer task students were asked to imagine planning a journey to a desert area in order to search for relics in Pueblo caves. Students who had simply read facts to remember them tended to be quite general in their responses. In contrast, students who had acquired the information in the problem solving condition were much more specific and gave greater evidence of being aware of various sets of constraints that they would need to consider. When discussing food, for example, most of them focused on the importance of its nutritional contents. When discussing water they emphasized the importance of calculating its weight.

Overall, students who received information in the context of problem-solving were much more likely to remember what they read and to spontaneously use it as a basis for creating new sets of plans. Similar effects were found with seventh and eighth grade students on knowledge of why it might be useful to know these concepts, and, to a lesser degree, on recall of science information (Sherwood, Kinzer, Hasselbring, & Bransford, 1987). These results suggested that an important goal for science materials would be to involve students in actual problem solving. These data, along with other data on the acquisition of science concepts from video-based materials (see for discussion, CTGV,1992c), convinced us of the potential for using video to create inquiry environments that can increase students' interest in science and facilitate their understanding of science concepts.

Design Modifications of the Scientists-in-Action Series

There are three features of the instantiation of the design principles for anchored instruction that are unique to the pilot episode of the *Scientist-in-Action* series (as contrasted with the instantiation of the design principles in our work in mathematics and literacy).

First, challenges are posed several times during the course of the story rather than at the end, as in the Jasper adventures. These "interruptions" to deal with a problem enable students to be a part of the problem solving process and to have multiple opportunities to work on the same problems that the scientists in the video are working on. When the video resumes, students can compare and contrast their solutions to what the scientists actually did. There is a tradeoff in this design as contrasted with the design of the Jasper adventures: the modification made in the *Scientist in Action* pilot provides fewer opportunities for students to gain experience formulating solutions to complex, ill-defined problems.

A second modification of the original anchored instruction design principles is that much of the data needed to solve the problem occur in ancillary materials rather than all of the data being embedded in the video, the design used in the Jasper series. These ancillary materials are authentic (network television news footage, United States Department of the Interior Geological Survey topographic maps and Department of Transportation manuals) and teachers are encouraged to help students conduct all or some of the laboratory tests in the classroom.

A third feature is that the *Scientist in Action* pilot was done with a very limited operating budget. In order to keep production costs reasonable, a decision was made to use only two sets and bring the "real world" in through a television set during the script's emergency news broadcasts. For the prototype, we became interested in whether this loss of production value had any impact on student engagement and whether we could show enough actual science to be helpful within such a limited format.

Initial Prototype

We were guided by these design principles when we developed the prototype of the *Scientists-in-Action* series, the episode *The Overturned Tanker.* The basic premise of the video is a "day-in-the-life" format: As the video opens, the viewer is welcomed to the office of Gina Davis (played by a white female), a hydrologist for the county. She begins to explain what a hydrologist does when a student intern (played by a white male about 22 years old) enters her office with some questions. A few moments later, they are interrupted by the hurried entry of Gina's secretary who is quite disturbed. It seems that a tanker truck has overturned on the highway and is spilling an unidentified chemical all over the highway. There is the additional danger that the liquid will spill into the river at the base of the highway; if it does it will be the hydrologist's responsibility to deal with the emergency. The tanker has a "dangerous chemical" sign on it but the driver is unconscious and the identity of the chemical is unknown. More information is provided about the properties of the liquid and they find out that the chemical is indeed running off into the river. At this point, the video fades to black and students are asked to hypothesize what the unidentified chemical might be, possible reactivity with water, and whether the chemical will flow toward the lake or toward the city (it flows downstream toward the city).

Students are provided with authentic materials — the same ones that real emergency teams use and that the student intern is given in the video — as they work on these problems. Once the students have finished, the video resumes and the students see how the hydrologist, the intern, and an additional member of the team, a fire chief, have attempted to solve the same problem. The team discusses the physical properties of the liquid, the precautionary measures needed to make to protect the population, and that the chemical is flowing toward the city and right past the town's water treatment plant. They enlist the help of a chemist (played by an African-American female) to determine the exact nature of the chemical. She is shown in her laboratory conducting a series of tests that identify the chemical as a highly toxic compound, phosphorus trichloride. Additional information about methods for preventing contamination of the town's water supply is provided and the video again fades to black as students are asked to help the team solve the problem of determining the best method of dealing with the spill. Again, after working on this problem, students return to the video to find out what the "experts" have figured out. There is a final problem posed that deals with determining the flow rate of the river and the time of day the phosphorus trichloride would be expected to reach the intake valves of the city's water treatment plant.

Prototype Pilot Research Results

After the original prototype episode was completed two sets of pilot studies were undertaken with students at various grade levels. Details of these studies are presented elsewhere (Goldman, et. al., in press; Sherwood, Petrosino, Goldman, Garrison, Bransford, & Pellegrino, 1993) but a summary is provided.

In a study with 5th grade students, a two group experimental design was used. One group saw only a seven minute segment of videotape compiled from network news footage that reported the spill of a pesticide into a California river system in 1991. The second group saw the network news footage followed by the prototype episode "The Overturned Tanker."

Two types of instruments were administered to both groups as pre- and posttests. First there were a set of "free response" content questions. These asked students (1) who would be involved in dealing with a chemical spill into a local river, (2) to list four or five steps that would need to be taken in the event of a spill, (3) what chemical could be used to make an acid spill less dangerous, (4) what hydrologists do, and (5) what fire chiefs do. Quantitative scores were assigned to responses.

Second, we administered a set of questions probing students' perceptions and attitudes towards science. Students were asked to respond on a 3 point scale (disagree-not sure-agree) to questions about where and how scientists work (e.g., in laboratories, alone, etc.), and whether they thought they (personally) could be a scientist. A final set of interest questions asked students to rate on a 4 point scale (not at all, a little, some, a lot) how much they wanted to learn more about what a variety of different kinds of scientists do. Hydrologists, chemists, biologists, and physicists were included in the list of scientists that were rated.

The results of the content questions indicate that the network news segments alone were not sufficient for students to acquire specific information about how scientists would deal with chemical spills. As expected students who viewed and solved the problems embedded in the *Overturned Tanker* had a more differentiated understanding of spills and the responsibilities of different kinds of scientists in an emergency situation such as a chemical spill.

For the attitude items results were somewhat mixed with both sets of students showing changes in their perception of scientists not only working in laboratories and some indication that the group that worked with both the news footage and the pilot was more interested in learning about some of the types of scientists in the videos.

In a second study, forty-nine adolescents enrolled in the ninth grade of a predominantly Hispanic, inner-city high school served as subjects. Materials used in our first study were duplicated for this study. In addition, transfer materials consisting of a topographic map and a set of questions relating to the topographic map were constructed for additional research. The study was conducted in three intact classrooms with a certified science teacher/researcher serving as the instructor over a three-day intervention. Three experimental conditions were carried out with the first two groups representing a replication of Experiment 1: (1) Overturned Tanker preceded by network news; (2) network news. The third group saw the network news and then the Overturned Tanker but did not solve the problems at each segment; instead this group watched the video "straight through."

The content questions, attitude, and interest questions were scored as described in the first experiment. As mentioned previously, students were administered a transfer task on topographic map reading after their respective treatments and posttesting were completed. The students were asked to identify characteristics such as slopes, elevations, and direction of potential water flow.

On the content questions, the results were quite similar to those observed in the middle school students who participated in our first study. This finding was encouraging since our population was different both in location and grade level. Changes in attitudes toward science tended to be weaker, perhaps due to the older ages of the students. We observed positive trends on the interest questions that were similar to those observed for the fifth graders. It was concluded that our initial findings in experiment one were validated by this successful replication. Furthermore, we were encouraged to see that our Scientists in Action intervention may be suitable for an older educational audience.

From the data we have collected thus far, it appears that multimedia environments can be effective in increasing specific content knowledge. We also want to pursue the effect multimedia has on transfer to scientific inquiry in new domains. We hope to examine the impact of generative activity on transfer, and learning, by continuing to pursue contrasts between watching a solution versus doing the solution.

In addition, while we were slightly embarrassed by the quality of the prototype and fully expected students to pan it, to date, we have found uniform acceptance of the video and have in fact been our own most severe critics of the production of the pilot episode. We suspect with the relative recent popularity of “real life” crime and emergency shows being broadcast nationally, a certain tolerance and authenticity is becoming associated with video of less than professionally polished quality.

Revisions of Design Principles for Planned Episodes

While our initial studies were encouraging, they also tended to show that our pilot episode was somewhat limited in the depth to which students became engaged in the problem solving process. The problems encountered by the students were identified in the video and they were asked to solve them. This lack of *generativity* in the design was a concern to the development team. Other issues such as the authenticity of the situation and the need to bring students into situations with more extended data analysis and hypothesis testing drew the development team to reconsider some of the design features mentioned previously and to modify them further. The current design reflects work in progress that has been used in the production of an episode during the late spring/summer of 1994. These design issues are still somewhat fluid and represent the thinking of the group at the time of the writing of this chapter (August, 1994).

The script for the first episode still reflects many of the design principles that were used in the pilot with some modifications. The first episode is a *narrative* with a *video format*, although the video format is being supplemented with more computer based materials such as digitized maps, data sets, simulations, and reference materials. While the problem(s) posed in the adventure offer students a chance to work within a problem space that is *complex*, they do not have answers that are based simply on one or two scientific principles. Data is still *embedded* in the video, however more data is available to the students from sources not directly in the video but referred to by the characters and utilized by the students to solve the overall problem. The data are accessible via CD-ROM, commercial database applications and over electronic networks. We are also experimenting with the inclusion of anomalous data (Chinn & Brewer, 1993) in order to facilitate the development of various critical thinking skills among middle school students. *Links across the curricula*, especially social issues related to science, have been established. The strength of the links has yet to be decided as is the *nature of the episode connections* since we have only completed one of a planned three episodes.

In order to discuss some of the more extensive modifications in the design principles being used in the first episode development, some indication of the nature of that episode is needed. As with the pilot, the first episode deals with the issue of water quality and the degradation of it by some action. However, it does have some major differences. The story line starts with a team of scientists and students at a local school conducting an extensive water quality survey of a local river system. Pairs of students and scientists conduct "electronic field trips" to various sites on the river and collect such data as chemical tests for dissolved oxygen, ammonia levels, pH, and temperature. In addition, biological surveys such as macro invertebrate sampling are performed to give a picture of overall water quality. Data, which may be in the form of raw numbers, video clips of various sections of the river, pictures of invertebrates, etc., are sent back to the "base" classroom through electronic means. The technology to do this in real life is not completely developed but may be available by the time the series is completed (1996).

Students who are using the program will have an opportunity to work in small groups, summarize raw data, develop graphs and charts for classroom discussion, and research the science behind the particular tests used in the video. In this manner, we believe they will develop a stronger understanding of the environmental system under study and be prepared to tackle the "incident" problem that will occur next in the video.

A noticeable change in the data collected from a new sample is an indication to the students that a change in the river system has occurred and they will need to collect additional data to compare to their baseline data to help isolate what has happened and to determine the source of the contamination. The story is designed so that the expected hypothesis generated by the students will be incorrect in terms of the source of the pollution. The video will then provide additional data that allows them to re-think their initial ideas to determine the actual source and cause of the problem. In this manner, we hope to allow students to experience some of the difficulties of doing science in natural systems; namely, that first hypotheses often need rejection or refinement.

As has been indicated, the use of both *embedded and external* data, the *electronic field trip* component, substantial work with *baseline data* and more *hypothesis testing* are all substantial modifications of the pilot episode and original design. These design changes along with the use of CD-ROM technology to replace the videodisc technology of our pilot and earlier "Jasper Woodbury" project will hopefully allow us to develop a product that will be both interesting and challenging to middle school science students.

The Role of Field Testing in Design and Development

We would like to illustrate how we integrate design and development processes with field experiments and testing in order to effectively modify design decisions. The script for the first episode, *The Stones River Mystery,* not only reflects the design principles discussed in the previous sections, but also takes into consideration our understanding of students' domain knowledge through field testing and review of existing literature. The goal is always to build on students' current understanding rather than simply prescribe instruction or learning activities designed to help students reach pre-set objectives by particular points in time. To do so, we held the research and design of the video concurrent so that our most timely understanding of students' domain knowledge and conceptual understanding can be incorporated in the development processes. In this way, we avoid simply repeating a scripted version of what worked someplace else and still keep the flexibility to make the design and development more situational and reflective of classroom needs (see Lin, Bransford, Hmelo, Kantor, Hickey, Secules, Petrosino, Goldman & CTGV, in press).

For example, throughout the design process we did several experiments with 6th and 7th graders on their knowledge about pollution. We designed various scenarios around concepts of dissolved oxygen, testing and cleaning pollution, etc. (e.g. Imagine you have an opened bottle of soda. If you don't open the soda, you can not see bubbles. However, when you open the bottle, we see a lot of bubbles trying to come out. Can you explain where do these bubbles come from?) to study how students understand those concepts. We found that most students did not understand the notion of dissolved oxygen and how dissolved oxygen is produced. When this finding was reported to the design team, we immediately revised the existing episode to add more situations where dissolved oxygen can be further explored and discussed. For instance, we added one incident to the script where one of the students in the Scientists In Action team accidentally slipped into the river. He asks the scientists in the team what algae is good for except making people slip. We took this opportunity to explain how the dissolved oxygen can be produced by having the field biologist explain that "when the sun shines on the algae, it produces oxygen that fish and invertebrates need." Further discussions around the notion were evoked by the incident among the students and the scientists.

In addition to our concerns over the domain knowledge acquisition in the design process, we also gave considerable attention to how students reacted to the materials, acting and the learning environments created in the video. We continued the field experiments even after the video was produced so that further revisions could be made according to the suggestions and reactions from teachers and students. These experiments helped guide our post-production editing. Additionally, we synthesized the results from our field experiments, our review of existing literature and the suggestions from teachers and students to further design the hands-on laboratory experiments and teacher guidelines to go with the episode. The upcoming research on the episode will further focus on the concerns and issues that have been raised throughout the design and development processes.

Other Projects Related to Science Using a Similar Design

Mission to Mars

This prototype learning environment (Hickey, Petrosino, Pellegrino, Bransford, Goldman & Sherwood, in press) is designed to lead students to generate problems about the scientific challenge of planning a human expedition to the planet Mars, and then support student inquiry into solving these problems. We feel that the Mars mission is an excellent problem space because it lends itself to subproblems from every academic domain, thus making it inherently cross-curricula. Current cognitive theories of learning are leading many to reconsider guided discovery learning with the assumption that cognitive and motivational processes are strongly bound to specific domains and content (Linn, 1986). Like others (e.g., Blumenfeld, et. al., 1991; Hawkins & Pea, 1987) we feel that the thoughtful application of these advances and the incorporation of emerging information technologies should allow us to build upon some advances made by earlier discovery learning programs (see Bredderman's 1983 meta-analysis). The Mission to Mars group is currently using a three-level conceptual model for our Mars Challenge (see Figure 1). The first level, *Problem Generation*, concentrates on students' problem posing, defining, and categorizing of problems within the problem space. The second level, *Knowledge Distribution & Teamwork*, centers on the cooperative environment based on individual efforts directed toward the larger groups' common goal. It is at this level where we attempt to model both the development of expertise in specific content areas and the collaborative nature of the broader research community. The third level, *Using Knowledge Tools*, concentrates on the educational and reference resources students will use to solve their self-selected problems. The efforts have focused on problem generation. This phase has the crucial role of defining the inquiry domain and developing the degree of interest needed to motivate students during subsequent self-directed inquiry and learning. Except for mathematical problem solving (e.g., Brown and Walters, 1991) and studies of creativity, (e.g., Dillon, 1983; Getzels, 1979), problem generation has received little attention. To facilitate problem generation, the group has developed a seven-minute video using existing NASA footage. The Mars Mission Challenge video visually suggests the wide variety of factors involved in planning and carrying out a human mission to the planet Mars. The video narration explicitly challenges students to pose problems within the domain of planning a mission, but is not so specific that it preempts problems that we expect students to pose. The program is recorded on videodisc and navigated with a HyperCard controller.

Over a series of pilot studies with middle school classrooms, the problem generation activity has evolved into several discrete steps that take place over three class periods. During the first period, the students view the Mars Mission Challenge videodisc, then individually pose as many problems as they can. The class then breaks into small groups and take turns reviewing the video and posing additional problems. Each group then aggregates problems across individuals to create a set for the entire group. During the second period, each group sorts the problems into self- specified categories. Finally, during the third period, the class regroups around the master categories according to individual student preferences; these groups then collect all of the problems in their master category. Students in each of these groups then use these problems to synthesize a prioritized set of problems for their category. One study (Hickey and Petrosino, 1992) examined the range of problems posed during the first activity period. They analyzed 319 problems posed by 11 small groups of students and categorized these problems because of

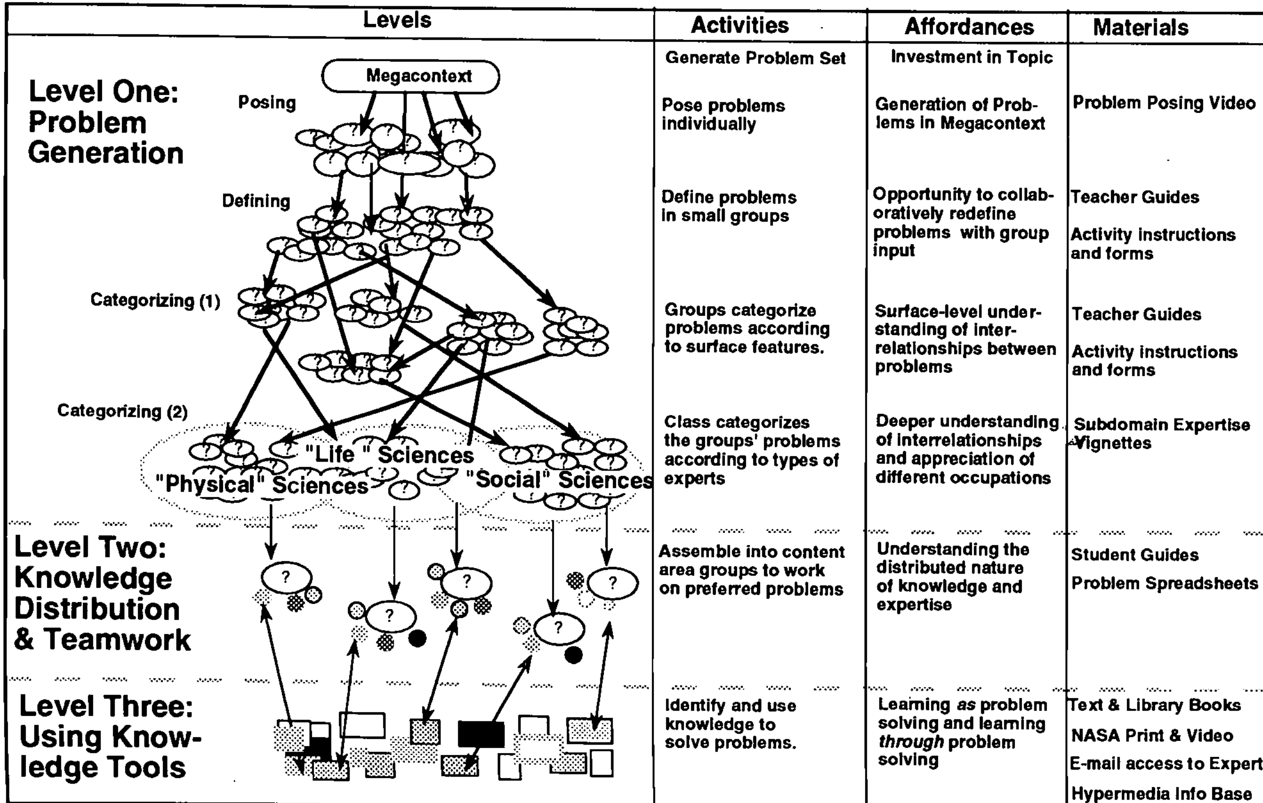


Fig. 1 *Mars Mission Challenge* Conceptual Framework

their anticipated role as anchors for instruction. Based on the materials available, we decided it would be reasonable to support inquiry in several broad content areas (Technology, Trip Route, Health, Life Support, Policy, and Surface Activities). We believe these broad content areas map over well with typical subjects taught in the middle school curriculum, namely computers, math, biology, history and geography.

Another pilot study with low/average-ability sixth graders allowed them to probe the categorization process. Given the students' limited domain knowledge and the novelty of the task, results were satisfactory. Six main categories emerged from this pilot, Ship and Transportation, Feelings about Mars, Landforms and Lifeforms, Safety, Fuel and Necessities. We recognize opportunities for improvement of this process. For example, students' difficulties in defining their own categories have led our group to restructure the categorization activities around predetermined subdomains based on different types of scientific expertise.

While the idealized environment that are proposed represent a fundamental reshaping of existing educational practices, these activities and materials are functioning well in existing classrooms (see the <u>Schools For Thought</u> section below). Meanwhile, we are considering ways of formally studying this environment. In particular, controlled studies of different activity formats, and the impact of problem generation on student knowledge, interests, and subsequent academic performance are being planned.

From a technological standpoint, the developers are considering reformatting the Mars Mission Challenge video into a QuickTime application and incorporating recent advances in CD-ROM technology to allow for small group (1-2 students) learning environments with virtual "experts" at the disposal of the individual learner and teaching professional. However, we still wish to maintain the collaborative nature of this activity and move toward more learning environments that utilize the principle of distributed expertise.

<u>The Schools For Thought Project</u>

In a project currently funded by the James S. McDonnell Foundation, three groups of researchers are cooperating with sites in three diverse geographic locations, to test an integration of projects that have showed potential for instruction. The three projects being integrated are:

1. <u>Fostering Communities of Learners</u>, developed by a team at the University of California at Berkeley (Brown & Campione, 1994; Brown, Rutherford, Nakagawa, Gordon, & Campione,

1993; Brown, 1992; Brown & Campione, 1990).

2. The CSILE project (Computer Supported Intentional Learning Environments) developed at the Ontario Institute for Studies in Education (Scardamalia, Bereiter, & Lamon, 1994; Scardamalia, Bereiter, Brett, Burtis, Calhoun & Smith, 1992; Scardamalia, Bereiter, McLean, Swallow, & Woodruff, 1989).

3. The Jasper Woodbury Project previously mentioned in this chapter (CTGV, 1990; 1992b; 1993a,b).

The joint experiences of our consortium have taught us that the materials and cognitive tools that we create as anchors play a very valuable role in promoting learning communities. By becoming the focal point of authentic discussion, inquiry, and debate, the anchors and tools help establish a cognitive/social environment that is stimulating for students and assists them in the participation as both experts and learners in activities that occur both in and out of the traditional classroom environment (Moore et. al., in press).

A more extensive report on the integration project can be found in Lamon (1993) but some of the aspects of each project and their interconnection are provided here. In Fostering Communities of Learners the developers take the position that disciplines are based upon sets of "deep principles" that are important for students to discover through in-depth research on particular problems or situations within that discipline. It is important to mention that effective learning communities are not simply "discovery" environments. Rather, there is a great amount of structure that is needed in order for the model to work optimally. For example, in a unit on biology, students are first given a "benchmark" lesson on a curriculum theme (e.g., changing populations). From this lesson, students will generate as many questions as they can think of (usually 100 or more are produced). The teacher and students categorize these questions into approximately five subtopics (e.g., extinct, endangered, artificial, assisted, and urbanized populations). About six students form a research group, each group takes responsibility for one of the five or so subtopics.

Using texts, magazines, newspapers, video, and electronic mail consultations with outside experts, students write up summaries of what they are learning. As well, students engage in small discussions of articles and texts relevant to the overall theme. Discussions are structured along the lines of four key strategic activities, summarizing, clarifying, questioning and predicting. These activities originated in a reading comprehension program called Reciprocal Teaching (Palinscar & Brown, 1984) in which students take turns leading the discussion.

In addition, the program uses a modified version of the Jigsaw method (Aronson, 1978) of cooperative learning. As students prepare preliminary drafts of their reports, they engage in "crosstalk" sessions (a whole class activity where the groups periodically summarize where they are in their research) and get input from other groups (see also Cohen, 1994). In the case of our Mission to Mars unit, students generated a number of interesting drawings attesting to their growing expertise in the domain.

Discussion between and within groups is greatly facilitated by the CSILE computer environment. The physical CSILE environment is a computer network of at least eight computers connected to a file server on which the CSILE software resides. The core of CSILE is a communal student generated database which encourages students to articulate their theories and questions, to explore and compare different perspectives, and to reflect on their joint understanding. In CSILE, students work individually and collaboratively, commenting and building upon one another's understandings. Anyone can add a comment to a note or attached a graphic note (e.g., picture, diagram) to another note, but only authors can edit or delete notes.

The Jasper Woodbury series, with its emphasis on complex problem solving in mathematics, complements the other two projects by providing topic areas for discussions between and with groups in the classroom. Students have the opportunity to discuss possible solutions and areas of understanding (or misunderstanding) through the CSILE network.

Three sites: Ontario, Canada; Nashville, Tennessee; and Oakland, California are attempting to use all three methods within classrooms during the 93-94 school year. As this is an experiment in progress, overall results are not yet available but site coordinators report several interesting changes in classroom environments as the students become more familiar with the methods under study. In the Nashville site, one of the first units being undertaken is based upon the previously described Mars Mission materials, which has been well received both by students and participating teachers. At the same time, students were doing Jasper units involving themes of flight and trip planning, as well as deep principles of optimization and rate. Using CSILE discussion notes they were studying Roman colonization in social studies as well as themes of social interdependence of people in literature and relating both of these to the Mars Mission. In this way the three projects, once separate, now contribute to the development of truly integrated curriculum (CTGV, 1993c).

Implications for Teachers using the Schools For Thought Design

Scardamalia & Bereiter (1991) in discussing knowledge building in children postulate that there are three different types of teachers in current instructional settings. The type "A" teacher holds a "task" model. This model fits the current style of many classrooms with a focus on seatwork (tasks) to be completed by the student. The type "B" in contrast holds a "knowledge-based" model. Scardamalia & Bereiter describe this model as: "The focus tends to be on understanding, and the teacher's role includes setting cognitive goals, activating prior knowledge, asking stimulating and leading questions, directing inquiry, and monitoring comprehension" (p.39). They note this model is one that is often taught as a model of exemplary teaching and is often found in teachers guide. There reservations about the model center on the fact that the teacher is very much in "control" of the learning process. Their model "C" teacher exhibits the characteristics of the model "B" with the addition of trying to turn over to the students the higher-level processes that the teacher would control in model "B." They note; "Thus, there is a concern with helping student to formulate their own goals, do their own activating of prior knowledge, ask their own questions, direct their own inquiry, and do their own monitoring of comprehension" (p. 39). They note that this model is followed in the reciprocal teaching model they have described and researched (Brown & Palincsar, 1989; Palincsar & Brown, 1984). This type "C" teacher is also consistent with Wheatley's (1991) call for the teacher to view the learner "more as a growing tree than a sponge" (p. 19).

In order for the teacher to effectively use the materials proposed in this chapter several things need to occur. First, the teacher's view of instruction must undergo some modification if they are in a type "A" mode and to some extent from the type "B". For some teachers this will be a very large step, for others it will seem a natural progression. Secondly, our work with the "Jasper" series (CTGV, in press) has pointed out that there are several "special challenges" that needed to be attended to in working with the "Jasper" materials. These include:

1. Allowing students to pursue "wrong" pathways that lead to alternative solutions that may not be optimum.

2. Knowing when to assist students versus letting them "struggle" a little longer.

3. Where to put alternative materials such as "Jasper" (and the ones currently under development) in the "regular" curriculum (see also Hofwolt, 1992).

We also realize that many of our teachers may need additional subject knowledge support. Our experiences indicate that this support is particularly beneficial at the lower grade levels (fifth and sixth) of our target grades (fifth through eighth). In some "Jasper" episodes teachers have reported difficulty with some of the statistics and geometry concepts. This is not surprising, given that they may not have taught these particular concepts and their instruction in them may have been several years ago.

We are attempting to address this issue through both video and print materials to assist teachers to become more comfortable with the mathematics and science that are involved in the episodes.

Concluding Remarks

While the projects described in this chapter are all relatively new and evaluative studies are only at the preliminary stages, it would appear that they offer promise for science instruction. They require school environments to be structured differently with both the teacher and student taking on roles that are not, in may ways, traditional. Teachers will need to be more like guides of instruction, helping students develop their understandings. Students will need to be able to act as active pursuers of knowledge not just passive acceptors of information.

References

Aronson, E. (1978). *The jigsaw classroom.* Beverly Hills, CA: Sage.

Barron, B., Vye, N. J., Zech, L. Schwartz, D., Bransford, J. D., Goldman, S. R., Pellegrino, J., Morris, J., Garrison, S., & Kantor, R. (1994). Creating contexts for community based problem solving: The Jasper challenge series. To appear in C. Hedley, P. Antonacci, & M. Rabinowitz (Eds.), *Thinking and literacy: The mind at work.* Hillsdale, NJ: Erlbaum.

Barrows, H. S. (1985). *How to design a problem-based curriculum for the preclinical years.* New York: Springer.

Blumenfeld, P.C., Soloway, E., Marx, R. W., Krajcik, J.S.,Guzdial, M., & Palincsar, A. (1991). Motivating project-base learning: sustaining the doing, supporting the learning. *Educational Psychologist, 26* (3&4), 369-398.

Bransford, J., Franks, J. J., Vye, N. J., & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470-497). New York: Cambridge University Press.

Bransford, J., Kinzer, C., Risko, V., Rowe, D., & Vye, N. (1989). Designing invitations to thinking:Some initial thoughts. In S. McCormick, J. Zutrell, P. Scharer, & P. O.Keefe (Eds.),*Cognitive and social perspectives for literacy research and Instruction.* Chicago, IL: National Reading Conference, 35-54.

Bransford, J. D., Sherwood, R. D., Vye, N., Rieser, J. (1986). Teaching thinking and problem solving. *American Psychologist, 41* , 1078-1089.

Bransford, J., & Stein, B. S. (1993). *The IDEAL problem solve*r (2nd ed.). New York: Freeman.

Bransford, J. D. & Vye, N. J. (1989). A perspective on cognitive research and its implications for instruction. In L. Resnick & L.E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research* (pp. 173-205). Alexandria, VA: ASCD.

Bransford, J. D., Vye, N., Kinzer, C., & Risko, V. (1990). Teaching thinking and content knowledge: Toward an integrated approach. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction: Implications for educational reform* (Vol. 1, pp. 381-413). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bredderman, T. (1983). Effects of activity-based elementary science on student outcomes: A qualitative synthesis. *Review of Educational Research*, *53*, 499-518.

Brown, A. L., Ash, D., Rutherford, M., Nakagawa, K., Gordon, A., & Campione, J. C. (1993). Distributed expertise in the classroom. In G. Salomon (Ed.), *Distributed Cognition*, New York: Cambridge University Press.

Brown, A. L. & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.) *Classroom lessons: Integrating cognitive theory and classroom practice*. Cambridge, MA: MIT Press/Bradford Books.

Brown, A. L. & Campione, J. C. (1990). Communities of learning and thinking or a context by any other name. *Human Development, 21*, 108-125.

Brown, A. L. & Palincsar, A. S. (1989). Guided cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.) *Knowing, learning and instruction: essays in honor of Robert Glaser* (pp. 393-451). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Brown, A. L., Ash, D., Rutherford, M., Nakagawa, K., Gordon, A. & Campione, J. C. (1993). Distributed expertise in the classroom. In G. Salomon (Ed.), *Distributed cognitions:Psychological and educational considerations.* (pp.188-228). New York, New York: Cambridge University Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Brown, S. I., & Walters, M. I. (1991). *The art of problem posing*. Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Glaser, R., & Farr, M. (1991). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Review of Educational Research, 63*(1), 1-49.

Clement, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception. *Journal of Research in Mathematics Education, 13*, 16-30.

Cognition and Technology Group at Vanderbilt. (1994). From visual word problems to learning communities: Changing conceptions of cognitive research. In McGilly, K. (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice*. Cambridge, MA: MIT Press/Bradford Books.

Cognition and Technology Group at Vanderbilt. (1993a). Anchored instruction and situated cognition revisited. *Educational Technology, 33*, 52-70.

Cognition and Technology Group at Vanderbilt. (1993b). The Jasper experiment: Using video to furnish real-world problem-solving contexts. *Arithmetic Teacher*, April, 1-5.

Cognition and Technology Group at Vanderbilt. (1993c). Toward integrated curricula: Possibilities from anchored instruction. In M. Rabinowitz (Ed.), *Cognitive Science Foundations of Instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cognition and Technology Group at Vanderbilt. (1992a). The Jasper experiment: An exploration of issues in learning and instructional design. In M. Hannafin & S. Hooper (Eds.), *Educational Technology Research and Development, 40*(1), 65-80.

Cognition and Technology Group at Vanderbilt. (1992b). The Jasper series as an example of anchored instruction: theory, program description, and assessment data. *Educational Psychologist, 27*, 291-315.

Cognition and Technology Group at Vanderbilt. (1992c). Anchored instruction in science and mathematics: Theoretical basis, developmental projects, and initial research findings. In R. Duschl & R. Hamilton (Eds.), *Philosophy of science, cognitive psychology and educational theory and practice*. NY: SUNY Press.

Cognition and Technology Group at Vanderbilt. (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher, 19*, 2-10.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64* (1), 1-35.

Dewey, S. (1933). *How we think: Restatement of the relation of reflective thinking to the educative process*. Boston: Heath.

Dillon, J. T. (1983). Problem finding and problem solving. *Journal of Creative Behavior, 16*, 97-111.

Duffy, T. M., & Bednar, A. K. (1991). Attempting to come to grips with alternative perspectives. *Educational Technology, 31*, 12-15.

Getzels, J. W. (1979). Problem finding: a theoretical note. *Cognitive Science, 3*, 167-172.

Goldman, S. R., Petrosino, A. J., Sherwood, R. D., Garrison, S., Hickey, D., Bransford, J. D., & Pellegrino. (in press). Anchoring science instruction in multimedia learning environments. In S. Vosniadou, E. De Corte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the psychological foundations of technology-based learning environments*. Hillsdale, NJ: Erlbaum.

Gragg, C. I. (1940, October 19). Because wisdom can't be told. *Harvard Alumni Bulletin*, 78-84.

Hanson, N. R. (1970). A picture theory of theory meaning. In R. G. Colodny (Ed.), *The nature and function of scientific theories* (pp. 233-274). Pittsburgh: University of Pittsburgh Press.

Hawkins, J. & Pea, R.D. (1987). Tools for bridging the cultures of everyday and scientific thinking. *Journal of Research in Science Teaching, 24*, 291- 307.

Hickey, D. T. & Petrosino, A. J. (1992, November). *Effects of generative video on students' scientific problem posing*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Knoxville, Tenn.

Hickey, D. T., Petrosino, A. J., Pellegrino, J. W., Bransford, J. D., Goldman, S. R., & Sherwood, R. D. (in press). The Mars mission challenge environment: A generative problem-solving school science environment. In Vosniadou, S. (Ed.) *Psychological and educational foundations of technology-based learning environments* (NATO ASO Series). Berlin: Springer-Verlag.

Hofwolt, C. (1992, March). *HyperScience for Middle School Students.* A paper presented at the annual meeting of the National Science Teachers Association, Boston, MA.

Kinzer, C. K., Risko, V. J., Goodman, J., McLarty, K., & Carson, J. (1990). *A study of teachers using videodisc anchors in literacy instruction.* Paper presented at the 40th annual meeting of the National Reading Conference, Miami, Fl.

Kinzer, C. K., Risko, V., Vye, N. J., & Sherwood, R. D. (1988). *Macrocontexts for enhancing instruction..* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Lamon, M. (1993). *Cognitive studies for restructuring middle school education.* Report of the St. Louis Science Center/St. Louis Public Schools Middle School Curriculum Collaborative, St. Louis, MO.

Lin, X. L., Bransford, J. D., Hmelo, C. E., Kantor, R. J., Hickey, D. T., Secules, T., Petrosino, A., Goldman, S. R. & CTGV (in press). Instructional design and development of learning communities: An invitation to a dialogue. *Educational Technology.*

Linn, M.C. (1986). Science. In R.F.Dillon and R. J. Sternberg (Eds.) *Cognition and Instruction.* (pp. 155-204). Orlando, Fl: Academic Press .

McLarty, K., Goodman, J., Risko, V., Kinzer, C. K., Vye, N., Rowe, D., & Carson, J. (1990). Implementing anchored instruction: Guiding principles for curriculum development. In J. Zutell & S. McCormick (Eds.), *Literacy theory and research: Analyses from multiple paradigms* (pp. 109-120). Chicago, IL: National Reading Conference.

Minstrell, J. A. (1989). Teaching science for understanding. In L.B. Resnick & L.E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research* (pp. 129-149). Alexandria, VA: ASCD.

Moore, J. L., Lin, X., Schwartz, D. L., Petrosino, A. J., Hickey, D. T., Campbell, O., Hmelo, C. & CTGV (in press). The relationship between situated cognition and anchored instruction: A response to trip. To appear in *Educational Technology.* Also to appear in H. McLellan (Ed.), *Perspectives on situated learning.* Englewood Cliffs, NJ: Educ. Technology Publishers.

Palinscar, A.S. & Brown, A.L. (1984). Reciprocal teaching of comprehension-fostering and comprehension monitoring activities. *Cognition and Instruction, 1,* 117-175

Perfetto, G. A., Bransford, J. D., Franks, J. J. (1983). Constraints on access in a problem solving context. *Memory and Cognition, 11,* 24-31.

Perkins, D. N. (1991). What constructivism demands of the learner. *Educational Technology, 31,* 19-21.

Rutherford, F. J., & Ahlgren, A. (1990). *Science for all Americans.* New York: Oxford University Press.

Scardamalia, M., Bereiter, C., & Lamon, M. (1994). Bringing the classroom into World III. In K. McGilly (Ed.) *Classroom lessons: Integrating cognitive theory and classroom practice.* Cambridge, Ma: MIT Press/Bradford Books.

Scardamalia, M., Bereiter, C., Brett, C., Burtis, P. J., Calhoun, C., & Smith L. N. (1992). Educational applications of a networked communal database. *Interactive Learning Environments, 2*(1), 45-71.

Scardamalia, M. & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *Journal of the Learning Sciences, 1*(1), 37-68.

Scardamalia, M., Bereiter, C., McLean, R. S., Swallow, J., & Woodruff, E. (1989). Computer supported intentional learning environments. *Journal of Educational Computing Research, 5,* 51-68.

Schoenfeld, A. H. (1989). Teaching mathematical thinking and problem solving. In L. B. Resnick & L. E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research* (pp. 83-103). Alexandria, VA: ASCD.

Schwab, J. J. (1978). *Science, Curriculum, and Liberal Education.* Chicago: The University of Chicago Press.

Schwab, J. J. (1960). What do scientists do? *Behavioral Science; 5,* 1-27.

Sherwood, R. D., Petrosino, A. J., Goldman, S. R., Garrison, S., Bransford, J. D., and Pellegrino, J. W. (1993, April). *An experimental study of a multimedia instructional environment in a science classroom.* Paper presented at the meeting of the American Educational Research Association, Atlanta.

Sherwood, R. D., Kinzer, C. K., Bransford, J. D., & Franks, J. J. (1987). Some benefits of creating macro-contexts for science instruction: Initial findings. *Journal of Research in Science Teaching, 24*, 417-435.

Sherwood, R., Kinzer, C., Hasselbring, T., & Bransford, J. (1987). Macro-Contexts for learning: Initial findings and issues. *Journal of Applied Cognition, 1*, 93-108.

Wheatley, G. H. (1991). Constructivist perspectives on science and mathematics learning. *Science Education, 75*(1), 9-21.

Williams, S. M. (1992). Putting case-based instruction into context: Examples from legal, and medical education. *Journal of the Learning Sciences , 2*, 367-427.

Analogies: Conceptual Tools for Problem Solving and Science Instruction

Shawn M. Glynn
University of Georgia

Reinders Duit
Institute for Science Education (IPN)
at the University of Kiel, Germany

Bruce K. Britton
University of Georgia

Introduction

In this chapter, we will examine how science teachers and textbook authors can help students to use analogies when solving problems. In the first part of this chapter, we will review the role of analogies in the history of science. Analogies have long served as important conceptual tools in problem solving and explanation; they have been influential both in thought experiments and actual experiments. We will consider how an analogy can be used to explain and even predict some aspects of a target concept or problem, but at some point every analogy breaks down and that is where misconceptions begin. We will point out that analogies play an important role not just in problem solving, but in the identification of new problems, the generation of hypotheses, and the explanation of solutions as well. Finally, we will recommend that science teachers train their students to strategically use a model of analogical problem solving. With such a model, students can plan, monitor, evaluate, and improve their problem-solving efforts.

Our view of analogical problem solving is based on recent findings in cognitive science concerning students' cognitive architecture (e.g., Anderson, 1989, 1990; Newell, 1990), problem solving (e.g., Anderson, 1993; Carey & Smith, 1993; Kaplan & Simon, 1990; Lavoie, 1993), development of expertise (e.g., Ericsson & Charness, 1994; Schmidt & Boshuizen, 1993), conceptual models (e.g., Anderson, Boyle, Corbett, & Lewis, 1990; White, in press), and persistent misconceptions (e.g., Duit, 1991a; Griffiths & Preston, 1992). In our view, students learn science meaningfully when they activate their existing knowledge, relate it to educational experiences, and construct new knowledge (Glynn & Duit, in press; Glynn, Yeany, & Britton, 1991). Analogies play an important role in the process of relating existing knowledge to new experiences (Glynn, 1991; Glynn, Duit, & Thiele, in press).

apter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for tific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

Role of Analogies in the History of Science

Analogies are valuable in actual experiments and the "thought experiments" that precede them. A thought experiment involves imagining experimental manipulations under simplified conditions that may not be attainable in the laboratory--at a given point in time. Thought experiments often reveal a phenomena or principle that is hidden by variables that cannot be practically removed or manipulated. Thought experiments, like actual experiments, play an important role in science instruction and famous scientists, such as Ernst Mach (1838-1916), often have spoken to this point: "Experimenting in thought is important not only for the professional inquirer, but also for mental development" (1905, reprinted 1975, p. 143).

Analogies played an important role in the thought experiments of Galileo Galilei (1564-1642). The conventional belief in his time was that heavier objects always fall faster than lighter ones. By means of an insightful thought experiment, Galileo argued that objects of different mass should fall (in a vacuum) with the same acceleration. Galileo's thought experiment was based on an analogy. He reasoned that if three identical blocks of iron--A, B, and C--were simultaneously dropped from the same height, they would fall and hit the ground together. This case is analogous, he believed, to another in which A and B are held together by a weightless chain and simultaneously dropped with C. In this second case, A-B would have double the mass of C, but A-B and C would still fall and hit the ground together. When the vacuum pump was eventually invented, Galileo's argument was validated: objects of different mass fell at the same rate in a glass cylinder and hit the bottom at the same time.

Johannes Kepler (1571-1630) often drew on analogies in his scientific thinking: "And I cherish more than anything else the Analogies, my most trustworthy masters. They know all the secrets of Nature..." (quoted in Polya, 1973, p. 12). Kepler believed that planetary motion, analogous to clockwork, could be explained by physical laws:

> I am much occupied with the investigation of the physical causes. My aim in this is to show that the celestial machine is to be likened not to a divine organism but rather to a clockwork...insofar as nearly all the manifold movements are carried out by means of a single, quite simple magnetic force, as in the case of a clockwork, all motions are caused by a simple weight. Moreover, I show how this physical conception is to be presented through calculation and geometry. (Quoted in F. J. Rutherford, Holton, & Watson, 1975, Unit 2, p. 68).

Benjamin Franklin (1706-1790), who did many experiments with electrical phenomena, drew on an analogy to explain them. In Franklin's view, an object acquired an electric charge by transferring an "electric fluid" that was present in all matter. According to him, when two materials, such as glass and silk, were rubbed together, some electric fluid passed from one into the other. One material then had an excess of fluid, while the other had a deficiency. The excess resulted in one type of charge, called "positive" by Franklin, while the deficiency resulted in another type of "negative" charge.

Joseph Priestley (1773-1804), a friend and colleague of Franklin's, took Franklin's ideas about electricity a step further. Priestley, reasoning by analogy, proposed the "law of electrical force," which was later experimentally tested by Charles Coulomb. Priestley arrived at the law in the following way:

> Priestley verified Franklin's results, and went on to reach a brilliant conclusion from them. He remembered from Newton's *Principia* that gravitational forces behave in a similar way. Inside a hollow planet, the net gravitational force on an object (the sum of all the forces exerted by all parts of the planet) would be exactly zero. This result also follows mathematically from the law that the gravitational force between any two individual pieces of matter is inversely proportional to the square of the distance between them. Priestley therefore proposed that forces exerted by charges vary inversely as the square of the distance, just as do forces exerted by massive bodies...We call the force exerted between bodies owing to the fact that they are charged 'electric' force, just as we call the force between uncharged bodies 'gravitational' force....
>
> Priestley's proposal was based on reasoning by analogy, that is, by reasoning from a parallel, well demonstrated case. Such reasoning alone could not *prove* that electrical forces are inversely proportional to the square of the distance between charges. But it strongly encouraged other physicists to test Priestley's hypothesis by experiment. (F. J. Rutherford et al., 1975, Unit 4, p. 35)

Shortly after Wilhelm Röntgen (1845-1923) discovered x-rays in 1895, Joseph John Thomson (1856-1940) found that when the rays pass through a gas they make it a conductor of electricity. J. J. Thomson (quoted in F. J. Rutherford et al., 1975, Unit 5, p. 52) drew an analogy to explain this effect as "a kind of electrolysis, the molecule being split up, or nearly split up by the Röntgen rays." Electrons were knocked loose from molecules and atoms in the gas. These molecules and electrons acquired a positive charge and were called "ions" because were are similar to the positive ions in electrolysis.

J. J. Thomson again drew an analogy when he proposed a popular model for the atom in 1904. According to him, an atom could be thought of as a spherical pudding of positive electricity, with the negative electricity embedded in it like raisins. Thomson's model proved useful until it was replaced by in 1911 by the more sophisticated model of Niels Bohr (1885-1962). Bohr also drew an analogy, with the nucleus compared to a planet and the electrons revolving like satellites in orbits about the nucleus.

The breakthrough that eventually led to Bohr's model was made by Hans Geiger, an assistant to Ernest Rutherford (1871-1937). In contrast to the predictions of the Thomson model, charged particles that were aimed at a thin foil were found to bounce back in unexpectedly large numbers. E. Rutherford drew an analogy when explaining this discovery:

> It was quite the most incredible event that has ever happened to me in my life. It was almost as incredible as if you fired a 15-inch shell at a piece of tissue paper and it came back and hit you. ...It was then that I had the idea of an atom with a minute massive center, carrying a charge. (quoted in F. J. Rutherford et al., 1975, Unit 5, p. 69).

Sheldon Glashow, who was awarded a Nobel Prize in 1979 for his work in nuclear physics, often mentioned the important role that analogies played in his development of a theory of unified weak and electromagnetic force interactions:

> I was led to the group SU(2) x U(1) by analogy with the approximate isospin-hypercharge group which characterizes strong interactions....Part of the motivation for introducing a fourth quark was based on our mistaken notions of hadron spectroscopy. But we also wished to enforce an analogy between the weak leptonic current and the weak hadronic current. (Glashow, 1980, p. 1320)

As these examples from the history of science indicate, analogies have long served as important conceptual tools in problem solving and explanation. Analogies have been influential both in thought experiments and actual experiments. Perhaps one of the most powerful and eloquent testimonials to the value of analogies in scientific thinking was that given by Robert Oppenheimer (1956):

> Analogy is indeed an indispensable and inevitable tool for scientific progress....Whether or not we talk of discovery or of invention, analogy is inevitable in human thought, because we come to new things in science with what equipment we have, which is how we have learned to think, and above all how we have learned to think about the relatedness of things. (pp. 129-130)

Defining an Analogy

An *analogy* is a correspondence in some respects between concepts (or problems) that are otherwise dissimilar. More precisely, it is a mapping between similar features of those concepts. When an analogy is drawn, ideas are transferred from a familiar concept to an unfamiliar one. The familiar concept is called the "analog" and the unfamiliar one the "target." Both the analog and the target have "features" (also called "attributes"). If the analog and the target share common or similar features, an analogy can be drawn between them. An abstract representation of an analogy, with its constituent parts can be represented in the following way:

```
                    Superordinate Concept
                              :
                              :
                   --------------------------
      :              :
            ANALOG .... compared with .... TARGET
      :              :
                 Feature ....   "    " .... Feature
      1  ....   "   " ....  1
      2  ....   "   " ....  2
      3  ....   "   " ....  3
      n  ....   "   " ....  n
```

As can be seen in this representation, the analog and target often are examples of a higher-order, or superordinate, concept.

For example, the concept of an electrical field can be explained by drawing an analogy to a gravitational field (see Figure 1):

> Just as the space around the earth and every other mass is filled with an *electrical field*--a kind of aura that extends through space....A gravitational force holds a satellite in orbit about a planet, and an electrical force holds an electron in orbit about a proton. In both cases there is no contact between the objects, and the forces are "acting at a distance." Putting this in terms of the field concept, we can say that the orbiting satellite and electron interact with the force fields of the planet and the proton and are everywhere in contact with these fields. In other words, the force that one electric charge exerts on another can be described as the interaction between one charge and the electric field set up by the other. (Hewitt, 1987, p. 496).

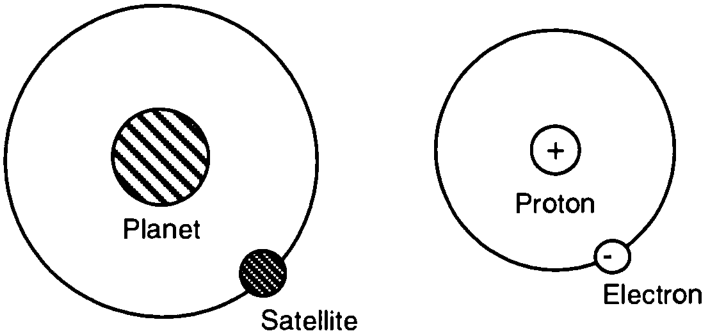


Figure 1. An analogy drawn between a gravitational field and an electrical field.

The superordinate concept in the preceding analogy is *field*, the analog is *gravitational field* and the target is *electric field*. The analogy is drawn by comparing similar features of these two kinds of fields: namely, (a) both fields embody forces, (b) the forces influence objects not in contact, and (c) the forces act at a distance. As can be seen in Figure 1, the planet corresponds to the proton, the satellite to the electron, and the gravitational force to the electrical force.

Analogies and Models

An analogy is a kind of *model* (Glynn, 1991, 1989a,b). A model is a simplified representation of an object (e.g., the human eye) or a process (e.g., photosynthesis). An analogy uses a familiar concept (e.g., a camera) as a model for a less familiar concept (e.g., the human eye). The term "familiar" is relative, of course. A photographer in a biology class might think of a camera as a model for the eye, whereas a biologist in a photography class might think of the eye as a model for a camera. Although an analogy is a model, a model need not be an analogy. For example, consider the stick-and-ball models of molecules that are used in chemistry--these are not analogies. When an analogy functions as a model, it plays a role in theory development, just as any model does:

> A theory is a reasonable explanation of observed events that are related. A theory often involves an imaginary model that helps scientists picture the way an observed event could be produced (Williams, Trinklein, & Metclafe, 1984, p. 3).

When explaining a scientific event, it sometimes is difficult to distinguish clearly where an analogy begins and ends, just as it is difficult to distinguish where a model, hypothesis, or theory begins and ends:

> There is no sharp general distinction among model, hypothesis, and theory. Roughly we can say that a model (whether mechanical or mathematical) is a rather limited conception to explain a particular observed phenomenon. A hypothesis is a statement that usually can be directly or indirectly tested. A theory is a more general construction, putting together one or more models and several hypotheses to explain many effects or phenomena that previously seemed unrelated. (F. J. Rutherford et al., Unit 4, p. 6)

Analogies and Examples

Sometimes examples of a concept are confused with an analogy. An example is an instance of a concept, not a comparison between similar features of two concepts. Consider the relationship between an electric spark and lightning. Lightning is not *like* a big spark, it *is* a big spark! So, lightning is an example of the concept of electric spark.

Analogies and metaphors

Although the terms *analogy* and *metaphor* sometimes are used interchangeably, analogy tends to be used more often in scientific contexts. Metaphor is used more often in literary contexts (e.g., He is a lion among men). In science instruction, an analogy (e.g., An atom *is like* a tiny planetary system) is less likely to be misunderstood than a metaphor (e.g., An atom *is* a tiny planetary system).

Good and Bad Analogies

The effectiveness of an analogy generally increases as the number of similar features shared by the analog and target increases. For example, a camera is often used as an analog when explaining the human eye, because the two concepts share so many similar features:

> In many respects the human eye is similar to the camera. The amount of light that enters is regulated by the **iris**, the colored part of the eye which surrounds the opening called the **pupil**. Light enters through the transparent covering called the **cornea**, passes through the pupil and lens, and is focused on a layer of tissue at the back of the eye--the **retina**--that is more sensitive to light than any artificial detector made....In both the camera and the eye, the image is upside down, and this is compensated for in both cases. You simply turn the camera film around

to look at it. Your brain has learned to turn around images it receives from your retina! (Hewitt, 1987, p. 450-451)

The analogy between the camera and the eye is a powerful one because of the many similar features shared by the two concepts. It is possible, however, to draw a good analogy on the basis of a few similar features, or even one feature, if that feature is directly relevant to the goals of the teacher or author. For example, consider the following analogy:

> Most carbohydrate molecules can be compared to a freight train that is made up of boxcars linked together. Carbohydrate molecules are usually long chains of simple sugars bonded together. We call these large carbohydrate molecules **polysaccharides**. (Smallwood & Green, 1968, p. 54)

The preceding analogy is simple, with the train and boxcars corresponding to the carbohydrate molecule and the sugars, respectively. Despite its simplicity, the analogy is effective because it maps a familiar mental picture, a freight train, onto the target concept, a carbohydrate molecule. The analogy helps one to quickly visualize the general structure of the molecule.

The value of an analogy decreases if it is difficult to identify and map the important features shared by the analog and the target. For example, in the Silver, Burdett and Ginn textbook *Physical Science* (1988), the gravity of "black holes" is explained:

> To understand black holes, astronomers study what happens when gravity is very strong. An object with a huge mass has such strong gravity that it bends space around it. The curved space bends any light that passes by. (p. 465)

The teacher's edition to this textbook provides teachers with the following problem and analogy, presumably to aid them in their explanation of the effects of gravity:

> Is space curved or does gravity "pull on" the light? Analogy: two men walk side by side, ten feet apart, each perpendicular to the equator. To their surprise, they collide upon reaching the north pole. Question: Did gravity pull them together, or is the earth curved? (p. 467)

The answer to the preceding problem is clear: the earth is curved. But what is not clear is how the features of the analog --two men walking, parallel to each other and perpendicular to the equator, until they collide at the north pole--correspond to the target concept, the effect of gravity on space and light? In this analogy, as in any where it is difficult to identify and map the corresponding features, there is a strong likelihood of misconceptions being formed.

Misconceptions Caused by Analogies

Analogical thinking is efficient, helping us to understand new phenomena and solve new problems by drawing upon our past experiences. This is the "bright side" of analogical thinking. There is a "dark side" as well. When one overgeneralizes and maps noncorresponding features of concepts, the results are misconceptions (Thagard, 1992). This dark side of analogical thinking is unfortunate, but a fact of life:

> It is important, for example, to guard against the danger of believing that a model or analogy is an exact representation of some physical system. One should always regard a model critically and remember that an analogy means no more than: under certain special conditions, the physical system being studied behaves as if....(Miller, Dillon, & Smith, 1980, p. 253).

Analogies are double-edged swords. An analogy can be used to explain and even predict some aspects of the target concept or problem; however, at some point, every analogy breaks down and, at that point, misconceptions may begin. Since two concepts or problems are never completely identical, differences always exist among their defining features. It is important to point out these differences when drawing an analogy, as in the case of the camera and the human eye:

> A principal difference between a camera and the human eye has to do with focusing. In a camera, focusing is accomplished by altering the distance between the lens and the film. In the human eye, most of the focusing is done by the cornea, the transparent membrane at the outside of the eye. Adjustments in focusing of the image on the retina are made by changing the thickness and shape of the lens to regulate its focal length. This is called *accommodation* and is brought about by the action of the *ciliary muscle*, which surrounds the lens. (Hewitt, 1987, p. 451)

In the preceding excerpt, Hewitt indicated where the analogy breaks down. He explained that focusing is accomplished differently in the camera than in the eye. By doing so, he reduced the likelihood that students would overgeneralize from the analog to the target concept and form some misconceptions.

Analogies sometimes can be used to clear up students' misconceptions. Clement (1993, 1989) and Brown (1993), for example, have had considerable success coping with students' misconceptions by using demonstrations that build upon the students' real world experiences. For example, many students find it difficult to believe that static objects can exert forces, such as a table exerting an upward force on a book sitting on the table. Students usually agree, however, that a spring exerts a constant force on one's hand when one holds it compressed. If students are presented with an intermediate, "bridging" example, such as a book resting on a flexible board, the students are then more willing to believe that the table exerts an upward force on the book. By drawing an analogy from the spring to the board to the book, students become more receptive to the general idea that even apparently rigid objects are springy to some degree.

Analogical Reasoning, Problem Solving, and Understanding

When scientists are solving problems, expressions such as "It's the same as," "It's no different than," It's just like," and "Think of it this way," are commonplace. These expressions are all ways of saying "This problem is analogous to that one." Analogies play an important role not just in problem solving, but in the identification of new problems, the generation of hypotheses, and the explanation of solutions as well. For this reason, analogical reasoning is viewed as fundamental cognitive process (Anderson & Thompson, 1989; Lawson, 1993; Piaget, 1962).

Cognitive Architecture

To discuss analogical problem solving, it will be helpful to formulate a model of the cognitive components the student brings to bear when problem solving. Cognitive scientists refer to the relatively permanent framework of the mind as the *architecture of cognition*. Our specific model for learning science is consistent with more general models of cognitive architecture (e.g., Anderson, 1990; Baddeley, 1990; Britton, Glynn, & Smith, 1985; Gagne, 1985). Our model of a student's cognitive architecture is intended only to introduce the concepts of metacognition, perception, working memory, and long-term memory in the context of solving science problems. More sophisticated cognitive architectures have been developed known as connectionist network models, or parallel distributed processing models, that are based on parallel rather than serial processing. In these connectionist models, information is distributed over many small units rather than being located in discrete memory stores.

In our *model of a student's cognitive architecture* (see Figure 2), the perception, storage, manipulation, and recall of information are all

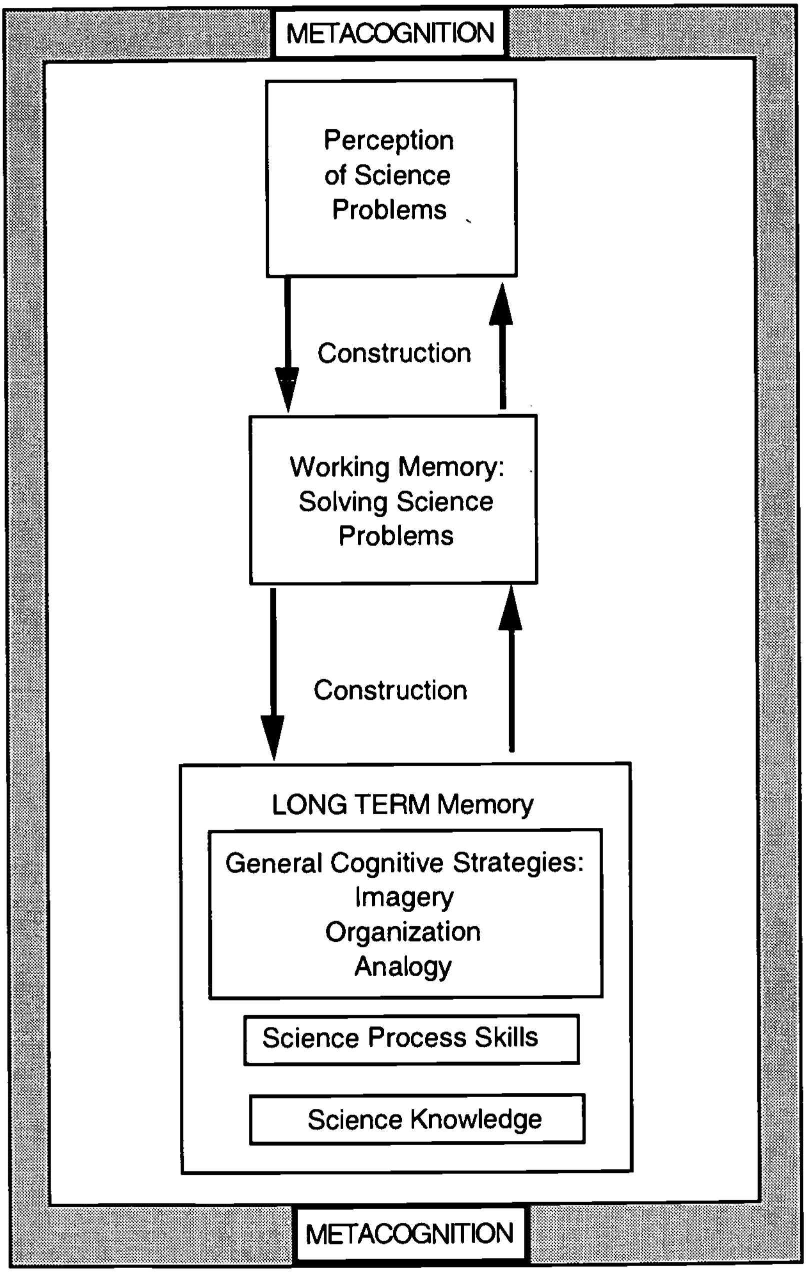


Figure 2. A student's cognitive architecture for solving science problems. Adapted from Glynn and Muth (in press).

constructive processes, not rote ones. These constructive processes are influenced by students' knowledge and experiences, but also by students' expectations, beliefs, values, and social-cultural background.

Our model includes the following interactive components: *metacognition*, *perception*, *working memory*, and *long-term memory*. The metacognitive component represents a student's awareness of his or her own "cognitive machinery" and how the machinery works (Meichenbaum, Burland, Gruson, & Cameron, 1985). Because students differ in their metacognitive knowledge, they differ in how effectively they solve problems. Metacognition has an executive function. It controls the other cognitive components the student brings to bear when problem solving.

Visual, auditory, or tactual information is perceived and then processed in working memory, where intellectual "products" (hypotheses, inferences, generalizations, elaborations, and solutions) are formed using *science knowledge*, *science process skills*, and *general cognitive strategies* retrieved from long-term memory. Working memory corresponds roughly to awareness and serves as a cognitive workspace. Working memory is limited in terms of how much information it can deal with at one time. Information is operated on in working memory and the products of these operations are stored in long-term memory. Once stored, the information can be recalled into the workspace to be used in subsequent operations. Long-term memory has an enormous capacity for storing categorized, hierarchically-organized information.

Learning science constructively requires both science knowledge and science process skills (Carey & Smith, 1993; Dunbar & Klar, 1989; Inhelder & Piaget, 1958; Kuhn, Amsel, & O'Loughlin, 1988). The knowledge engages the skills that, in turn, refine the knowledge. Science knowledge includes formal theories, laws, conceptual models, principles, and facts. This is the "minds on" aspect of science learning and corresponds to what cognitive scientists call *declarative* knowledge ("knowing that").

Science process skills are those procedures routinely performed by practicing scientists in many disciplines. This is the "hands on" aspect of science learning that corresponds to what cognitive scientists call *procedural* knowledge ("knowing how"). In the report *Project 2061: Science for All Americans* (American Association for the Advancement of Science, 1989), the science process skills are identified as: computation, estimation, manipulation, observation, communication, and critical-response. Earlier compilations of science process skills distinguished between basic and integrated skills (Funk, Okey, Fiel, Jans, & Sprague, 1979; Gagne, 1967; Yap & Yeany, 1988). The basic skills are: observation, classification, communication, metric measurement, prediction, and inference. The integrated skills are: identifying variables, constructing a table of data, constructing a graph, describing relationships between variables, acquiring and processing data, analyzing investigations, constructing hypotheses, defining variables operationally, designing investigations, and experimenting.

Meaningful problem solving involves integrating new knowledge with existing knowledge. This process is complex and the result of an interaction of general cognitive strategies, such as imagery, organization, and analogy (Anderson & Thompson, 1989). The interaction of these strategies facilitates the construction of conceptual relations.

Conceptual Relations

Scientific problem solving involves constructing conceptual relations among new knowledge and existing knowledge (Glynn, 1991; Glynn & Muth, in press; Glynn, Yeany, & Britton, 1991). Studies done of experts and novices in domains such as physics (Chi, Feltovich, & Glaser, 1981) and biology (Feltovich, 1981) have shown that expert problem solvers are experts, not just because they know more facts than novices, but because their knowledge exists in the form of interrelated networks. Conceptual relations in science are of many kinds, including hierarchical, exemplifying, sequential, temporal, comparative, contrasting, causal, temporal, additive, and adversative (e.g., see Mayer, 1985).

Students should think of scientific concepts and problems as organized networks of related information, not as collections of facts. Many science teachers know this, of course, but are not sure how to facilitate relational thinking in their students, particularly when the number of students in a class is large and the concepts and problems are complex--and complex is the rule rather than the exception in biology (e.g., photosynthesis and mitosis-meiosis), chemistry (e.g., chemical equilibrium and the periodic table), physics (e.g., gravitational potential energy and electromagnetic induction), earth science (e.g., plate tectonics and precipitation), and astronomy (e.g., the sun and planetary motion).

Students who are thinking constructively will find a scientific problem challenging, struggle with it, and try to make sense of it by integrating it with what they already know. The students will construct a representation, or "mental model," of the problem, often taking the form of an analogy. This constructed representation has the advantage of being meaningful to the student, and therefore more memorable and applicable. At the same time, this representation may incorporate reasonable misconceptions about the content of the problem. The formation of "reasonable misconceptions" is a normal consequence of meaningful thinking. Reasonable misconceptions result when students construct new knowledge by integrating existing knowledge with that in their long-term memories. These misconceptions could be avoided through rote problem solving, but this kind of problem solving is not easily generalized.

Story Problems and Four-Term Analogies

Analogical problem solving is relatively easy within a conceptual domain, but can be very difficult between domains (Holyoak, 1991). In a series of experiments, Gick and Holyoak (Gick & Holyoak, 1980, 1983; Holyoak, 1985) demonstrated that 75% of the college students they tested were able to solve a story problem by applying previously learned information *after they received a hint to apply it*, but only about 30% were able to solve the problem without a hint. The problems used by Gick and Holyoak were embedded in stories from a variety of domains, such as, military, medical, and fire-fighting.

Story problems are more difficult than the four-term problems of the form A:B::C:D (e.g., Einstein:Relativity::Darwin: ?) investigated by Sternberg (1977, 1986) and Rumelhart and his colleagues (Rumelhart, 1989; Rumelhart & Abrahamson, 1973; Rumelhart & Norman, 1981). In the four-term problems, there is a built-in prompt: the student knows that the A-B relationship is relevant and must be mapped to the C-D relationship. With Gick and Holyoak's story problems, in effect the student is given "C," the problem statement, and asked for "D," the problem solution. To solve the problem, the student must notice that some apparently unrelated information ("A-B"), received earlier in the experiment, is relevant to the problem and must apply it without any prompt or hint. Regarding the difficulty of problems that require the spontaneous recognition and application of potentially relevant knowledge from a different domain, Gick and Holyoak (1983) said:

> It should be noted that in all of our experiments the critical prior analogs were presented in a context in which their problem-oriented character was incidental. Subjects were never explicitly encouraged to use the stories to learn about a novel kind of problem. In many situations, such as an instructional context, more directive guidance in the application of an analogy is often given. It is quite likely that more intentional learning procedures could improve transfer performance in our paradigm. In particular, explicit guidance might facilitate transfer from a single analog. In the absence of such guidance, failure to devise a general schema from a single instance may only reflect appropriate conservatism; without either further examples or direct instruction, the person may have no principled way to isolate the essential causal aspects of the situation....Given the difficulty of schema abstraction from a single analog (at least without the guidance of a teacher), one might ask how anyone could spontaneously notice an analogy between one initial analog and a semantically remote transfer problem. (p.32)

Thus, the research conducted with analogical story problems, when extended to real-life science problems, suggests that it is easier to draw analogies within a specific domain (e.g., heat) than between related domains (e.g., heat and electricity), and it is easier between related domains than distant domains (e.g., heat and respiration). At the same time, caution is called for when extending these laboratory findings:

> Almost all research on analogical transfer has used laboratory paradigms in which subjects are provided with a relatively simple source analog, after which transfer to a novel target analog is studied....Such experiments allow control over subjects' knowledge of the analogs; however, it is unclear whether laboratory experiments capture the richness and complexity of naturalistic analogical reasoning. (Spellman & Holyoak, 1992, p. 913.)

If caution is called for when using analogical story problems to assess real-life analogical problem-solving ability, then extra caution should be in order when using simple four-term problems (A:B::C:?) for this purpose. There is little evidence to suggest that successful performance on these simple four-term problems is related to successful analogical problem solving in real life. Until recently, these simple four-term problems comprised a significant portion of the verbal sections of the Scholastic Aptitude Test (SAT), a major test used for college admission and placement decisions.

A future direction for research in cognitive science involves the support of students' analogical problem solving by means of an *intelligent tutoring system* (ITS). In general, an ITS is a computer-based environment intended to support students' problem solving and concept learning. For example, *ThinkerTools* is an ITS that teaches students to understand basic principles of Newtonian mechanics and solve problems using these principles (White, in press). An ITS consists of various modules, one of which is the *learning environment* defined as "that part of the system specifying or supporting the activities that the student does and the methods available to the student to do those activities" (Burton, 1988, p. 109). These activities should include analogical problem solving, if the goal is to construct a learning environment that supports constructive scientific thinking.

In order to design ITS learning environments that incorporate analogical problem solving, researchers must determine exactly what expert problem solvers actually do. In determining this, researchers will rely increasingly on qualitative research methods, such as, task analysis, systematic interviewing, introspection ("thinking out loud"), and the analysis of written protocols (e.g., see Glynn, Muth, & Britton, 1990; Wiggs & Perez, 1988).

Analogical Thinking and Science Instruction

The role of analogical thinking in science instruction has received increasing attention in recent years (Brown, 1993; Clement, 1989, 1993; Donnelly & McDaniel, 1993; Gentner, 1989; Lawson, 1993). Research findings suggest that teachers and textbook authors often use analogies, but inconsistently (Duit, 1991b; Duit & Glynn, 1992; Glynn, Britton, Semrud-Clikeman, & Muth, 1989; Halpern, 1987; Halpern, Hansen, & Riefer, 1990; Harrison & Treagust, 1993; Thagard, 1992; Thiele & Treagust, 1991; Vosniadou & Schommer, 1988). Sometimes the analogies that teachers and textbook authors use are clear and well developed, while other times they are vague and confusing. What teachers and textbook authors need are models to guide their construction of instructional analogies--analogies that are custom-tailored to the particular background knowledge of their students. One solution, of course, would be to advise teachers and authors not to use analogies. That would be unrealistic because teachers and authors, like all human beings, are predisposed to think analogically. Consciously or unconsciously, teachers and authors will use analogies during problem solving and explanation. The better solution is to introduce teachers and authors to models that use analogies systematically in problem solving.

An Analogical Problem Solving Model

In order to examine how analogy is used as a tool for problem solving and explanation in science instruction, a task analysis of elementary school, middle school, high school, and college science textbooks was performed; the analysis identified how the authors of 43 textbooks used analogies to explain concepts and facilitate problem solving (Glynn, 1991; Glynn et al., 1989). Task analysis, a knowledge acquisition technique, was used "with the intent of modelling an individual expert's thinking" (Wiggs & Perez, 1988, p. 267). Task analysis (Gagne, 1985; Gardner, 1985), also called a "procedural analysis," is a technique that "identifies and structures the basic processes that underlie task performance... in procedural analysis, you are trying to document the basic processes that are involved in performing a cognitive task" (Goetz, Alexander, & Ash, 1992, p. 360). A task analysis of how experts perform a cognitive task leads to a representation of the experts' knowledge and, eventually, to a model of the task that includes the operations carried out in the performance of the task. The model that results from a task analysis can be used to help novices acquire expertise in the performance of a skill.

In the task analysis carried out on the analogies in 43 textbooks, the operations performed in each analogy were listed and tallied. All of the authors were considered "experts." For those textbook problems involving analogies, three main operations were found to be performed with regularity; each of these main operations had two suboperations associated with it. These operations form the basis of what we call the *Analogical Problem Solving* (APS) Model.

Analogical Problem Solving Model

1. **State Problem**
   a. **Represent Problem**
   b. **Identify Important Features**
2. **Retrieve Analog**
   a. **Search for Analog with Similar Features**
   b. **Select Analog**
3. **Solve Problem**
   a. **Map Solution**
   b. **Verify Solution**

Operations in the APS Model

The operations in the APS Model can be explained and illustrated by applying the Model to the problem-based analogies drawn by Paul Hewitt (1987) in his textbook *Conceptual Physics*. The task analysis of textbooks identified his analogies as being among the best.

In the Appendix to this chapter, there are several excerpts from Hewitt (1987) about thermal energy and electric potential energy, plus a problem he poses where he draws an analogy between these concepts. His answer to the problem and explanation also are provided. As these excerpts illustrate, students can be strategically prompted to use analogies to connect related bodies of knowledge for purposes of solving a problem. In this problem, Hewitt cued students to retrieve the appropriate analog from memory. If Hewitt wished to provide the students with further prompts, making the problem "easier," he also could have identified the important features of the analog and target, and mapped them for the students.

The APS Model can be applied to Hewitt's electric potential energy problem. In the state-the-problem operation, students put the problem in familiar terms for themselves. To do this they must perform two suboperations, the first of which is to form a mental representation of the problem. In general, problems can be represented either as diagrams, or sets of logical statements, or algebraic equations, or geometric figures, or simpler cases, or as examples. In fact, the same problem can often be represented in several ways. In the electric potential energy problem, it is important to represent electric potential as electric potential energy *per charge* (PE/charge). It also is important to recognize that the electric potential energy of one body is to be compared with the electric potential of *another body*. If a student did not attend to these features of the problem, he or she might think the problem is asking about what the effect of doubling the electrical potential energy would be on the electric potential of the *same body*.

In the second suboperation of the APS Model, students identify the important features of the problem. The key features of this problem are the relationships between the two central concepts, electric potential energy and electric potential, both of which have to do with electric energy. Electric potential energy refers to the *total* potential energy of a charged body, whereas electric potential refers to an *average*, the electric potential energy *per charge*. This average is important because it permits a definite value for electric potential energy per charge (also called "voltage") to be assigned to a location, say, in an electric circuit, regardless of whether or not a charge exists at that location.

Analog retrieval is the students' next major operation. It consists of two suboperations: searching for an analog with similar features and selecting an analog for evaluation. In the electric potential energy and electric potential problem, the author performed both of these suboperations for the students by "hinting" that the relationship between thermal energy and temperature is analogous to that between electric potential energy and electric potential. Had he not done that, students would have had to search their memories for analog candidates. They would recognize these candidates on the basis of similar features and select the best candidates for evaluation.

It is hoped that students would have recognized that some of the relationships between thermal energy and temperature are similar to those between electric potential energy and electric potential. Thermal energy and temperature also have to do with energy. More precisely, thermal energy refers to the *total* random kinetic energy of atoms and molecules in a body, whereas temperature is a measure of the *average* kinetic energy of atoms and molecules in a body.

The final operation, solving the problem, consists of two suboperations, the first of which is mapping a solution. If students know that a body with more thermal energy than another does not necessarily have a higher temperature, then they can map this information onto the problem as a tentative solution. However, even if the students do not have this piece of knowledge immediately available, it is still possible for them to induce it and then map this induction to solve the problem.

One way to induce a solution is by generating and mapping an example of the analog to an example of the target. In this case, for thermal energy and temperature, a student might generate an example such as this: a huge pot of warm soup may have a great deal more thermal energy than a bowl of hot soup, but it does not have a higher temperature. Then, by way of analogy, the student might generate an example for the electric potential energy and electric potential problem: a charged 12-volt automobile battery may have a great deal more electrical potential energy than a balloon rubbed on a hairy head, but it does not have more electric potential. In fact, the balloon might be charged to several thousand volts! However, it has only a tiny amount of electric potential energy, because the charge it carries is less than one millionth of a coulomb.

The final suboperation in the APS Model is to verify the solution generated. For the electric potential energy and electric potential problem, this can be done by testing the solution by means of experiments.

Errors in Analogical Reasoning

For each of the six operations in the APS Model, there is associated a potential error that can lead students to either no solution or to an incorrect solution. First, students could misrepresent the problem. Second, even if students correctly represent a problem, it is possible for them to miss an important feature of the problem. Third, the search for an analog could prove fruitless because students did not search certain domains. For example, with the electric potential energy and electric potential problem, if students had searched only the domain of mechanics instead of heat, they might not have found a useful analog. Fourth, students might select and use a relatively poor analog simply because it was the first one they encountered in their search. Fifth, students might select a good analog, but fail to correctly map important features of the analog to the solution of the problem. Finally, students might fail to verify that the solution they mapped is, in fact, a valid solution.

Rethinking the Problem--Analogically

If, at any operation in the model, students detect that they have made an error, it is possible for them to cycle back to an earlier operation and correct their error. For example, a solution that cannot be verified may have been mapped incorrectly, so students should reexamine their solution mapping. If a solution cannot be mapped, this may be the result of a poor analog, so readers should select a new analog. If students are unable to find a good analog, it may be the result of searching the wrong domains, so readers should examine other domains. If readers have difficulty finding the right domains, it may be because they have failed to identify important features of the problem or they have focused on unimportant features. And, finally, if students are not identifying important features, it may be because they have misrepresented the problem, and if so, they should try some alternative ways of looking at it.

Problem Finding and Hypothesis Construction

Analogies can help students to discover new problems and hypothesize about their solutions. Here are some examples in which questions are posed which prompt problem finding and hypothesis construction in students. The questions are based upon an analogy drawn between gravity and electricity:

* It is said that a gravitational field, unlike an electric field, cannot be shielded. But the gravitational field at the center of the earth cancels to zero. Isn't this evidence that a gravitational field can be shielded?

* How are a gravitational and an electric field similar?

* How is an electric field different from a gravitational field?

* The vectors for the gravitational field of the earth point *toward* the earth; the vectors for the electric field of a proton point *away* from the proton. Explain. (Hewitt, 1987, pp. 501, 507-508)

In addition to responding to such questions from teachers and textbook authors, students can and should be trained to generate questions for themselves; that is, to discover their own problems through analogical reasoning rather than rely entirely upon others to pose them. Having discovered their own problems, the students can then proceed to hypothesize about possible solutions.

Conclusions

This chapter has focused on how teachers and textbook authors can help students to use analogies when solving problems, finding problems, and generating hypotheses. A future direction for research is to determine how students can construct effective analogies for themselves, independently of teachers and textbook authors (e.g., Glynn et al., in press; Wong, 1993a,b). If students are trained to ask certain questions of themselves and perform key analogical operations, they can prompt themselves to use analogies effectively.

We recommend that science teachers train their students to strategically use a model of analogical problem solving, such as the APS Model. With the APS Model, the students can plan, monitor, evaluate, and improve their problem-solving efforts. Science teachers should look for opportunities during lessons, demonstrations, and laboratory activities to apply the APS Model. For example, a lively class discussion, in which a problem is dissected by students who understand the APS Model, could help the students to better understand the target problem and, at the same time, help the teacher to diagnose misconceptions the students might have about it. Teachers should point out to their students when an analogy might be used to gain insight into a problem.

Teachers also should demonstrate how to apply each of the operations in the APS Model to the problem at hand.

Students should be shown how to criticize the analogies they construct, as well as those constructed by the teacher and the textbook author. Those students who learn how to apply the APS Model for themselves will have a powerful intellectual tool they can bring to bear on complex science problems.

Critics of analogies might argue that students sometimes will construct incorrect relations between problems or concepts, looking for relations where none are known to exist. Critics might say that the resulting relations will impede students' understanding. Critics of analogies might recommend that students not only be cautioned about the potential pitfalls of analogies, but actually should be discouraged from using them. Such a recommendation would be based on a naive view of human cognition. The process of relating problems and concepts by means of analogy is basic to human cognition; in effect, analogical reasoning is "hard wired" and it is unreasonable to expect students not to use it. They will use it. It is the responsibility of science teachers to help students use it effectively. One way to help students use it effectively is to train them in an analogical problem-solving model such as APS, thereby ensuring that they understand analogical reasoning, the errors that can result from it, and the ways to correct those errors.

It is true that students' analogical relations will be incorrect at times, but most of the time the relations will be correct, or at least reasonable. What is most important is that students are encouraged to engage in the *process* of connecting and constructing knowledge. To reject the process because it sometimes leads to incorrect relations would be like throwing the baby out with the bath water.

Acknowledgement

The work reported herein was prepared with partial support to the first and third authors from the National Reading Research Center (NRRC) of the Universities of Georgia and Maryland. It was supported under the Educational Research and Development Centers Program (PR/AWARD NO. 117A20007) and administered by the Office of Educational Research and Improvement, U.S. Department of Education. The findings and opinions expressed here do not necessarily reflect the position or policies of the National Reading Research Center, the Office of Educational Research and Improvement, or the U.S. Department of Education. For additional information, contact Shawn M. Glynn, Depts. of Educational Psychology and Science Education, 325 Aderhold Hall, University of Georgia, Athens, GA 30602.

References

American Association for the Advancement of Science. (1989). *Project 2061: Science for all Americans.* Washington, DC.

Anderson, J. R. (1989). A theory of human knowledge. *Artificial Intelligence, 40*, 313-351.

Anderson, J. R. (1990). *Cognitive psychology and its implications* (3rd. ed.). New York: Freeman.

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist, 48*, 35-44.

Anderson, J. R., Boyle, C. F., Corbett, A., & Lewis, M. W. (1990). Cognitive modelling and intelligent tutoring. *Artificial Intelligence, 42*, 7-49.

Anderson, J. R., & Thompson, R. (1989). Use of analogy in a production system architecture. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning.* (pp. 267-297). Cambridge: Cambridge University Press.

Baddeley, A. (1990). *Human memory: Theory and practice.* Boston: Allyn and Bacon.

Britton, B. K., Glynn, S. M., & Smith, J. W. (1985). Cognitive demands of processing expository text: A cognitive workbench model. In B. K. Britton & J. Black (Eds.), *Understanding expository text* (pp. 227-248). Hillsdale, NJ: Erlbaum.

Brown, D. E. (1993). Refocusing core intuitions: A concretizing role for analogy in conceptual change. *Journal of Research in Science Teaching, 30*, 1273-1290.

Burton, R. R. (1988). The environment module of intelligent tutoring systems. In M. C. Polson & J. J. Richardson (Eds.), *Foundations of intelligent tutoring systems.* Hillsdale, NJ: Erlbaum.

Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *American Psychologist, 28*, 235-251.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Clement, J. (1989). Learning via model construction and criticism. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *A handbook of* *creativity: Assessment, theory, and research.* (pp. 341-381). New York: Plenum.

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching, 30,* 1241-1257.

Donnelly, C. M., & McDaniel, M. A. (1993). Use of analogy in learning scientific concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 975-987.

Duit, R. (1991a). Students conceptual frameworks: Consequences for learning science. In S. M. Glynn, R. H. Yeany, & B. K. Britton (Eds.), *The psychology of learning science* (pp. 65-85). Hillsdale, NJ: Erlbaum.

Duit, R. (1991b). On the role of analogies and metaphors in learning science. *Science Education, 75,* 649-672.

Duit, R., & Glynn, S. M. (1992). Analogien und Metaphern, Brücken zum Verständnis im schülergerechten Physikunterricht. (Analogies and metaphors, bridges to understanding in student-oriented physics instruction.) In P. Häußler (Ed.), *Physikunterricht und Menschenbildung.* Kiel, Germany: Institut für Pädagogik der Naturwissenschaften.

Dunbar, K., & Klahr, D. (1989). Developmental differences in scientific discovery processes. In D. Klahr, & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 109-143). Hillsdale, NJ: Erlbaum.

Ericsson, K. A., & Charness, N. (1994). Expert performance. *American Psychologist, 49,* 725-747.

Feltovich, P. J. (1981). *Knowledge based components of expertise in medical diagnosis* (Tech. Rep. No. PDS-2). Pittsburgh: University of Pittsburgh Learning Research and Development Center.

Funk, H. J., Okey, J. R., Fiel, R. L., Jans, H. H., & Sprague, C. S. (1979). *Learning science process skills.* Dubuque, IA: Kendall/Hunt.

Gagne, R. M. (1967). *Science--A process approach: Purpose, accomplishments, expectations.* Washington, DC: Commission on Science Education, Association for the Advancement of Science.

Gagne, R. M. (1985). *The conditions of learning and theory of instruction* (4th. ed.). New York: Holt, Rinehart & Winston.

Gardner, M. K. (1985). Cognitive psychological approaches to instructional task analysis. In E. W. Gordon (Ed.), *Review of research in education* (pp. 157-196). Washington, D.C.: AERA.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning.* (pp. 199-241). Cambridge: Cambridge University Press.

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Gick, M. L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1-38.

Gilbert, S. W. (1989). An evaluation of the use of analogy, simile, and metaphor in science texts. *Journal of research in science teaching, 26,* 315-327.

Glashow, S. L. (1980). Toward a unified theory: Threads in a tapestry. *Science, 210*, 1319-1323.

Glynn, S. M. (1989a, January). *Conceptual explanations in high school physics textbooks: A Teaching-with-Analogies Model.* Paper presented at the meeting of the American Association of Physics Teachers, San Francisco.

Glynn, S. M. (1989b). The Teaching-with-Analogies (TWA) Model: Explaining concepts in expository text. In K. D. Muth (Ed.), *Children's comprehension of text: Research into practice* (pp. 185-204). Newark, DE: International Reading Association.

Glynn, S. M. (1991). Explaining science concepts: A Teaching-with-Analogies Model. In S. M. Glynn, R. H. Yeany, & B. K. Britton (Eds.), *The psychology of learning science* (pp. 219-240). Hillsdale, NJ: Erlbaum.

Glynn, S. M., Britton, B. K., Semrud-Clikeman, M., & Muth, K. D. (1989). Analogical reasoning and problem solving in science textbooks. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *A handbook of creativity: Assessment, theory, and research.* (pp. 383-398). New York: Plenum.

Glynn, S. M., & Duit, R. (in press). Learning science meaningfully: Constructing conceptual models. In S. M. Glynn & R. Duit (Eds.), *Learning science in the schools: Research reforming practice.* Hillsdale, NJ: Erlbaum.

Glynn, S. M., Duit, R., & Thiele, R. (in press). Teaching science with analogies: A strategy for constructing knowledge. In S. M. Glynn & R. Duit (Eds.), *Learning science in the schools: Research reforming practice.* Hillsdale, NJ: Erlbaum.

Glynn, S. M. & Muth, K. D. (in press). Reading and writing to learn science: Achieving science literacy. *Journal of Research in Science Teaching.*

Glynn, S. M., Muth, K. D., & Britton, B. K. (1990). Thinking out loud about concepts in science text: How instructional objectives work. In H. Mandl, E. De Corte, S. N. Bennett, & H. F. Friedrich (Eds.), *Learning and instruction: European research in an international context* (Vol. 2, pp. 215-223). Oxford, England: Pergamon press.

Glynn, S. M., Yeany, R. H., & Britton, B. K. (1991). A constructive view of learning science. In S. M. Glynn, R. H. Yeany, & B. K. Britton (Eds.), *The psychology of learning science* (pp. 219-240). Hillsdale, NJ: Erlbaum.

Goetz, E. T., Alexander, P. A., & Ash, M. J. (1992). *Educational psychology.* New York: Merrill.

Griffiths, A. K., & Preston, K. R. (1992). Grade-12 students' misconceptions relating to fundamental characteristics of atoms and molecules. *Journal of Research in Science Teaching, 29,* 611-628.

Halpern, D. (1987). Analogies as a critical thinking skill. In D. Berger, K. Pezdek, & W. Banks (Eds.), *Applications of cognitive psychology* (pp. 75-86). Hillsdale, NJ: Erlbaum.

Halpern, D. F., Hansen, C., & Riefer, D. (1990). Analogies as an aid to understanding and memory. *Journal of Educational Psychology, 82,* 298-305.

Harrison, A. G. & Treagust, D. F. (1993). Teaching with analogies: A case study in grade-10 optics. *Journal of Research in Science Teaching, 30,* 1291-1307.

Hewitt, P. G. (1985a). *Conceptual physics.* Boston: Little, Brown.

Hewitt, P. G. (1985b). *Instructor's manual to accompany conceptual physics.* Boston: Little, Brown.

Hewitt, P. G. (1987). *Conceptual physics.* Menlo Park, CA: Addison-Wesley.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 19, pp. 59-87). Orlando, FL: Academic Press.

Holyoak, K. J. (1991). Symbolic connectionism: Toward third-generation theories of expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 301-355). Cambridge, England: Cambridge University Press.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*, 374-419.

Kaskel, A., Hummer, P. J., & Daniel, L. (1988). *Biology: An everyday experience.* Columbus, OH: Merill.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills.* Orlando, FL: Academic.

Lavoie, D. R. (1993). The development, theory, and application of a cognitive-network model of prediction problem solving in biology. *Journal of Research in Science Teaching, 30*, 767-785.

Lawson, A. E. (1993). The importance of analogy: A prelude to the special issue. *Journal of Research in Science Teaching, 30*, 1213-1214.

Mach, E. (1976). On thought experiments. *Knowledge and error.* (pp. 134-147). Dordrecht: Reidel. (Original 1905).

Mayer, R. E. (1985). Structural analysis of science prose: Can we increase problem-solving performance. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text.* Hillsdale, NJ: Erlbaum.

Meichenbaum, D., Burland, S., Gruson, L, & Cameron, R. (1985). Metacognitive assessment. In S. Yussen (Ed.), *The growth of reflection in children* (pp. 3-30). Orlando, FL: Academic Press.

Miller, F., Dillon, T. J., & Smith, M. K. (1980). *Concepts in physics.* New York: Harcourt Brace Jovanovich.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Oppenheimer, R. (1956). Analogy in science. *American Psychologist, 11*, 127-135. Pasachoff, J. M., Pasachoff, N., & Cooney, T. M. (1983). *Physical science.* Glenview, IL: Scott, Foresman.

Physical science. (1988). Lexington, MA: Silver, Burdett & Ginn.

Piaget, J. (1962). *Play, dreams, and imitation in childhood.* New York: Norton.

Polya, G. (1973). *Mathematics and plausible reasoning* (Vol. 1). Princeton, NJ: Princeton University Press.

Rumelhart, D. E. (1989). Toward a microstructural account of human reasoning. In S. Vosniadou and A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 298-312). Cambridge: Cambridge University Press.

Rumelhart, D. E., & Abrahamson, A. A. (1973). A model for analogical reasoning. *Cognitive Psychology, 5*, 1-28.

Rumelhart, D. E., & Norman, D. A. (1981). Analogical process in learning. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 335-339). Hillsdale, NJ: Erlbaum.

Rutherford, F. J., Holton, G. & Watson, F. G. (1975). *Project physics.* New York: Holt, Rinehart & Winston.

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review, 5*, 205-221.

Smallwood, W. L. & Green, E. R. (1968). *Biology.* New York: Silver Burdett.

Spellman, B. A., & Holyoak, K. J. (1992). If Saddam is Hitler then who is George Bush? Analogical mapping between systems of social roles. *Journal of Personality and Social Psychology, 62*, 913-933.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning.* Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1985). *Intelligence applied.* San Diego: Harcourt, Brace, Jovanovich.

Thagard, P. (1992). Analogy, Explanation, and education. *Journal of Research in Science Teaching, 29*, 537-544.

Thiele, R. B., & Treagust, D. F. (1991). Using analogies in secondary chemistry teaching. *The Australian Science Teachers Journal*, 1991, *37*, 10-14.

Vosniadou, S., & Schommer, M. (1988). Explanatory analogies can help children acquire information from expository text. *Journal of Educational Psychology*, *80*, 524-536.

White, B. (in press). The ThinkerTools Project; Computer microworlds as conceptual tools for facilitating scientific inquiry. In S. M. Glynn & R. Duit (Eds.), *Learning science in the schools: Research reforming practice*. Hillsdale, NJ: Erlbaum.

Wiggs, C. L. & Perez, R. S. (1988). The use of knowledge acquisition in instructional design. *Computers in Human Behavior*, *4*, 257-274.

Williams, J. E., Trinklein, F. E., & Metcalfe, H. C. (1984). *Modern physics*. New York: Holt, Rinehart and Winston.

Wong, E. D. (1993). Understanding the generative capacity of analogies as a tool for explanation. *Journal of Research in Science Teaching*, *30*, 1259-1272.

Wong, E. D. (1993). Self-generated analogies as a tool for constructing and evaluating explanations of scientific phenomena. *Journal of Research in Science Teaching*, *30*, 367-380.

Yap, K. C., & Yeany, R. H. (1988). Validation of hierarchical relationships among Piagetian cognitive modes and integrated science process skills for different cognitive reasoning levels. *Journal of Research in Science Teaching*, *25*, 247-281.

Appendix

Questions and hints can be used to prompt students to solve problems by means of analogies. Read the following excerpts, one about temperature and heat, and the other about electrical potential, and then solve the problem posed.

Thermal Energy and Temperature

We are going to investigate more closely the effects of the chaotic and haphazard motion of atoms and molecules that we call *thermal motion*. We begin by considering that which a body has by virtue of this energetic motion: *thermal energy*.... The quantity that tells how warm or cold something is with respect to a standard body is called *temperature*. We say temperature is a measure of random translational motion of atoms and molecules in a body; more specifically, it is a measure of the *average kinetic energy* of atoms and molecules in a body. We know, for example, that there is twice the thermal energy in 2 liters of boiling water as in 1 liter of boiling water, because 2 liters of water will melt twice as much ice as 1 liter. But the temperatures of both amounts of water are the same because the average kinetic energy of molecules in each is the same. So we see there is a difference between thermal energy, which is measured in joules, and temperature, which we measure in degrees. (Hewitt, 1985a, pp. 220-221)

Electric Potential Energy and Electric Potential

Rather than dealing with the total potential energy of a charged body, it is convenient when working with electricity to consider the *electric potential energy per charge*. We simply divide the amount of energy in any case by the amount of charge....The concept of potential energy per charge is called electric potential; that is,

Electric potential = Energy/Charge

(Hewitt, 1985a, pp. 333)

Problem

Can we say that a body with twice the electric potential energy of another has twice the electric potential? Why or why not? (Hint: Consider the analogy with thermal energy and temperature.) (Hewitt, 1985a, pp.337)

Answer and Explanation

We cannot say that a body with twice the electric potential energy of another has twice the electric potential, just as we cannot say that a body with more thermal energy than another has a higher temperature. For example, a barrel full of warm water may have considerably more thermal energy than a cup of hot water. But it doesn't have a higher temperature, or greater KE/molecule. Likewise, the body with twice the electric potential energy (PE) does not necessarily have the greater PE/charge. It is important to distinguish between electric potential energy (PE) and electric potential (PE/charge). The key difference between the two is "*per charge.*" (Hewitt, 1985b, p. 156)

The Role of Representations in Problem Solving in Chemistry

George M. Bodner and **Daniel S. Domin**
Department of Chemistry
Purdue University
West Lafayette, IN 47907

Introduction

Efforts to understand the cognitive processes involved in problem solving have been underway for at least 100 years (Helmholtz, 1894). One approach to this task has focused on differences between expert and novice problem solvers (Chi, Feltovich, and Glaser, 1981; Larkin, McDermott, Simon, and Simon, 1980; Schoenfeld and Herrmann, 1982). Smith (1992) has criticized this expert-novice dichotomy as unjustly equating expertise with success. He argued that successful problem solvers often share more procedural characteristics that distinguish them from unsuccessful subjects than do experts when compared to novices.

This chapter focuses on the differences between successful and unsuccessful problem solvers (Camacho and Good, 1989; Smith and Good, 1984). Our goal is a better understanding of the process by which individuals disembed relevant information from the statement of a problem and transform the problem into one they understand – in other words, how they build and manipulate the representation they construct of the problem. Particular attention will be paid to differences between both the number and kind of representations built by successful versus unsuccessful chemistry problem solvers. A theoretical model based on schema theory will then be described that offers a possible explanation for the role that representations play in determining the success or failure of the problem-solving process.

What Is a Representation?

The first step toward understanding the role that representations play in problem solving in any discipline involves building an adequate definition of what we mean by the term *representation*. Simon (1978) argued that the following is noncontroversial.

Chapter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for scientific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six*. Manhattan, KS: Ag Press.

> The human brain encodes, modifies, and stores information that is received through its various sense organs, transforms that information by the process that is called "thinking," and produces motor and verbal outputs of various kinds based on the stored information.

What is highly controversial, he concluded, is "*how* information is stored in the brain; in the usual terminology, how it is 'represented' ..."

Simon uses the term representation in the sense of an *internal* or *mental* representation – information that has been encoded, modified, and stored in the brain. Martin (1982) uses the term in the same sense when he says that representations "signify our imperfect conceptions of the world."

Estes (1989) makes an important point when he reminds us that "a representation stands for but does not fully depict an item or event." He notes that representations are attempts the brain makes to encode experiences. Thus, a representation is very different from a photograph, which preserves all the information in the scene up to the resolving power of the film.

It is tempting to define an internal or mental representation as the *mental image* that an object or event evokes in the individual who experiences it. Purists would note, however, that there is some question about whether mental representations can be stored as images (Pylyshyn, 1978). Within the context of research on problem solving, it is therefore useful to rely on an operational definition in which an internal or mental representation is assumed to be the understanding an individual constructs about the problem being solved.

Greeno (1978, 1980) has proposed three characteristics that can be used to evaluate a mental representation: coherence, connectedness, and correspondence. A representation is *coherent* when it is internally consistent. It is *connected* when it is related to other concepts (or schemata) the individual has constructed. *Correspondence* reflects the extent to which the representation is accurate because it matches reality. (From a constructivist perspective, correspondence might be better defined as the extent to which the representation fits what Cobb (1989) has defined as knowledge "... taken-for-granted by *specific* communities of knowers.")

The modifiers *internal* or *mental* have been added to the term representation to distinguish the information stored in the brain from *external representations*, which are physical manifestations of this information. An external representation can take the form of a sequence of words the individual uses to describe the information that resides in his or her mind. In other situations, it takes the form of a drawing or a list of information that captures particular elements of the mental representation. Within the context of problem solving in chemistry, it can include the equation – such as $PV = nRT$ or $E = E° - RT/nF \ln Q$ – an individual writes that shapes the way information is processed in subsequent steps in the problem-solving process.

Understanding the Problem: The Early Stages In Problem Solving

Ten years ago, we began a series of experiments to study the relationship between students' performance on tests of spatial ability and their performance on hour exams they took while enrolled in college-level chemistry courses. We expected to find a correlation between students' performance on the spatial ability tests and their performance on exam questions that involved the manipulation of three-dimensional images. We found, however, that the magnitude of this correlation was equally strong for all questions that probed the students' problem-solving skills (Bodner & McMillen, 1986).

Subsequent experiments with students in both general chemistry (Carter, LaRussa, and Bodner, 1987) and organic chemistry (Pribyl & Bodner, 1987) showed that correlations with tests of spatial ability were strongest for exam questions that differed significantly from any the students had seen previously. Thus, the tests of spatial ability correlated best with the student's performance on *novel problems*, rather than *routine exercises* (Bodner, 1991), regardless of the type of question that was asked.

Because the spatial tests used in these experiments were tests of disembedding and cognitive restructuring in the spatial domain we concluded that there were preliminary stages in the problem-solving process that involved disembedding the relevant information from the statement of the problem and restructuring or transforming the problem into one the individual understands. We described the goal of these early stages of the problem-solving process as trying to *understand the problem* or to *find the problem*. Larkin (1985) reached similar conclusions when she concluded:

> To work on the problem, the solver must convert the string of words with which he is presented into some internal mental representation that can be manipulated in efforts to

solve the problem. Understanding the problem then means constructing for it one of these internal representations.

The preliminary stages in the problem-solving process in which students begin to understand the problem can therefore be thought of as stages in which the first step is taken toward building a mental representation of the problem.

Representation Systems

Whereas the work in general chemistry focused on multiple-choice exams, our study of student performance in organic chemistry involved the analysis of answers to free-response questions, such as predicting the product of the following reaction.

$$PhCOOH + SOCl_2 \rightarrow$$

Students who scored well on the tests of spatial ability were more likely to draw preliminary structures in which the "Ph" or phenyl group was represented by a six-member ring and the "COOH" carboxylic acid group was represented by an OH group attached to a C=O double bond, and they were more likely to score well on this question.

Students with low scores on the spatial tests were more likely to write equations such as:

$$PhCOOH + SOCl_2 \rightarrow PhCl + SO_2 + HCl$$

or:

$$PhCOOH + SOCl_2 \rightarrow PhCOOCl + SO_2 + HCl$$

For our purposes, the most important characteristic of these answers is the fact that they are "absurd" – they carry no resemblance to the physical reality of what can happen to the molecules involved in the reaction. In the first example, a carbon and oxygen atom mysteriously disappear. In the second example, atoms are conserved, but there is no way to connect the starting materials to the products of this reaction by a rational process of breaking and forming chemical bonds.

We concluded that one of the differences between students who are successful in organic chemistry and those who are not is their ability to switch from one representation system to another. Students who do poorly in organic chemistry often have difficulty escaping verbal/linguistic representation systems. They tend to handle chemical formulas and equations that involve these formulas in terms of letters and lines and numbers that aren't symbols because they don't represent or symbolize anything that has physical reality. Thus, they see nothing wrong with transforming PhCOOH into PhCl.

Students locked in a verbal/linguistic representation system might recognize that the linguistic and symbolic representations in Figure 1 describe the same compound. But they aren't likely to switch from the representation on the left to the one on the right, or vice versa. Other students – who tend to do better in the course – switch back and forth between these representation systems as needed.

$PhCO_2H$

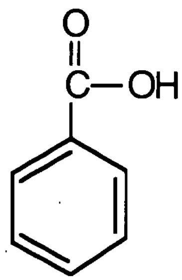


Figure 1. Verbal/linguistic and symbolic representations of benozoic acid

If this hypothesis is correct, similar external representations might be written by individuals with very different internal representations. Consider the following equation, for example.

$$CH_3CH_2CH_2\overset{\overset{O}{\|}}{C}CH_3 + CH_3MgBr \xrightarrow{Et_2O}$$

Students believe that when they write this equation in their notebooks it is a direct copy of what the instructor writes on the blackboard. An external observer, comparing the two, would agree that the students' notes seem to be direct copies of what the instructor wrote. In spite of this agreement, there is a fundamental difference between what the instructor and some of the students write. The instructor writes *symbols*, which represent a physical reality. All too often, students write *letters and numbers and lines*, which have no physical meaning to them

Students for whom chemical formulas are examples of a verbal/linguistic representation system are more likely to write "absurd" formulas, such as the product shown in the following equation.

$$\mathrm{CH_3CH_2CH_2\overset{\overset{\displaystyle O}{\|}}{C}CH_3 + CH_3MgBr \xrightarrow{Et_2O} CH_3CH_2CH_2\overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle CH_3}{|}}{C}}CH_3}$$

It is only when the letters, numbers, and lines used to write these equations are symbols, which represent a physical reality, that students recognize why this answer is absurd.

Multiple Representation Systems

Our understanding of the role of representation systems in organic chemistry was clarified by a study of the problem-solving behavior of students enrolled in an advanced-level graduate course in organic synthesis (Bowen & Bodner, 1991). This study used a variety of tasks, including those shown in Figures 2 and 3, to probe different phases of the process involved in designing the synthesis of complex organic compounds.

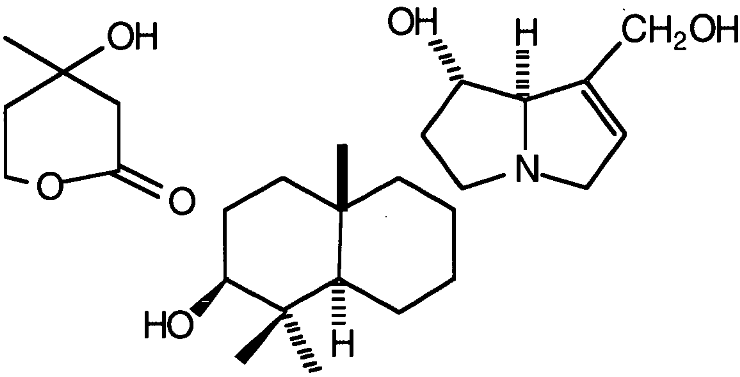
Analysis of the data obtained in this study suggested that the 2nd and 3rd year graduate students used seven different representation systems. Three of these representation systems play an important role from the very beginning of a student's struggle to understand organic chemistry: verbal/linguistic, symbolic, and methodological. Thus, it isn't surprising that these representation systems were the ones that were most commonly used by the graduate students in this study.

Verbal/linguistic representations in organic chemistry can take many forms. They include:

The names of functional groups, such as aldehyde versus ketone, or amine versus amide.

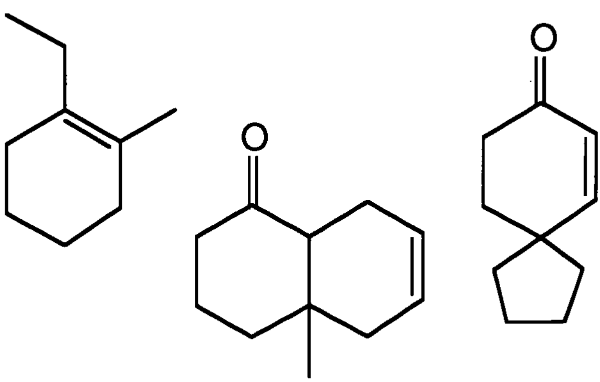
The names of compounds, such as acetone, ethyl acetoacetate, or *cis*-2-methyl-5-hexanolide.

Categories of chemical reactions, such as Michael addition, Diels-Alder reactions, Wolff-Kishner reduction, and so on.



What would be the most difficult portion of theses molecules to construct and why?

Figure 2. A task designed to explore the process by which an individual prepares to solve an organic synthesis problem.



Using readily available starting materials, propose a synthesis for one of these three compounds.

Figure 3. A task designed to probe the process by which a solution to a synthesis problem is generated.

Symbolic representations also take many forms. The structure of benzoic acid shown on the right side of Figure 1, and the line structures in Figures 2 and 3 are examples of symbolic representations. So are the various techniques developed to depict the stereochemistry of a molecule, such as the structure of the **L**-amino acid in Figure 4.

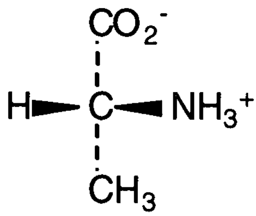


Figure 4. Chemical sructure of L-alanine.

The simplest *methodological* representation system is a chemical equation. The following equation, for example, captures more than the structure of the starting materials and the products of the reaction; it also describes the method by which the transformation occurs.

$$CH_3CH_2Br \xrightarrow[Et_2O]{Mg} CH_3CH_2MgBr$$

Other forms of the methodological representation system – such as the tendency to design an organic synthesis starting from the product and working backwards – are more important for the practicing organic chemist.

The Role of Representation Systems in Determining the Success of Problem Solving in Organic Chemistry

Analysis of answers to free-response questions has provided useful information about the problems students face when building and manipulating representations of organic molecules. Several years ago, students in an organic chemistry class were interviewed shortly after the following question appeared on an hour exam.

A graduate student once tried to run the following reaction to prepare a Grignard reagent. Explain what he did wrong, why the yield of the desired product was zero, and predict the product he obtained.

$$CH_3CH_2Br \xrightarrow[CH_3CH_2OH]{Mg} CH_3CH_2MgBr$$

There is nothing wrong with the starting material, which is a common reagent used to prepare Grignard reagents. There is nothing wrong with the product of the reaction, which is a typical Grignard reagent, or with using magnesium metal to prepare this reagent. The only possible source of error was the solvent: CH_3CH_2OH.

Students for whom the equation was a symbolic representation that had physical meaning were primed to consider what they know about the solvent and recognize that this solvent is a source of H^+ ions that would destroy the Grignard reagent produced in this reaction. Students for whom such equations were verbal/linguistic representations were often unable to answer the question, and were more likely to express the opinion after the exam that this wasn't a "fair" question.

Analysis of student responses to exam questions has also provided insight into the problem that students face when they have to switch between different symbolic representation systems before they can answer a question. Consider the reaction in Figure 5, for example.

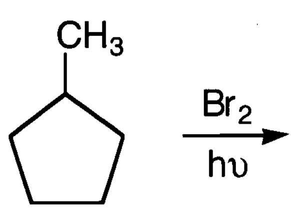


Figure 5. The representation system in which the reaction was presented.

The students were asked to predict the major products of this reaction, estimate the ratio of these products that would be formed if Br· radicals are just as likely to attack one hydrogen atom as another, and use the relative stability of alkyl radicals to predict which product is likely to occur more often than expected from simple statistics. Many students –

particularly those with a history of handling formulas as combinations of letters, numbers, and lines – predicted that this reaction would give the three products shown in Figure 6, in the ratio of 3:2:2.

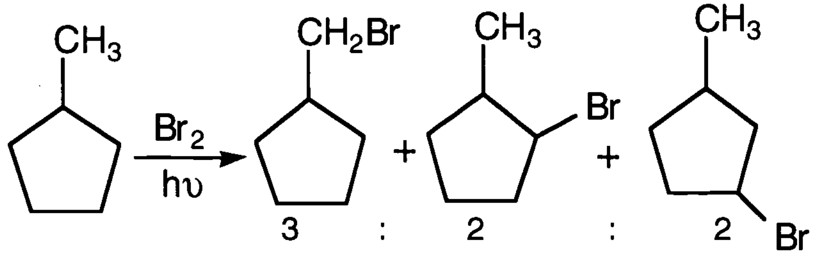


Figure 6. The most common answer.

These students were able to work within the representation system given to them – the line drawing of the starting material. They recognized that attack by the Br· radical at any one of the three hydrogen atoms in the CH_3 group would give the first product. They also recognized that the molecule is symmetrical, and it therefore doesn't matter whether reaction occurs on the right or left side of the molecule when the second and third products are formed

Unfortunately, they failed to recognize that there are two hydrogen atoms on the carbon atoms at which attack occurs to give the second and third products. They therefore failed to recognize that statistics predicts a 3:4:4 ratio for the three products they listed.

A relatively small proportion of the students translated the line drawing for the starting material into a drawing that showed the positions of all the hydrogen atoms in this compound, as shown in Figure 7

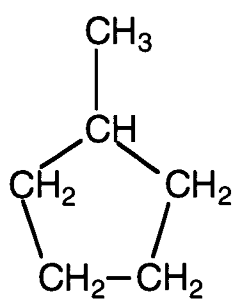


Figure 7. The alternative representation of the starting material drawn by virtually every student who obtained the correct answer to this question.

Students who did this, however, invariably recognized that the reaction actually gives the four products shown in Figure 8.

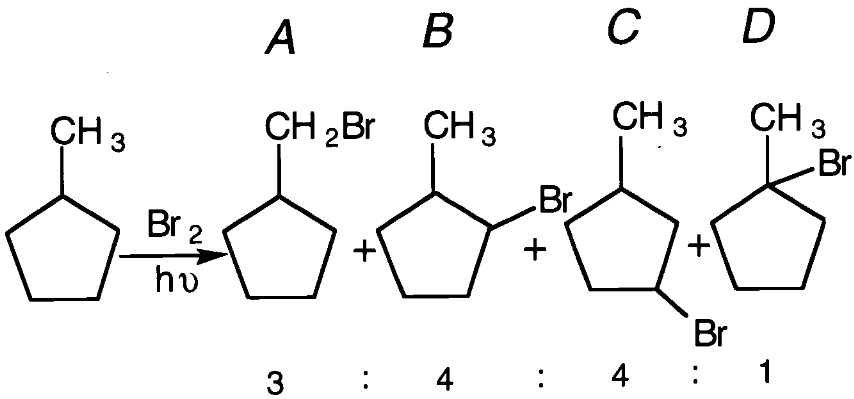


Figure 8. The accepted answer to this question.

These students also recognized that simple statistics would give a product distribution of 3:4:4:1. More importantly, these students came to the correct conclusion that compound *D* is the most likely product to be formed in this reaction because of the stability of the 3° radical formed when the Br· radical attacks this carbon atom.

<u>Differences in the Number and Kind of Representations Constructed During Problem Solving</u>

As we have seen, an essential component of an individual's problem-solving behavior is the construction of a mental representation of the problem, which can contain elements of more than one representation system. Domin & Bodner (in press) therefore studied differences in both the number and types of representations constructed by successful and unsuccessful problem solvers among a population of 1st and 2nd year graduate students faced with questions that dealt with two-dimensional nuclear magnetic resonance spectroscopy (2D-NMR).

The NMR experiment involves irradiating a sample in a magnetic field with electromagnetic radiation in the low-energy, long-wavelength radiofrequency (RF) portion of the spectrum. Nuclei (such as ^{1}H or ^{13}C) that have a "spin" analogous to the spin of an electron interact with the magnetic field to produce two or more spin states at different energies.

The population of these spin states is not quite the same; there is a slight excess of spin in the lowest-energy spin state. The NMR experiment was originally done by slowly changing the frequency of the RF radiation with which the sample is irradiated until it carries just enough energy to induce the nuclei to change spin states. When this happens, the system is in resonance, and a small, but detectable amount of RF energy is absorbed until the population of the spin states is the same and the system is said to be "saturated." The frequency at which resonance occurs depends on the identity of the nucleus being studied. More importantly, there is a very small, but once again detectable, difference in the frequency at which similar nuclei in a molecule absorb. Thus, for example, it is possible to differentiate between the position at which the three hydrogen environments in ethanol (CH_3CH_2OH) absorb radiation.

The FT-NMR experiment involves irradiating the sample with a burst of RF energy, which is equivalent to exciting all of the possible spin-state transitions at the same time. A detector then measures the change in the magnetization of the sample as it decays from saturation back to an equilibrium distribution of spin states. The signal collected from this experiment is then subjected to a Fourier analysis, which transforms the signal from the time domain – in which it is collected – to a frequency domain spectrum identical to the result of the original NMR experiment.

2D-NMR is a two-dimensional NMR experiment that plays an important role in the process by which the individual peaks in the NMR spectrum of a molecule are assigned to specific environments within the molecule. This content domain was chosen because multiple representations not only can but must be used to understand this Fourier-transform (FT) NMR experiment.

The data obtained in this study were consistent with the notion that the ability to switch between representations or representation systems plays an important role in determining success or failure in problem solving in chemistry. Successful problem solvers constructed significantly more representations than unsuccessful problem solvers.

The two groups also differed in the nature of the representations they constructed. Among the successful problem solvers, the most common representations were those that are best described as symbolic. These representations were characterized by a reliance on symbols or highly symbolic equations that might include fragments of a phrase or sentence. The most common representations constructed by the unsuccessful problem solvers were those best described as verbal. These representations, which were expressed either orally or in writing, contained intact sentences or phrases, such as: "the number of spin orientations of a spin-active nucleus is equal to two times the spin-quantum number plus one."

Schema Theory as a Basis for Explaining the Difference Between Successful and Unsuccessful Problem Solvers

Although the successful problem solvers throughout our work constructed significantly more representations than those who weren't successful, neither group constructed very many representations while solving the problems. In the 2D-NMR experiment, for example, the successful problem solvers constructed an average of about 2 representations per problem, whereas those who were unsuccessful constructed an average of just more than 1 representation per problem. A possible explanation for the difference between successful and unsuccessful problem solvers, which might provide insight into the role of mental representations in problem solving, can be found in the schema theory of cognitive structures.

Schema theory views cognitive structure as a general knowledge structure used for understanding (Rumelhart & Ortony, 1977). Schema, also referred to as frames (Minsky, 1975) or scripts (Schank & Abelson, 1977), relate to one's general knowledge about the world. Schema are activated or triggered from an individual's perceptions of his or her environment and they provide the context on which general behaviors are based. Because they don't include information about any exact situation, the understanding of a situation they generate is incomplete. But, by including both facts about a *type of situation* and the *relationship* between these facts, they provide a structure that allows one to make inferences (Medin & Ross, 1992).

Within a given context, problem solving requires the activation of an appropriate schema that contains an algorithm or heuristic that guides the individual to the correct solution to the problem. The construction of the first representation is an effort by the individual to activate the appropriate schema. Thus, the first representation establishes a context for understanding the statement of the problem. In some cases, this representation contains enough information to both provide a context for the problem and to generate a solution to the problem. In other cases, additional representations may be needed since the solution may require more than one algorithm or heuristic. But the first representation provides the context in which the other representations are built.

Unsuccessful problem solvers seem to construct initial representations that activate an inappropriate schema for the problem. This can have three different consequences, each of which leads to an unsuccessful outcome: (1) the initial representation doesn't possess enough information to construct additional representations that contain algorithms or heuristics that might lead to the solution, and the individual gives up, (2) the initial representation leads to the construction of additional representations, but these representations activate inappropriate algorithms or heuristics, and eventually, an incorrect solution to the problem, or (3) the unsuccessful problem solver may never actually achieve an understanding of the problem, in spite of the number of representations that were constructed in an effort to establish a context for the problem.

Implications for the Teaching of Chemistry

Although most of our work on representation systems has focused on organic chemistry, a similar phenomenon exists in general chemistry. Perhaps the best way to illustrate this is to ask the reader to consider the following question: "Which weighs more, dry air at 25°C and 1 atm, or air at this temperature and pressure that is saturated with water vapor?" (Assume that the average molecular weight of air is 29.0 g/mol.)

Most students (and some of their instructors) are convinced that air that has been saturated with water must weigh more than dry air. (It seems reasonable that adding water vapor to air must increase its weight.) Many of these individuals change their mind, however, when they are confronted with Figure 9.

Figure 9 illustrates an important point: *Different representations differ in the information they convey.* Encouraging students to use different representation systems when solving a problem might therefore simply be a way of helping them recognize what information is important in generating the answer to this question.

The symbolic/pictorial representation in Figure 9 prompts us to consider the implications of Avogadro's hypothesis, which assumes that equal volumes of different gases contain the same number of particles. Because the molecular weight of water (18.015 g/mol) is significantly smaller than the average molecular weight of air (29.0 g/mol), water that has been saturated with air actually weighs less than dry air.

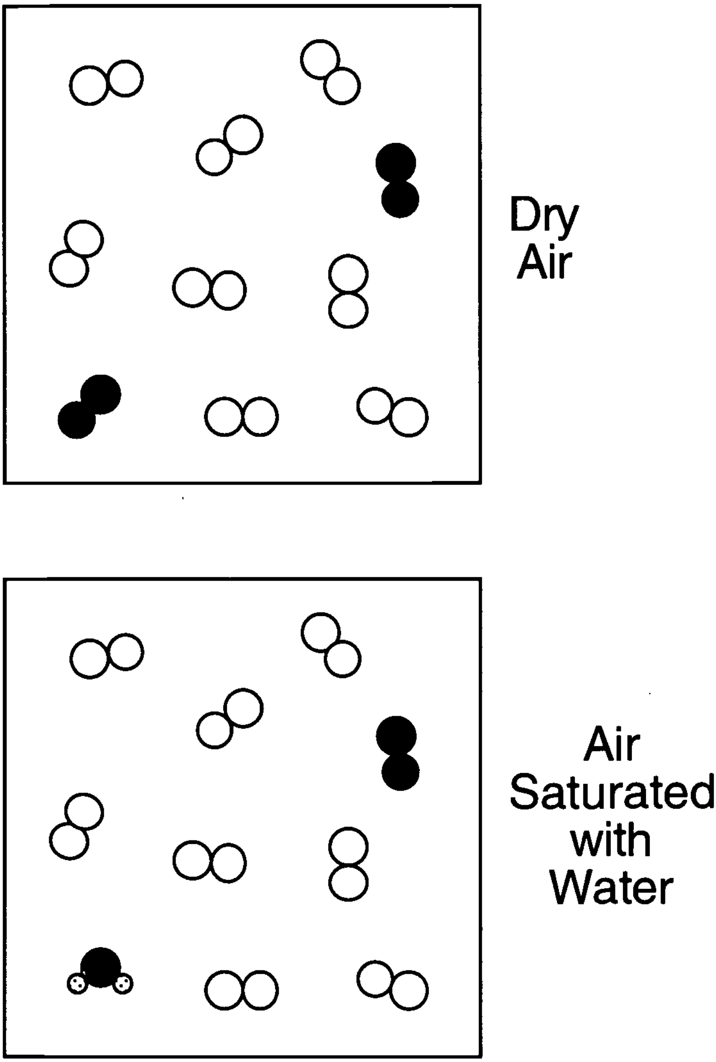


Figure 9. A symbolic representation of the difference between dry air and air saturated with water.

Future Research Questions for Chemistry Problem Solving

Research, by its very nature, raises more questions than it answers. The following is a brief list of questions for future research that might be generated from the hypotheses presented in this chapter.

Does the ability to switch between representations or even representation systems play a particularly important role in chemistry, or is it a significant factor determining the success or failure of problem solving in other domains?

What is the effect on problem-solving performance of changes in the number and types of representations used during instruction?

What is the relationship (if any) between the initial representation constructed during problem solving and subsequent representations?

Are symbolic representations intrinsically more powerful than verbal/linguistic representations, or does their power come from the fact that they are simply an alternative way of representing the information in a problem?

Is the shift from a view of chemical formulas as verbal/linguistic representations to one in which they become symbolic representations an example of *conceptual change* (Posner, Strike, Hewson, & Gertzog, 1982)?

Is it possible to develop an instructional strategy to facilitate this change?

References

Bodner, G. M., & McMillen, T. L. B. (1986). Cognitive restructuring as an early stage in problem solving. *Journal of Research in Science Teaching, 23*, 727-737.

Bodner, G. M. (1991). A View From Chemistry. In M. U. Smith (Ed.), *Toward A Unified Theory of Problem Solving: Views From the Content Domains* (pp. 21-34). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bowen, C. W., &Bodner, G. M. (1991). Problem-solving processes used by students in organic synthesis. *International Journal of Science Education, 13*, 143-158.

Camacho, M., & Good, R. (1989). Problem solving and chemical equilibrium: Successful versus unsuccessful performance. *Journal of Research in Science Teaching, 26*(3), 251-272.

Carter, C. S., LaRussa, M. A., & Bodner, G. M. (1987). A study of two measures of spatial ability as predictors of success in different levels of general chemistry. *Journal of Research in Science Teaching, 24*, 645-657.

Chi, M., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Sciences, 5*, 121-152.

Cobb, P. (1989). Experiential, cognitive, and anthropological perspectives in mathematics education. *For the Learning of Mathematics, 9*(2), 32-42.

Domin, D. S., & Bodner, G. M. (in press). Differences in the number and types of representations constructed by successful and unsuccessful problem solvers. *Journal of Research in Science Teaching*, submitted for publication.

Estes, W. K. (1989). Learning theory. In A. Lesgold and R. Glaser (Eds.), *Foundations for a Psychology of Education*, (pp. 1-49). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J. G. (1978). Understanding and procedural knowledge in mathematics instruction. *Educational Psychologist, 12*, 262-283.

Greeno, J. (1980). Trends in the theory of knowledge for problem solving. In D. Tuma and F. Reif (Eds.), *Problem solving and education: Issues in teaching and research.* Hillsdale: Lawrence Erlbaum Associates.

Helmholtz, H. v. (1894). *Vortrage und reden (5th ed.).* Braunschweig: Vieweg.

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

Larkin, J. H. (1985). Understanding, problem representations, and skill in physics. In S. F. Chipman and J. W. Segal (Eds.), *Learning and Thinking Skills, Volume 2, Research and Open Questions,* (pp 142-159). Hillsdale, NH: Lawrence Erlbaum Associates.

Martin, J. E. (1982) Presentationalism: Toward a self-reflexive psychological theory. In W. B. Weimer and D. S. Palermo (Eds.), *Cognition and the Symbolic Process, Volume 2,* (pp 69-129). Hillsdale, NJ: Lawrence Erblaum Associates.

Medin, D. L. & Ross, B. H. (1992). *Cognitive psychology* . Fort Worth: Harcourt Brace Jovanovich.

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

Pylyshyn, Z. W. (1978). Imagery and artificial intelligence. In C. W. Savage (Ed.), *Minnesota Studies in the Philosophy of Science, Volume IX, Perception and Cognition; Issues in the Foundations of Psychology.* Minneapolis: University of Minnesota Press.

Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982). Accomodation of a scientific conception: Towards a theory of conceptual change. *Science Education, 66,* 211-227.

Pribyl, J. R., & Bodner, G. M. (1987) Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of Research in Science Teaching, 24,* 229-240.

Rumelhart, D. E. & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Ed.), *Semantic factors in cognition.* Hillsdale, NJ: Lawrence

Erlbaum Associates.

Schank, R. C., & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8,* 484-494.

Simon, H. A. (1978.) On the forms of mental representation. In C. W. Savage (Ed.), *Minnesota Studies in the Philosophy of Science, Volume IX, Perception and Cognition; Issues in the Foundations of Psychology*. Minneapolis: University of Minnesota Press.

Smith, M. U. (1992). Expertise and the organization of knowledge: Unexpected differences among genetic counselors, faculty, and students on problem categorization tasks. *Journal of Research in Science Teaching, 29*(2), 179-205.

Smith, M. U., & Good, R. (1984). Problem solving and classical genetics: Successful versus unsuccessful performance. *Journal of Research in Science Teaching, 21,* 895-912.

Research Methods For Investigating Problem Solving in Science Education

Victor L. Willson
Texas A&M University
Educational Psychology
Texas A&M University
College Station, TX 77843-4225

Introduction

A difficulty in summarizing research methods employed for any topic is the inclination to attempt to recapitulate the entire corpus of the topic in the process. Since this work is contributory to a monograph in which many of the issues in science problem solving appear, and the works all contribute in some sense to a scientific problem-solving methodological perspective, the work in this chapter was iterative and relied on them as well as independent investigation.

Some Organizing Principles

A book by Helga Rowe (1985) based on her dissertation provides a useful organizer for problem-solving methods, based on psycho-logical research over the last century. She examined problem solving from the perspectives of Gestalt theory, behaviorism, psychometrics, and information processing. Each has employed preferred methods of research, and all have examined science learning and problem solving to some extent. Rowe herself was a student in a program whose faculty are oriented toward cognitive psychology, and this perspective is present throughout her work.

An arbitrary but defensible decision was made to limit examination of studies to post-1970. In practice this was not restrictive because few studies of interest were found through computer selection, reference backtracking, or nomination that predated 1970. That is not to say that there were no important or interesting studies before 1970, but they were subsumed in important ways into reported research that was done later. It is not coincidental that this date approximates the major shift in learning psychology from a behavioristic to cognitive focus (Willson, 1989). Some have argued that the beginnings occurred two decades, or even more, prior to the 1970's, but the last 25 years have clearly been dominated by cognitive perspectives in psychology.

Another organizer was discipline or field of study. It is safe to say that problem solving in science education does not exist as an independent field; it is linked instead either to broader fields, such as

A chapter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for scientific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

cognitive science and philosophy, or to derivative fields, such as physics education. The organizers employed were cognitive science, psychology, science education, and humanities. Under cognitive science were included the study of expertise from cognitive science; the study of intelligence and thinking were included under psychology, the study of physics, chemistry, and biology learning under science education, and the study of history and linguistics under the humanities. In many instances the studies and organizers are overlapping, so that no definitionally tight claim is made for the structure presented, merely that it is helpful in thinking about methodology. It will become clear to the reader how cognitive psychology has infiltrated most of the categories over the last several decades, and where there is no apparent link, I will provide readers with points for reflection on the appropriateness of syntheses for them.

Finally, an organizing principle for methodology was based on previous work that I have done on examining use of research methods and techniques (Willson, 1980; 1989). This chapter differs from those in that in this work I have not dwelt on statistical counts of methods used. Instead I have used research studies as examples of the dominant methodology I found for that area. My own philosophical and research perspective places constraints on the conclusions that have been drawn. Readers will get some sense of this from papers I have written for the *Journal of Research in Science Teaching* (Willson, 1989, 1987), which I would characterize as post-positivist in the vein of Toulmin and D.C. Phillips. Further, I clearly identify in my research with cognitive psychology principles. While some have argued that those not within a paradigm cannot criticize it (Edelsky, 1990), I believe that it is indeed helpful to obtain outside perspectives, at least as long as the perspectives are identified by origin and purpose. The purpose here is to identify the conceptual map of topics studied with methods used as well as the explanatory shells of the knowledge claims derived from those studies. The overt criticisms from my perspective should be considered from the framework of the reader; are the points indeed commensurable, and if not, is there a correspondence that allows a useful consideration of the point at issue, whether or not it is now considered a weakness in the research?

In the summary and conclusions I will bring in yet another organizer, the structure of the problem being solved, to help illuminate the methods employed by various researchers under different discipline banners. This is itself a major element of cognitive psychology.

Gestalt Problem Solving

Rowe (1985) concluded that Gestalt psychology was "the oldest of the interpretive frameworks within which problem solving was investigated" (p.41). The Gestaltists rejected the atomistic-behaviorist mechanisms for problem solving, instead emphasizing perceptual organization of experience, a precursor of the current connectionist movement in cognitive psychology. Perceptual organization led to the principle of figure-ground dichotomization of stimuli, in which externally-based sensations are sorted mentally into either stimuli attended to or stimuli ignored perceptually. Kîhler (1927), Koffka (1935), and Wertheimer (1945) were notable for their work on problem solving, and as Miller (1986) has noted, borrowed extensively from physics. Kurt Lewin, a student of Kîhler's, was himself a physicist who extended Gestalt theory into social psychology (Miller, 1986). Wertheimer, until his death in 1943, corresponded extensively with Albert Einstein, and Wertheimer investigated Einstein's creative processes in the development of the theory of special relativity. This led to Wertheimer's incorporation of the reconstructed conversations into Productive Thinking (Wertheimer, 1945), published posthumously. Miller's analysis of this reconstruction is itself historical research on scientific problem solving in which he examines the role of imagery and visualization.

Wertheimer's explanation, based on Gestaltist principles that were consciously built to mimic physics as the apex of science, was summarized by Miller as follows: facts (data from experiments) are published or circulated, and problem situations are born, since the relations among the facts are not evident. For the problem solver particularly salient facts are apprehended (perception) or are created within the mind. The mind focuses on these facts (figure) over others (field), but they remain connected; resolving the problem requires restructuring the field. This provides both consistency in the new field and better understanding of the relations among parts of the old field. Wertheimer's interpretation of Einstein's work was that Einstein focused on the relativity of the observer and created the required invariance of light's velocity to solve the problems of time related to, among other things, the Michelson-Morley experiment. Wertheimer concluded that human minds tend toward Gestalt resolution of problems, implying a genetic disposition.

Miller criticizes Wertheimer's work as being a reconstruction of the letters and conversations he had with Einstein from a Gestalt perspective rather than a historical interpretation of Einstein's thoughts that supported Gestalt principles. In particular, Wertheimer argued that the Michelson-Morley experiment on the invariance of the velocity of light was crucial to Einstein's figure-field conception of invariance as a product of the Gestalt problem. Miller concluded from Einstein's own writings that the experiment was interpreted by Einstein after realizing that the concept of distant simultaneity was necessary for a consistency between mechanics and electromagnetic field theory invariant to choice of inertial system. Thus, for Einstein the principle of relativity came first, not the invariance of the velocity of light, and Wertheimer's interpretation was invalidated. Wertheimer went further and concluded that Einstein proposed crucial experiments to test his theory, which Miller also asserted is false. The role of the crucial experiment, explicated in Popper (1959), was important to Wertheimer to resolve the figure-field problem, just as the earlier crucial experiment of Michelson-Morley was important in generating the problem. Miller concluded that neither played the role assigned and that Wertheimer's analysis failed because it was too dependent upon sequences and orders of experiments and actions that were not consistent with Einstein's own contemporary writings and reconstructions. Miller allowed Wertheimer the legitimacy of formulating a regularity or "PrÑgnantz" principle in which scientific discovery tends toward "simplicity, symmetry, and regularity (p.209)." Miller's contributions to problem solving from the imagery perspective will be considered in the section on methods from the humanities.

Rowe (1985) concluded that the Gestalt approach to problem solving, certainly close to science problem solving, had methodological problems due to "inaccessibility of the phenomena being investigated" (p.45). Further, she noted that there was virtually no recognition of the individual differences that might arise in such research. That is, insights into Einstein's processes might be quite irrelevant to those of less gifted persons, even within science problem solving. In that regard, the criticism is similar to that of Willson's (1989) with respect to expert-novice research: the comparison of the most highly developed expertise with novice performance is subject to many threats to validity and tells us nothing about the development of expertise, or of the competence level expected for most students.

A related field that has studied science problem solving is creativity and creative thinking. While the field has drawn on cognitive principles, such as lateralization of brain function, it has done little to incorporate directly modern theory. Brainstorming (Osborn, 1983) is an example from this field. Much of the research in the field is nonexperimental and often suffers from the Hawthorne effect and a lack of consistency in findings. Tests of creativity have been developed (Getzels & Jackson, 1962; Torrance, 1966), but their predictive validity and relation to real-life creative and problem solving has not been demonstrated after many years of research (Tannenbaum, 1983). One of the clarifying distinctions that has come from cognitive science is that much of the effort in creativity research has been in solving <u>ill-structured</u> problems, as opposed to <u>well-structured</u> problems, and that methods to measure outcomes have been poorly developed and understood. Both predictor (creativity) and outcome (ill-structured problem solution) remain elusive.

A recent edited work by Brandwein, Passow, Fort, and Skoog (1988) explored various dimensions of young gifted children and their relation to science learning and interest. Interestingly, most of the articles dealt with curriculum and identification; Brandwein, Morholt, and Abeles

(1988) presented the most focused piece related to problem finding and problem solving. They concentrated on programs for junior high and high school gifted students and reported several decades' work on students' development of experience in experimentation. They suggested that problem finding was not so difficult for students as finding problems that are doable in the context of high school labs and resources. Most of Brandwein, Morholt et al's (1988) chapter, however, is better classified as a program prescription rather than substantive work on problem solving by students.

An interesting distinction between the Gestalt/creativity approach to problem solving, essentially a precursor to cognitive psychology, and that of information-processing is found in the recent writings of Herbert Simon (1989, p.376):

> Moreover, the insight that is supposed to be required for such works as discovery turns out to be synonymous with the familiar process of recognition; and other terms commonly used in discussion of creative work- such terms as <u>judgment</u>, <u>creativity</u>, or even <u>genius</u> - appear either to be wholly dispensable or to be definable, as insight is, in terms mundane and well-understood concepts.

This comment itself helps to distinguish between Gestaltist creativity and information-processing approaches to problem solving.

<u>Behaviorist Problem solving</u>

Little behaviorist problem solving research remains from its peak in the 1950's except in disguised form. That is, the legacy of behaviorism rests in several concepts retained in more recent cognitive models, such as goal strength, complex association hierarchies, and generalization or transfer (Rowe, 1985). Problem solving was never adequately represented in the S-R or operant models, since the assumption that complex behaviors were built of chains and hierarchies of simple behaviors was never well tested. Further, even fairly simple learning situations were found that did not conform to the S-R models (Goetz, Alexander, & Ash, 1992). The complexities of science learning and problem solving were not investigated except as extensions of simpler theory.

Wheatley (1991) linked behaviorism with cognitive science in contrast to constructivism. The former were identified by him with a philosophical separation of mind from environment, in contrast to a constructivist orientation that mind exists to organize experience and adapt the organism to changes in environmental conditions. That is, Wheatley saw cognitive psychology following in the Cartesian dualistic tradition, while constructivism follows a more wholistic line in which mind cannot be separated from that which it was genetically developed to interpret. Wheatley's major conclusion with respect to behaviorism was that it still shapes science classroom experience through the model and reinforcers of the behaviorist account of learning: codified knowledge rather than problem-solving activity. He advocated a problem-centered learning model with task, groups, and sharing as the defining characteristics. Relatively little research in problem solving was linked to this thought-piece, however. Cognitive psychology was a fundamental break with behaviorism, and it shares much more in common with constructivism. Constructed knowledge is a fundamental tenet of most cognitive theories. Wheatley's emphasis on social construction diverges more from mainstream cognitive theories than does the concept of construction per se.

Psychometric Problem Solving

Problem solving has had a long history within the psychometric framework in psychology, and it appears in the earliest mental tests of Binet and Simon (1905), as well as all the successors over the decades that followed. Puzzles, block designs, and other nonverbal tasks were included in intelligence test batteries. Most of the processes invoked in these tasks have been incorporated into cognitive theory, and some, such as analogies, have become central components (Sternberg, 1977).

Science problems were not explicitly evaluated until Piaget began his seminal work on the nature of errors in children's responses to Binet's tasks. Piaget the geneticist studied children as an ontogenetic substitute for phylogenetic development in humans, a fact little discussed in education. He was specifically interested in scientific problem solving as a highly developed human mental activity. While this research is usually presented under developmental psychology due to its stage model, it best fits under the psychometric tradition because of the emphasis on the tasks fundamental to the theory. Piaget's formalizations of his observations resulted in experimental methods for exploring various conservation tasks built around physical characteristics of objects in the child's world. This, in turn, has led to research over the last several decades, largely experimental, into students' ability to solve problems in physics, chemistry, and genetics based on their stage of development.

Lawson and others have developed various tests for estimating Piagetian stage in students. In a study exemplary of the body, Lawson (1992) recently argued that the stages were better represented in adolescents as intuitive or reflective thinking, based on experimental method in which various Piagetian tasks were represented in demonstration form for which students responded in written form. He also administered logic items to assess inferential ability of the college students involved. He argued that the Piagetian tasks better reflect the ability of a student to use general reasoning that generates alternative antecedent conditions for problems. A possible difficulty in Lawson's study is that student prior knowledge for the various Piagetian tasks was not examined. Prior knowledge has become an important working concept in cognitive theories. Since each task used different content knowledge, it is possible for scores of each student to be high or low due to their prior content (general domain) knowledge for the various domains tested, yet the group means to reflect appropriately the average population characteristics. This might well result in the situation in which population means do not represent any individual in the population. That has been a recurrent criticism of the Piagetian tasks - the situated aspect of knowledge related to the problem formulation and solution. The criticism has come from adherents of cognitive problem-solving perspectives (cf. Perkins & Salomon, 1989).

Other psychometric models for problem solving have followed from the investigation of intelligence. Guilford's Structure-Of-the-Intellect (SOI) model (Guilford, 1967) included process functions or productions for which he has developed tests over the last several decades. Cattell (1963) developed a two-factor hierarchical model that posited problem solving under fluid intelligence. Little specific research related to science problem solving has resulted from this work, however. In particular, these models have virtually no developmental or instructional links, which significantly limits their application to science education (or any other field). Modern theories of intelligence that incorporate information-processing theory, however, have much more potential to assist and interpret science problem solving.

Cognitive Science Problem Solving

Cognitive science may be considered the parent field for modern science education problem solving, since a dominant portion of the current literature stems from it. Bransford, Sherwood, Vye, and Rieser (1986) provided a recent overview focused on the historical roots of mental discipline and problem solving. They represented much of the discussion that follows here as a problem in knowledge acquisition, and discussed five parts: problem identification, problem definition, problem exploration, problem action and revision, and the role of metacognition. Carey (1991) gave yet another account of the cognitive perspective, emphasizing the role that research on reading acquisition has played in the concepts of schema theory, which itself can be traced to Piaget's work, discussed elsewhere in this paper. Carey emphasized the role of vocabulary and schema in problem solving and invoked a literature quite overlapping with that in this work. She made a point that weak versus strong restructuring is a major consideration in understanding how students solve problems in science education.

A very recent monograph by diSessa (1993) gives yet another view. While it is focused on physics learning, it can serve as an organizer for a particular theory of science learning in which problem solving is an inherent part. diSessa argued that students build a loosely organized knowledge system from their world experience that is the primary means to inform their school experience. He used the concept of cognitive mechanism, a "structured priority" in mental processing, as a connectionist system to explain mental activity. This characterization is a structural model that has relatively little to say about problem solving except that solutions follow from the degree of organization of the knowledge system, the cognitive mechanism. He focused on the distinction between algorithmic solution of well-defined problems, reported to him by students as a capability divorced from mechanistic understanding. His research was based largely on beginning physics college students and employed interview method.

Ueno (1993), in response to diSessa's (1993) monograph, argued that the mental mechanism model of diSessa is exemplary of a dated conception of learning, and that situated cognition better reflects the current understanding. In this theory problem solving in physics, the primary focus of both works, requires changing or expanding the context that the student understands the problem to be related to. This highlights the notion of situated cognition as a constantly constructed mental activity. Every time stored knowledge is accessed, it is reconstructed and reinterpreted in the context of the environment that currently exists. Particular knowledge may not be accessed at all within some environments, even though they might be appropriate. This is used to explain the contextualized form of student problem solving, in which the same problem that was solved in one class is not recognized in another.

Artificial Intelligence Roots

When the research on artificial intelligence shifted from power (search exhaustive) to strategy (search selective) models, research on expertise as a subfield in turn developed. Early memory work was done with chess masters and novices (de Groot, 1965; Chase & Simon, 1973a, 1973b). None of the early chunking work was done specifically in science domains, although research with electronic technicians (Egan & Schwartz, 1979) and architects (Akin, 1980) might be considered domain-related.

Information-Processing Perspective

A second area of cognitive science study directly corresponds to the emphasis of this paper, strategies for problem solving. That line of inquiry followed Newell and Simon's (1972) theory of means-ends analysis. A general exposition of this perspective was given by Hegarty (1991), based on Newell and Simon (1972): a problem has three main components, a given state, goal state, and set of operators for transforming the given state into the goal state. Note that this fits the well-defined problem. Ill-defined problems are mostly concerned with defining the goal state.

The converging lines of research (Glaser & Chi, 1988) indicated that:

1) experts excel mainly in their own domain;

2) experts perceive large, meaningful patterns in their domain;

3) experts solve problems quickly with little error, much faster than can novices;

4) experts have superior short and long term memory related to their domain;

5) experts see and represent problems in their domain at a deeper level (principled) more so than do novices;

6) experts spend a great deal of time analyzing a problem qualitatively; and

7) experts have strong self-monitoring skills.

Another categorization was made by Schultz and Lochhead that overlapped the Glaser and Chi list, but to which can be added:

8) experts organize quantitative calculations through an understanding of qualitative relations;

9) experts represent problem situations via diagrams or drawings;

10) experts organize knowledge according to principles that bear on the solution of the problem at hand; and

11) experts evaluate the validity of a provisional physical (or other) model through an analogy or chain of analogies.

Each point will be examined from a methodological perspective below.

1. <u>Experts excel mainly in their own domain</u>. The only evidence in science education supporting this point that I found was from the

reference by Glaser and Chi to Voss and Post (1991), who compared political scientists to chemists in their problem-solving approach to a political science problem. The use of experimental-comparative method is appropriate to such research. It is less clear that problem solving within science will conform to this point. Do physicists approach biology problem solving more like experts than novice biologists? Is there a hierarchy within science that is relational with respect to problem solving, or are the knowledge structures required for expertise so critical that novice-like behavior will be observed? Causal-comparative methodology for this area seems quite appropriate, particularly if linked with training studies that demonstrate (or not) that experts can become more expert-like in another domain more quickly than novices in the second domain, as a theoretical analysis might suggest, since the experts already know the strategies associated with problem solving.

Early research on knowledge structure focused on puzzle problems and logic problems with definite solutions in which the demand for domain knowledge was low or nonexistent. Related work focused on physics; Shavelson (1974) examined within-student change, and unlike most expertise literature, demonstrated changes in students' reported organization of physics content. Later work directly compared reported expert and novice knowledge structures (Larkin, 1979).

It is important to emphasize that these studies do not directly provide evidence for <u>how</u> knowledge is structured or represented in the brain; everything from the think-aloud method is the result of the linear filter of verbal report. Concept-mapping (Armbruster & Anderson, 1984; Novak, 1981), webbing (Norton, 1989) and networking (Holley & Dansereau, 1984) are all similar procedures that allow a recapitulative multidimensional response: subjects can overtly make multiple links to their concepts.

Comparative experiments have been conducted over the last decade comparing between- and within-student change in concept-mapping, primarily in science education (see Wallace & Mintzes, 1990, for a review). Again, these procedures do not necessarily represent how knowledge is actually stored, which has been the object of much speculation and research from many perspectives, including artificial intelligence (Rumelhart, McClelland, and the PDP Research Group, 1985) and neuroanatomical psychology (Wise, Chollet, Hadar, Friston, Hoffner, & Frackowiack, 1991). Nevertheless, how experts represent the knowledge they possess appears to differentiate them from novices, and demonstrable changes in novices have been consistently documented. The Wallace and Mintzes study is significant because it demonstrated experimentally a half-standard deviation effect advantage for concept-mapping on a traditional achievement test. Much of the earlier literature had been inconclusive that concept-mapping improved achievement test scores.

Hegarty (1991) distinguished between domain-specific conceptual and procedural knowledge. Alexander and Judy (1988), not referenced by Hegarty, distinguished among declarative, procedural, and conditional knowledge: what can be accessed, and how and when to access that knowledge. Hegarty discussed conceptual knowledge as principled (see point 4 below). She noted that the problem representation state includes both the text and the problem solver's knowledge accessed at that point. She then discussed three types of conceptual knowledge: intuitive, practical, and theoretical. Intuitive knowledge research is typically based on comparing responses of children of different ages (Kaiser, McCloskey, & Proffitt, 1986) based on verbal report when presented a problem. Practical knowledge reflects experience with machines. Hegarty noted the long experience with mechanical comprehension from the psychometric perspective. She posited a hierarchical requirement for understanding lower level systems necessary to understand the working of a given mechanism or process based on Forbus and Gentner's (1984) and others' research based on interviews of people's understanding of mechanical systems. Mayer (1989) and others have used experimental variation in descriptions of mechanical systems to show change in the causal models of the systems.

Hegarty, Just, and Morrison (1988) reported an experiment on pulley problems in which high and low scoring students differed in preference for using internal rules, with low scorers exhibiting inconsistency in rule selection. This research was causal comparative, employing both testing and protocol analysis.

A related finding of the expert knowledge representation research was that the organization may facilitate understanding of related text (Anderson & Schifrin, 1980) and in incidental learning (Stanovich, 1986). Specific research in science domains is limited to the Anderson and Schifrin study of children's knowledge about spiders.

Another subject related to knowledge structure first studied during the late 1970's was that of misconceptions, also termed alternative conceptions or preconceptions in science (Driver, 1990; Mestre, in press). Much of this research was descriptive, and many paper and pencil tests were developed in the area of mechanics to evaluate students' conceptions (Hestenes & Wells, 1992; Hestenes, Wells, & Swackhammer, 1992). One enduring conclusion was that students' conceptions of physics and physical laws that develop outside the classroom are remarkably resistant to change (Champagne, Klopfer, Desena, & Squires, 1981; Clement, 1982). Students may correctly solve problems in a physics course yet revert to prescientific conceptions to frame a problem in the out-of-school world. This has contributed to the concept of situated cognition. This research has been based primarily on psychometric methods. Various testing procedures have been employed, including multiple choice, extended problem solution, think-aloud, and Piagetian task. Biology cross-age studies by Arnaudin and

Mintzes (1985) and Trowbridge and Mintzes (1985) characterized children's developing concepts.

2. Experts perceive large, meaningful patterns in their domain. Chi, Feltovich, and Glaser (1981) used expert-novice research to suggest that knowledge organization differs with expertise level. Experts organize their knowledge into problem schemata based on physical principles that include problem-solving procedures (Hegarty, 1991). Schmidt and Boshuizen (1993) argued that there are two competing theories about expert knowledge development: weak restructuring and radical restructuring (Vosniadou & Brewer, 1987). The former assumes an accretion model of knowledge, while the latter assumes a restructuring of knowledge that permits deep understanding in a series of phases. The Schmidt and Boshuizen studies examined development of expertise in medical practice, sufficiently related to our general theme of science education to permit in-depth analysis. They used cross-sectional designs to support their hypotheses that a) expertise progresses through a series of transitory phases; b) knowledge structures change from causal networks to scripts; c) knowledge is layered in memory with expertise; d) earlier layered knowledge is available when later learned knowledge fails in problem representation; and e) episodic traces of earlier problems (clinical cases) are used in new problem representation and solution. Gadd and Pople (1990) discuss a weak restructuring model based on work by Simon, Lesgold, Feltovich and others in which knowledge is accumulated and reorganized. In their framework experts are better at discerning what is missing or left out of information they need. It is interesting to note that the experts discussed by Gadd and Pople were clinician-teachers, while those in the Schmidt and Boshuisen studies often were practicing physicians. This distinction about the kind of experts studied has not been explored and represents another potential generalization problem for this area.

diSessa's (1993) work fits with this section; while he draws on others' work listed here, his methodology of choice is interviews oriented around physics problem situations.

3. Experts solve problems quickly with little error, much faster than can novices. I could find no studies directly related to science that support this conclusion, which is based on studies in chess, taxi driving, and the like.

4. Experts have superior short-term and long-term memory. Again, I found no science studies that attend to this characteristic.

5. Experts see and represent problems in their domain at a deeper level (principled) more so than do novices. In this area studies in science abound. Chi, Feltovich, & Glaser (1981) discussed the

categorizations of physics experts around principles and categories while novices used problem-specific features for their organization. Larkin, McDermott, Simon, and Simon (1980) showed that physics experts spend a good deal of time creating an adequate representation of the problem; while novices may also create diagrams and representations (Larkin, 1980), they frequently contain both errors of relationship or fact as well as incompleteness (Heller & Reif, 1984). diSessa (1993) emphasized the structured nature of mature physics understanding; it is this structure that leads to problem solving.

It has been less well documented that novices develop toward the kind of expertise represented in the research in this section, although Larkin has conducted training studies that support such conclusions (Willson, 1990). With much of the research conducted at the college level, implications for precollege instruction and understanding are much less clear than is often represented.

6. Experts spend a great deal of time analyzing a problem qualitatively. I located no science problem-solving research related to this conclusion. Schultz and Lochhead (1991) discuss in some detail how this can occur but provide no references or data.

7. Experts have strong self-monitoring skills. The rise of research on comprehension monitoring and awareness, also termed metacognition, was associated with many studies of physics experts and novices: Simon and Simon (1978); Chi, Glaser, and Rees, 1982; and Larkin (1983). Experts were more accurate at judging the difficulty of problems, more likely to check their work, and more likely to abandon unsure problem approaches before time-consuming computations were carried out. Again, Willson's (1990) criticisms of the selection threat affect the ecological validity of many of these findings with respect to inferring change in novice's problem solving toward expertise. We have little evidence that the experts, when they were novices, were in similar states to the novices with whom they are compared. Thus, there is little to indicate that the typical novice will develop toward the self-monitoring skills exhibited by the experts.

8. Experts organize quantitative calculations through an understanding of qualitative relations. Again, I found no clear science education research that appeared focused on this point.

9. Experts represent problem situations via diagrams or drawings. Schultz and Lochhead (1991) discuss this point without references or data. Hegarty and Just (1989) concluded that low mechanical ability subjects rely less on diagrams when text knowledge is incomplete than do high ability subjects, nor can they gain information from diagrams (Hegarty, 1991). These and related studies employed causal-comparative method. Diagramming is closely related, however, to

the issue of mental imagery discussed by Miller (1986), reviewed in this chapter under the headings of Gestalt psychology earlier and of imagery discussed below.

10. Experts organize knowledge according to principles that bear on the solution of the problem at hand. Schultz and Lochhead (1991) suggest that schema reorganization characterizes experts' knowledge and cite a computer-based teaching program by Mestre, Dufresne, Gerace, Hardiman, and Touger (1988) that assists students in selecting effective solution paths for mechanics problems. Much of the discussion in point 1 above fits this point.

11. Experts evaluate the validity of a provisional physical (or other) model through an analogy or chain of analogies. Sternberg (1977) pursued a theory of intelligence based on analogical reasoning, and during the ensuing decade a great deal of research on analogies took place. Little of it directly related to science education, although Alexander and her colleagues used science domain texts to explore verbal analogy training and representations (Alexander, Kulikowich, & Pate, 1989; Alexander, Pate, Kulikowich, Farrel, & Wright, 1989). They examined the limitations of domain knowledge in analogy, and began the study of science interest related to the texts. Their studies are largely experimental with random assignment or covariance adjustment for treatment group differences. Again, the linkage of text to both cognitive and affective processes appears to be fruitful and important to future science education problem-solving research from the cognitive-science perspective.

Schultz and Lochhead (1991) reported that Clement and his colleagues (Brown, 1988; Clement, 1987; Schultz, Murray, Clement, & Brown, 1987) have used analogies in both tutorial and computer-based instruction to help students understand principles at a deeper level. The methodologies of these works are largely nonexperimental or baseline (within-subject) comparisons.

Subject Matter Focused Research

Physics

In numerous references physics is made almost synonymous with problem solving. That linkage, however much physicists would like to make it, is not true from any of the problem-solving perspectives studied. Physics has a well-defined body of declarative knowledge and an almost algorithmic set of procedures for solving standard physics problems in mechanics, optics, and electromagnetics. These occupy most of the standard physics coursework. If physics teachers think they are teaching problem solving, their pedagogy and texts pays little attention to the eleven items listed above that currently define expertise in problem solving. This assumption may well blind physics teachers to the tasks demands of their domain as well as limit their thinking about what the aims of physics are for the typical high school or college student. The methods used to evaluate problem solving in physics have in the last two decades been dominated by cognitive think-aloud, with little systematic thought given to the methodological requirements to study the implementation of interventions related to the characterizations of expert problem solvers. Some researchers have examined strategy training, and experimental methodology has appeared appropriate to evaluate training efficacy.

In addition to the literature cited in the discussions under cognitive science, other recent works related to problem solving include McMillan and Swidener's (1991) study of six novices given one problem. The six were categorized into one of three types, and a broad conclusion was reached that students do not use qualitative thinking. This represents a rather unfortunate trend in science education research to generalize on the basis of few cases under the guise of qualitative, naturalistic, or constructivist research tradition. A few cases, no matter how representative or extraordinary, remain inert; we have little understanding if the results are exhaustive or knowledge of prevalence.

Researchers who have explored instruction to help novices become more like experts are Reif, Larkin, and their associates. I have criticized some of the methodology related to this research elsewhere (Willson, 1990) and will not reiterate it here except to continue to insist on careful explanatory linkage between observation, intervention, and generalization in studies based on quantitative or qualitative methods.

Novak (1982) was the only researcher I discovered who followed up artificial intelligence models to simulate high school and college physics problem solving. This work did not appear to be continued.

Chemistry

Direct instruction methods related to chemistry problem solving have been investigated by Bunce and Heikkinen (1986), and Bunce, Gabel, and Samuel (1991) using multivariate analysis of covariance methods. This application seemed appropriate to investigate intervention effects.

In a study of expert representations among ten graduate students, Bowen (1990) produced seven categories. The limited sample size renders such results speculative and unconfirmed at best.

Biology and Genetics

Most of the problem solving in biology has centered on genetics.

Slack and Stewart's (1990) article represents some middle ground in terms of the qualitative-quantitative methodology separation. They examined 30 biology students, splitting the group into a subgroup of 12 for whom 4 problems afforded them the basis for hypothesis formation. The remaining 18 became a cross-validation sample. While they reported response tendencies there was a lamentable lack of presentation of even descriptive percentages. Smith and Good's (1984) paper represents the other direction for research, evaluating 11 students' solutions to seven problems and creating 32 tendencies. Since there were only 77 possible, such categorizations are quite troublesome methodologically, similar to the difficulty I and my colleague Cecil Reynolds (Willson & Reynolds, 1982) found in classification studies in learning disabilities research: when the number of variables approaches the number of subjects, random classifications are extremely likely, and cross-validation in new samples predict zero relationship. Researchers must take care in producing new variables from old when the degrees of freedom are limited. This is not restricted to quantitative studies and can be unfortunately disguised in qualitative work.

Other genetics problem-solving research focused on misconceptions (Browning & Lehman, 1988), which included computer-based instruction. The study was basically descriptive. Smith and Sims (1992) evaluated the necessity for Piagetian reasoning in problem solving and concluded that formal reasoning is not needed for most high school problems. This paper basically employed content analysis, the methodological procedures for which have become reasonably consistent over the last several decades.

In more general biology content Faryniarz and Lockwood (1992) examined environmental problems via computer simulation in an experimental framework; with nonrandom assignment and no use of covariance or other description of groups, the tenability of the t-tests reported is questionable. Criticism of causal attribution in nonrandom two group comparisons goes back to Campbell and Stanley (1966) and before.

Other Cognitive Approaches

Imagery and Spatial Cognition

A competing model to prevalent information-processing (linear) models is dual coding, based on the work of Paivio and colleagues (Paivio, 1971, 1986). It is the only theory that can at present accommodate information processing, emotion, and imagery. Unfortunately, none of the work in this area can be directly related to science education.

My review of recent research on spatial cognition (Cohen, 1985;

Lohman, 1988) turned up little in science education, although spatial knowledge and spatial problem solving are often represented as important in various science domains. Lohman (1988) commented on spatial problem solving that subjects simply do not consistently solve figural tasks the same way, but instead use different processes as the problems increase in difficulty. It has long been known from psychometric research on mechanical reasoning, for example, that spatial ability correlates highly with it. Larkin and Simon (1987) used experimental method to demonstrate that computer simulated spatial representations reduce search time in physics problem solving. Other studies reviewed by Hegarty included self-report of spatial imaging in design by engineers (Ferguson, 1977).

Constructivism

An important new orientation has been adopted in science education around the general framework of constructivism (von Glasersfeld, 1984). Wheatley (1991) contrasted orientations to learning based on behavioral-cognitive principles and on problem-centered principles derived from constructivism. Thus, for him problem solving becomes the central focus of science education. He listed three components: task, groups, and sharing. The selection of literature in support of his thesis is incomplete, and from my perspective, incorrectly assigned. Rather than a dichotomy between constructivism and cognitive psychology, I believe that a better case can be made that constructivism is merely a radical derivative of cognitive science and part of the topic of situated cognition. It is nowhere nearly so completely characterized as is suggested in Wheatley's presentation of constructivism, particularly with respect to the role of prior knowledge versus generalized scripts and frameworks. Nevertheless, Wheatley's construction with respect to learning and tasks is consistent with the literature discussed above. The same cannot be said with respect to the components of groups and sharing. It is simply premature to assume that the literature on problem solving has demonstrated these as necessary components. The literature cited can be logically linked to some of the assumptions, such as learning as a social construction or cooperative learning, but that linkage, for example, is incomplete and theoretical rather than empirical. A competing theoretical position can be taken that knowledge construction at the level of problem solving needs neither other students nor information-sharing in order for students to be successful.

Stewart and Hafner (1991) use some similar references in a companion article. Their emphasis is on models in problem solving, and they expend considerable effort in contrasting constructivist, or at least nonpositivist, thinking with positivistic formulations of scientific problem solving. They emphasize the discrepancy with the positivist formalism of science problem solving; this is not new, as the literature on neo- and post-positivist philosophy has long acknowledged. They comment correctly, however, that the positivist formalism remains the primary account that students receive in college education, and they propose alternatives based on mental models. Stewart and Hafner emphasized problem finding, problem-solving process, and discovery over the theory-laden positivist account of problem solving. They provide description of mental models and problem solving with respect to elaboration and revision. The elements within these are useful for exploration of their own theory but go beyond available research evidence, in my opinion.

Humanistic Methods In Science Problem solving

Historical Method

An excellent example of historical method applied to science problem solving was referenced earlier under the discussion of Gestalt psychology. Miller (1986) examined the role of imagery in scientific problem solving through an evaluation of the writings of scientists who created modern physics: Poincare, Boltzmann, Einstein, Bohr, and Heisenberg. Miller contended that visual thinking was a seminal requirement for physics dating back to the 18th century and Kant. Miller also convincingly argued that modern physics has been represented, particularly by the Gestaltists, as requiring a break with the imagery aspect of the problem-solving method advocated by the German physicists of the 19th century. He also documented the role of psychology and philosophy as central to the thinking of these physicists as they created the new physics. Miller argued that the cultural orientation of the German scientific tradition was both misunderstood by those outside it and must be understood to interpret properly the problem solving they undertook. He also examined the psychological studies undertaken related to the physics of the turn of the 19th century, in particular Gestalt psychology and Piagetian genetic epistemology, and found them inadequate in explaining the important role of mental imagery and concomitant written understandings by the scientists. As an aside, Miller also criticized the oversubscription to the idea of scientific revolution (e.g.. Kuhn, 1962) as failing to explain the links that scientists made to the older theory. Specifically, he suggested that each of the successful "revolutions" brought about by Einstein, Bohr, and Heisenberg were in fact conscious attempts to reconcile and apply the classical models in mechanics and electromagnetism, in which imagery played a large role.

Miller's work is classical historical method, employing primary documents he discovered to support his position, critically reevaluating earlier historical conclusions in the context of new theory and facts. This book is an excellent example of the method and is highly recommended.

Another example of historical method is Shepard's (1989) examination of the use of imagery in Tesla and Galton's thoughts as discussed by Ferguson (1977). These appear, however, to be relatively uncritical reports of earlier accounts.

A general concern with the few historical studies in problem solving is that they may not generalize well. A study of Einstein, Heisenberg, and Bohr may have little to do with the problem solving of most scientists or even any students. What occurs with the extreme of a population may give us insight into the possibilities but little insight into the norm.

Linguistics

In considering problem solving as a cultural activity, the previous discussion by Miller is relevant, as he concluded that culture affected both the problem solvers and their historian-psychologist-interpreters. This method offers the potential for fresh insights of use to science education, while at the same time risking the tautological traps discussed by Phillips (1987), and reviewed by me (Willson, 1989, 1990). Specifically, the concern is that agenda-oriented perspectives have already reached the conclusion the data being examined are supposed to inform. Marxist, feminist, and hermeneutical frameworks were lumped into those perspectives by Phillips as guilty of such errors.

Summary And Conclusions

What can be made from the previous pages? Method in all disciplines is inextricably bound to what is known and how it is known, what is valued and what is not. Table 1 incorporates the idea first mentioned in the Gestalt-creativity section, that problems are either well-structured or ill structured, and summarizes the various approaches used. It is clear that the dominant methodological approach during the last two decades has been think-aloud protocol analysis. Age comparisons and ability comparisons have supplemented think-aloud. In a few cases experimental designs have been used, but they are conspicuous by their rarity.

While criticism of expert-novice differences at this point may be moot, my earlier criticism (Willson, 1990) has not lost its force for me. Selection remains a major obstacle to developmental theories of problem solving unless the degree of expertise is carefully integrated into the theory of problem solving. Several authors now discuss competence instead of expertise. The current theory is a patchwork, as are most theories, of work done at different levels, but often with college students compared with professors. For precollege problem solving the notion of expertise changes to high versus low scorers or high versus low ability.

Table 1. Summary of methodological approaches to problem solving for ill- and well- structured problems across domains.

PROBLEM-TYPE	
Well-structured	Ill-structured
Behaviorism	Gestaltism
Little relevant research.	Historical method and text analysis, vague constructs and problems with circularity in method.
Cognitive Science	Creativity
Emphasis on think-aloud method; problems with selection threats and generalizations; typically small sample sizes and small problem sets for the number of categorizations generated. Developmental approaches use cross-sectional causal-comparative methods; some experimental design in training studies.	Testing approach; suffers from vagueness in definitions, lack of predictive validity.
Piagetian theory is well-developed with good methods; limitations are the specificity of the tasks and their nongeneralization.	Humanities
	Historical methods applied for a few studies; potential problems with specificity of the examples.
	Imagery has been a major emphasis of this work.
Psychometry-intelligence	
Well-developed measures, large sample sizes are plusses; the inadequacy of theories to explain development and lack of specificity for science problem solving are the major problems.	

The training studies examining within-subject change have the least ecological threat and it was good to see that in the late 1980's a number of such studies were being conducted. Perhaps their followups will just now come into print.

Glaring in its omission is the affective domain. The roles of interest, anxiety, social desirability, and attitude are being explored in other aspects of cognitive science, and there is no reason to believe that science problem solving can progress far without them. The cold cognition reported in this chapter will be yet again reconceptualized as the affective component is included.

There do not appear to be significant differences in methods used across subject areas, although physics learning and expertise have dominated the topic space and their methods have become the model for investigations in biology and chemistry. That so much work has been done in physics seems related to the centrality of problem solving to the discipline, compared to the heavier declarative knowledge load in chemistry and biology. This calls into question exactly how problem solving occurs in science domains in which the declarative knowledge demands on problem solving are so great. The 11 problem-solving characteristics of experts have been developed almost exclusively within the physics context. The methodology and results do not give clear guidance for conclusions in heavy declarative knowledge demand fields as opposed to the equation/relation knowledge structure in physics.

The emphasis on upper secondary and college level coursework may also represent the amount of problem solving actually done in science education. Many national studies have called for more problem solving at the elementary and middle school level, yet there is almost no research in problem solving that is coherent and links to the expert-novice cognitive models that dominate upper education levels. This was a surprising finding for me, and must be addressed immediately, since elementary and middle school teachers will find themselves asked to train students in methods that are not yet research based nor closely linked to what we do know about problem solving.

A related issue in the precollege curriculum is the absolute lack of research in ill-structured problems in science education. While Simon's dismissal of ill-structure may suffice for theory, in practice such problems may arise in various curricula, particularly under the science-society-technology focus. A few programs emphasize such problem solving, notably Odyssey of the Mind and other gifted student competitions. These programs have almost no research base to them that justifies their approach except testimonial. If we begin performance assessment (Doran, Helgeson, & Kumar, in press) in problem solving, we must examine carefully the theory invoked as well as method employed.

Finally, problem-solving research in geology and earth science is almost nonexistent. Perhaps the physics problem solving conclusions hold, but it may be helpful for someone to explore those domains. The comments regarding declarative knowledge load apply to them also. Critics of rationalistic, quantitative methodology were right to focus on the centrality of method to epistemology in education. How we know cannot be separated from what we know, just as a zipper ceases to exist if one set of teeth is removed. Problem solving in science education retains strong links to cognitive psychology and its scientific orientation. The tug of alternative epistemologies and their methodologies will create troublesome discontinuities for the field in the future; some sort of correspondence principles will need to be invoked if separate, competing, and noninteractive formulations are to avoided.

Problem-solving methodology is at present insufficiently robust for researchers to be complacent. They are encouraged to consider the points made in this chapter and to select methods that support adequately the hypotheses formulated. Few studies examined here were simply descriptive; most philosophers now agree that no studies are untheoretical - the selection of variables to study and instruments or observation methods serve as markers for theoretical position. Researchers are encouraged to make explicit the theories they espouse and to support their choice to methods as useful in supporting the conclusions they reach. It is not acceptable, however, to dismiss all criticism as simply theory-laden or incommensurable, as Edelsky (1990) attempted to do. The science education community must reach some agreement as to what constitutes acceptable method for supporting conclusions about data. This social agreement is common to all scientific disciplines, and even to most social science fields, although not at present. Since problem solving has much broader roots than just in science education, researchers will need to incorporate or reject a wide range of methods from very diverse sources. This will prove a significant challenge.

References

Akin, O. (1980). *Models of architectural knowledge.* London: Pion.

Alexander, P. A., & Judy, J. E. (1988). Domain specific and strategic knowledge. *Review of Educational Research, 58,* 357-404.

Alexander, P. A., Kulikowich, J. M., & Pate, P. E. (1989). An educology of instruction: Training domain-specific and strategic knowledge. *International Journal of Educology,* 3, 104-125.

Alexander, P. A., Pate, P. E., Kulikowich, J. M., Farrell, D. M., & Wright, N. L. (1989). Domain-specific and strategic knowledge: Effects of training on students of differing ages or competence level. *Learning and Individual Differences, 1,* 283-325.

Anderson, R. C., & Schifrin, Z. (1980). The meaning of words in context. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.),*Theoretical issues in reading comprehension.* Hillsdale, NJ: Erlbaum.

Armbruster, B. B., & Anderson, T. H. (1984). Mapping: Representing informative text diagrammatically. In C. D. Holly & D. F. Dansereau (Eds.), *Spatial learning_strategies: Techniques, applications, and related issues* (pp. 189-221). New York: Academic Press.

Arnaudin, M. W., & Mintzes, J. J. (1985). Students' alternative conceptions of the human circulatory system: A cross-age study. *Science Education, 69,* 721-733.

Binet, A., & Simon, T. (1905). Methods nouvelles pour le diagnostic du niveau intellectuel des anormaux. *AnnCePsychologique, 11,* 191-244.

Bowen, C. W. (1990). Representational systems used by graduate students while problem solving in organic synthesis. *Journal of Research in Science Teaching, 27,* 351-370.

Brandwein, P. F., Morholt, E., & Abeles, S. (1988). Apprenticeship to well-ordered empiricism: Teaching arts of investigation in science. In P. F. Brandwein, A. H. Passow, D. C. Fort, & G. Skoog, (Eds.), *Gifted young in science: Potential through performance* (pp. 273-319). Washington, DC: National Science Teachers Association.

Brandwein, P. F., Passow, A. H., Fort, D. C., & Skoog, G. (1988). *Gifted young in science: Potential through performance.* Washington, DC: National Science Teachers Association.

Bransford, J., Sherwood, R., Vye, N., & Rieser, J. (1986). Teaching thinking and problem solving: Research foundations. *American Psychologist, 41*, 1078-1084.

Brown, D., & Clement, J. (1987, April). Overcoming misconceptions in mechanics: A comparison of two example-based teaching strategies. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Browning, M. E., & Lehman, J. D. (1988). Identification of student misconceptions concerning genetics problem solving via computer program. *Journal of Research in Science Teaching, 25*, 747-761.

Bunce, D. M., Gabel, D. L., & Samuel, J. V. (1991). Enhancing chemistry problem-solving achievement using problem categorization. *Journal of Research in Science Teaching, 28*, 505-521.

Bunce, D. M., & Heikkinen, H. (1986). The effects if an explicit problem-solving approach on mathematical chemistry achievement. *Journal of Research in Science Teaching, 23*, 11-20.

Carey, S. (1986). Cognitive science and education. *American Psychologist, 41*, 1123-1130.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand-McNally.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1-22.

Champagne, A. B., Klopfer, L. E., Desena, A. T., & Squires, D. A. (1981). Structural representations of students' knowledge before and after science instruction. *Journal of Research in Science Teaching, 18*, 92-111.

Chase, W. G., & Simon, H. A. (1973a). Perception in chess. *Cognitive Psychology, 4*, 55-81.

Chase, W. G., & Simon, H. A. (1973b). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence (Vol. 1,* pp. 7-76). Hillsdale, NJ: Erlbaum.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50,* 66-71.

Clement, J. (1987). Overcoming students' misconceptions in physics: The role of anchoring intuitions and analogical validity. *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics,* (Vol. 3, pp. 84-97). Ithaca, NY: Cornell University.

Cohen, R. (1985). *The development of spatial cognition.* Hillsdale, NJ: Erlbaum.

de Groot, A. (1965). *Thought and choice in chess.* The Hague: Mouton.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10,* 105-225.

Doran, R. L., Helgeson, S. L., & Kumar, D. D. (in press). Problem solving assessment in science. In D. R. Lavoie (Ed.), *Towards a cognitive science perspective for scientific problem solving.* Washington, DC: National Association for Research in Science Teaching.

Driver, R. (1990). The teaching and understanding of concepts in practice. In N. Entwhistle (Ed.), *Handbook of educational ideas and practices* (pp. 763-770). London: Routledge.

Edelsky, C. (1990). Whose agenda is this anyway? A response to McKenna, Robinson, and Miller. *Educational Researcher, 8,* 7-11.

Egan, D. E., & Schwartz, B. J. (1979). Chunking in recall of symbolic drawing. *Memory and Cognition, 7,* 149-158.

Faryniarz, J. V., & Lockwood, L. G. (1992). Effectiveness of microcomputer simulations in simulating environmental problem solving by community college students. *Journal of Research in Science Teaching, 29,* 453-470.

Ferguson, E. S. (1977). The mind's eye: Nonverbal thought in technology. *Science, 197,* 827-836.

Forbus, K. D., & Gentner, D. (1984). Learning physical domains: Toward a theoretical framework. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: Recent progress.* Los Altos, CA: Morgan Kaufmann Publishers, Inc.

Gadd, C. S., & Pople, Jr., H. E. (1990). Evidence from internal medicine teaching rounds of the multiple roles of diagnosis in the transmission and testing of medical expertise. In N. Frederiksen, R. Glaser, A. Lesgold, & M. G. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (89-112). Hillsdale, NJ: Erlbaum.

Getzels, J. W. & Jackson, P. W. (1962). Creativity and intelligence: Explorations with gifted children. New York: John Wiley and Sons.

Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The Nature of expertise* (pp. xv-xxviii). Hillsdale, NJ: Erlbaum.

Goetz, E. T., Alexander, P. A., & Ash, M. J. (1992). *Educational Psychology: A Classroom Perspective.* New York: Merrill.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Hegarty, M. (1991). Knowledge and processes in mechanical problem solving. In R. J. Sternberg & P. A. Frensch (Eds.), Complex problem solving (pp. 253-285). Hillsdale, NJ: Erlbaum.

Hegarty, M., (1991). Knowledge and processes mechanical problem solving. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving* (pp. 253-285). Hillsdale, NJ: Erlbaum.

Hegarty, M., & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. Levin, (Eds.), *Knowledge acquisition from text and picture* (pp. 171-194). Amsterdam: North Holland.

Hegarty, M., Just, M. A., & Morrison, I. R. (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. *Cognitive Psychology, 20,* 191-236.

Heller, J. I., & Reif, R. (1984). Prescribing effective human problem-solving processes: Problem description in physics. *Cognition and Instruction, 1,* 177-216.

Hestenes, D., & Wells, M. (1992). A mechanics baseline test. *Physics Teacher, 30,* 159-170.

Hestenes, D., Wells, M., & Swackhammer, G. (1992). Force concept inventory. *Physics Teacher, 30,* 141-158.

Holley, C. D., & Dansereau, D. F. (1984). Networking: The technique and empirical evidence. In C. D. Holly & D. F. Dansereau (Eds.), *Spatial learning strategies: Techniques, applications, and related issues* (pp. 189-221). New York: Academic Press.

Kaiser, M. K., McCloskey, M., and Proffitt, D. R. (1986). Development of intuitive theories of motion: Curvilinear motion in the absence of external forces. *Developmental Psychology, 22*, 67-71.

Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt, Brace & World.

Kîhler, W. (1927). *The mentality of apes*. London: Routledge.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Larkin, J. H. (1979). Information processing models and science instruction. In J. Lochhead & J. Clement (Eds.), *Cognitive process instruction*. Philadelphia, PA: Franklin Institute Press.

Larkin, J. H. (1980). Skilled problem solving in physics: A hierarchical planning approach. *Journal of Structured Learning, 6*, 121-130.

Larkin, J. H. (1983). The role of problem representation in physics. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 75-100). Hillsdale, NJ: Erlbaum.

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

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth 10,000 words. *Cognitive Science, 11*, 65-100.

Lawson, A. E. (1992). What do tests of "formal" reasoning actually measure? *Journal of Research in Science Teaching, 29*, 965-983.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 181-248). Hillsdale, NJ: Erlbaum.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research*, 59, 43-64.

McMillan, C., & Swadener, M. (1991). Novice use of qualitative versus quantitative problem solving in electro-statics. *Journal of Research in Science Teaching, 28*, 661-670.

Mestre, J. P. (in press). Cognitive aspects of learning and teaching science. In S. J. Fitzsimmons & L. C. Kerpelman (Eds), *Pre-college teacher enhancement in science and mathematics: Status, issues, and problems.* Washington, DC: National Science Foundation.

Mestre, J. P., Dufresne, R., Gerace, W., Hardiman, P., & Tougher, J. (1988). *Promoting expertlike behavior among beginning physics students* (Technical Report No. 178). Amherst, MA: University of Massachusetts, Scientific Reasoning Research Institute Department of Physics and Astronomy.

Miller, A. I. (1986). *Imagery in scientific thought.* Cambridge, MA: MIT Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Norton, D. E. (1989). *The effective teaching of language arts* (3rd ed.). Columbus, OH: Merrill.

Novak, G. S. (1982). *Cognitive processes and knowledge structures used in solving physics problems.* Final Technical Report. Austin, TX: Department of Computer Science. National Science Foundation, Washington, DC, NSF/SED-82032.

Novak, J. D. (1981). Applying learning psychology and philosophy of science to biology teaching. *American Biology Teacher, 43*, 12-20.

Osborn, A. F. (1983). *Applied imagination: Principles and procedures of creative problem solving* (3rd ed.). New York: Scribner.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart, & Winston.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 18*, 16-25.

Phillips, D. C. (1987). *Philosophy, science, and social inquiry.* New York: Pergamon Press.

Popper, K. R. (1959). *The logic of scientific discovery.* New York: Basic Books.

Rowe, H. (1985). *Problem solving and intelligence.* Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (1985). *Parallel distributed processing: Explorations in the microstructure of cognition.* Volume 1: Foundations. Cambridge, MA: MIT Press.

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review, 5*, 205-221.

Schultz, K., & Lochhead, J. (1991). A view from physics. In M. U. Smith (Ed.), *Toward a Unified Theory of Problem Solving* (pp. 99-114). Hillsdale, NJ: Erlbaum.

Schultz, K., Murray, T., Clement, J., & Brown, D. (1987). Overcoming misconceptions with a computer-based tutor. *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics* (Vol. 3). Ithaca, NY: Cornell University, 84-97.

Shavelson, R. J. (1974). Methods for examining representations of a subject-matter structure in a student's memory. *Journal of Research in Science Teaching, 11*, 231-249.

Shepard, R. N. (1978). The mental image. *American Psychologist, 33*, 125-137.

Simon, H. A. (1989). The scientist as problem solver. In D. Klahr & K. Kovotsky, (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp.375-398). Hillsdale, NJ: Erlbaum.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. D. Siegler (Ed.), *Children's thinking: What develops?* (pp.325-348). Hillsdale, NJ: Erlbaum.

Slack, S. J., & Stewart, J. (1990). High school students' problem-solving performance on realistic genetics problems. *Journal of Research in Science Teaching, 27*, 55-67.

Smith, M. U., & Good, R. (1984). Problem solving and classical genetics: Successful vs. unsuccessful performance. *Journal of Research in Science Teaching, 21*, 895-912.

Smith, M. U., & Sims, Jr., O. S. (1992). Cognitive development and genetics instruction: A critical review. *Journal of Research in Science Teaching, 29*, 701-713.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning.* Hillsdale, NJ: Erlbaum.

Stewart, J. H. & Atkin, J. A. (1982). Information processing psychology: A promising paradigm for research in science teaching. *Journal of Research in Science Teaching, 19*, 321-332.

Stewart, J. H., & Hafner, R. (1991). Extending the concept of "Problem" in problem-solving research. *Science Education, 75*, 105-120.

Tannenbaum, A. J. (1983). *Gifted children: Psychological and educational perspectives.* New York: Macmillan.

Torrance, E. P. (1966). *Torrance tests of creative thinking.* Princeton: Personnel Press.

Trowbridge, J. E., & Mintzes, J. J. (1985). Students' alternative conceptions of animals and animal classification. *School Science and Mathematics, 85*, 304-316.

Ueno, N. (1993). Reconsidering P-prims theory from the viewpoint of situated cognition. *Cognition and Instruction, 10*, 239-248.

von Glasersfeld, E. (1984). An introduction to radical constructivism. In P. Watzlawick (Ed.), *The invented reality.* New York: Norton, 17-40.

Vosniadou, S., & Brewer, W. F. (1987). Theories of knowledge restructuring in development. *Review of Educational Research, 57*, 51-67.

Voss, J. F., & Post, T. A. (1991). On the solving of ill-structured problems. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The Nature of expertise* (pp. 261-285). Hillsdale, NJ: Erlbaum.

Wallace, J. D., & Mintzes, J. J. (1990). The concept map as a research tool: Exploring conceptual change in biology. *Journal of Research in Science Teaching, 27*, 1033-1052.

Wertheimer, M. (1945). *Productive thinking.* New York: Harper & Row.

Wheatley, G. (1991). Constructivist perspectives on science and mathematics learning. *Science Education, 75*, 9-21.

Willson, V. L. (1980). Research techniques in AERJ articles: 1960 - 1978. *Educational Researcher, 9*, 5-11.

Willson, V. L. (1987). Theory-building and theory-confirming observation in science and science education. *Journal of Research in Science Teaching, 24*, 279-284.

Willson, V. L. (1989). Evaluation of learning strategies research methods and techniques. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 263-274). New York: Academic Press.

Willson, V. L. (1989). Review of Philosophy, science, and social inquiry by D. C. Phillips. *Journal of Research in Science Teaching, 26*, 653-655.

Willson, V. L. (1990). Methodological limitations for the use of expert systems techniques in science education research. *Journal of Research in Science Teaching, 27*, 69-77.

Willson, V. L., & Reynolds, C. R. (1982). Methodological and statistical problems in determining membership in clinicalpopulations. *Clinical Neuropsychology, 4*, 134-138.

Wise, R., Chollet, F. Hadar, U., Friston, K., Hoffner, E., & Frackowiack, R. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain, 199*, 1803-1817.

Synthesis and Commentary: Toward a Cognitive-Science Perspective for Scientific Problem Solving

Rita W. Peterson
Department of Education
University of California
Irvine, CA 92717

Introduction

Teaching people to solve problems has always been somewhat of a puzzle. Many in the general public assume that some individuals have a natural talent for solving problems while others have only modest talent, and still others have limited problem-solving talent -- an impression undoubtedly based upon recollections of their own school experience. Nowhere are differences in problem-solving performance more evident than in classrooms where the teacher's job is to teach students how to solve problems. With regularity teachers report that some students fail to understand how to solve problems in spite of their best instructional efforts, while other students "see" how to solve problems before the teacher has completed introductory instruction. The consistency of teachers' reports and national test results confirming variations in students' problem-solving performance has led the National Science Foundation and other funding agencies to support studies that promise a better understanding of problem solving as an intellectual process, and experimental programs that hold promise for improving problem-solving performance.

In the interest of furthering knowledge about problem solving, the authors of chapters in this monograph have shared their expertise on scientific problem solving from a cognitive-science perspective.

The field of cognitive science itself developed rapidly, owing its progress to a diverse community of scholars which included anthropologists, cognitive psychologists, computer scientists, linguists, philosophers, and others. Their interdisciplinary endeavors have provided new models for understanding problem solving as a cognitive process, and at the same time, reinforced the idea that advances gained in one field of inquiry often suggest models, concepts, tools, or technologies that can be adapted to explain phenomena in another field of inquiry.

In that tradition, this final chapter opens one more window to challenge assumptions about scientific problem solving. But first, three dominant themes that are heard throughout the volume are identified, considered, and then woven into a series of empirical questions that are, at the end of the present chapter, placed in a neuro-cognitive framework

Chapter in Lavoie, D. R. (Ed.). (1995). *Toward a cognitive-science perspective for scientific problem solving: A monograph of the National Association for Research in Science Teaching, Number Six.* Manhattan, KS: Ag Press.

to suggest new lines of inquiry concerning the range of human capacities for solving scientific and other kinds of problems.

Description and Analysis of Dominant Themes

There are many conceptions of problem solving, and the richness of the examples in this special volume makes clear the fact that scientific problem solving is viewed not as a unitary ability or behavior controlled by a limited number of factors, but rather that scientific problem solving is considered to be multidimensional in nature by all of the authors. Through a number of themes running throughout several chapters, one sees the varied conceptual models and approaches to research which have shaped and characterized the multidimensional nature of scientific problem solving. Three dominant themes will be discussed here; they are *Constructivism*, the *Cognitive Structure of Knowledge*, and the *Contextual Nature of Problem Solving*.

Constructivism

The constructivist nature of scientific problem solving is an important theme that one finds embedded in accounts of recent research and in research which has a rich history of inquiry spanning nearly half a century. One hears harmonics of Piaget's and Inhelder's (1958) original protocols which illuminated the world-view conceptions held by children and adolescents, as Piaget and Inhelder provided readers with their first glimpses of the evolving constructions of reality held by children and adolescents who solved problems to explain natural phenomena.

The essence of the constructivist theme is captured in Wheatley's opening statement, "Constructivists see the individual as trying to make sense of [his/her] experiences." To illustrate, Wheatley describes students who, when given non-routine problems involving uncertainty, more often pursue irregular paths of inquiry than they follow organized, linear problem-solving procedures described by Polya and others. This same theme is heard in Roth's discussion of Lave: "...individuals experience themselves as in control of their activities, interacting with the setting, generating problems in relation with the setting and controlling problem-solving processes." In these and in many other examples throughout this volume, one finds the embodiment of constructivism in descriptions of individuals, whether on their own or in the company of others, who are actively engaged in inquiry and the construction of meaning from their inquiry and experience. A hallmark of the constructivist view portrayed here is the consistent agreement among authors in their interpretation of those recognizably distinctive behaviors, associated with students solving problems, as evidence of individuals who are constructing knowledge for themselves.

A view frequently expressed is one in which students' problem-solving success is influenced considerably by their opportunities to explore the world around them on their own terms -- phrasing and rephrasing problems; engaging in manipulations of ideas, objects and materials; and testing the vitality of their highly personal *constructions* of reality. Through such exploratory opportunities students appear to develop confidence in their abilities to frame problems, to find answers to questions and problems, and to learn more about the world through problem solving (see this monograph: Doran, Helgeson, & Kumar; Lavoie; Linn & Clarke; Roth; Hauslein & Smith; and Wheatley).[1]

The constructivist theme also is seen in the comparisons of expert and novice knowledge. "Experts differ from novices in both the quantity and organization of their knowledge, consistent with Piaget's notion of assimilation and accommodation," Hauslein and Smith observe. In this same vein, Kuhn (1968), in another volume, argued: "we learn science by accepting collections of concrete problem solutions." One interpretation of Kuhn's term "accepting" is "assimilating or accommodating" solutions. Willson hints at this interpretation when he speculates about models of "accretion" and "restructuring" knowledge.

One of the most intriguing curiosities about discussions of the constructivist theme in this volume is the unexpected silence concerning the developmental nature of problem solving during childhood and adolescence. From a cognitive-science perspective as well as a pedagogical perspective, it is important to know how problem solving capacities change from early childhood through late adolescence. Most of the studies discussed here focus on students aged 11 to 20 years, with few examples of the problem-solving capacities of children between the ages 5 and 10 years. Doran, Helgeson & Kumar provide an exception, with the K-12 model by Strang, Daniels, and Bell which suggests ways to think about problem complexity.

Cognitive scientists have had a long-standing interest in the study of children's linguistic and sociological development. Science educators might consider linking questions about the developmental nature of problem solving to the rich databases that have resulted from the study of children's linguistic and social development.

The Structure of Knowledge

A second dominant theme to reflect a cognitive-science perspective of problem solving is that which focuses on the nature of the

---

[1] All reference to authors in this monograph are made by name without a date.

internal organization of knowledge associated with solving particular problems or classes of problems. The primary dichotomy commonly employed to discuss the structure of knowledge is one which separates the identifiable elements of knowledge in terms of what one knows (meaning the facts or concepts that one knows in a domain), commonly referred to as *declarative knowledge*; and what one knows how to do (including various intellectual processes, science processes, metacognitive strategies, and general and specific heuristics for solving problems), referred to as *procedural knowledge*. Later in this paper, it will be suggested that a third kind of knowledge be considered in analyses of problem solving behaviors; studies of children's *episodic memory* may have much to contribute to a broader understanding of the structure of knowledge as it pertains to problem solving in science.

Another dichotomy used in discussions of the structure of knowledge focuses on the quantity of one's knowledge (i.e., how many things one knows) and the organizational structure of one's knowledge (i.e., its cohesiveness, congruence, and inter relatedness of ideas and processes). This latter dichotomy is useful because both the quantity and the structure of knowledge are observed to differentiate between experts (or successful problem solvers) and novices (or unsuccessful problem solvers), according to Hauslein & Smith; Lavoie; Doran, Helgeson & Kumar; and others.

The tools and technologies described in this volume are a valuable first step toward understanding the structure of knowledge; many were created to assess declarative and procedural knowledge related to problem solving, and were designed to assess both the amount and the organization of that knowledge. For example, concept maps or concept trees and semantic networks are standard approaches for the assessment of declarative knowledge; they provide information about both the amount and organization of one's declarative knowledge in particular domains. Similarly, one finds descriptions of a variety of standardized approaches for assessing procedural knowledge that is associated with problem solving, including written tests and lab-like performance tests that focus on science process skills (e.g., observing, comparing, applying, analysis, etc.), logical thought (e.g., seriation, controlling variables, proportional reasoning, probability, etc.) and other problem-solving strategies and heuristics.

Even though the present levels of validity and reliability of some instruments are of concern (Hauslein & Smith), exploration should be encouraged during this early stage in the study of the structure of knowledge.

Still in flux are a number of very interesting ideas about the structure of knowledge, and of particular interest are those ideas concerned with restructuring declarative and/or procedural knowledge (refer to this monograh: Hauslein & Smith; Doran, Helgeson & Kumar; Lavoie). Information-processing paradigms presume that the brain attends to and selects incoming information from the environment, interprets the new information within a context of existing knowledge, and constructs new meaning by integrating the new information into existing structures. Such paradigms also presume that existing cognitive structures are either expanded, revised, or destroyed. The connectionist model, with a neurological basis, suggests that connections between neurons and neural networks are either strengthened, weakened, or extinguished as individuals respond to environmental demands. However, neurocognitive-educational dialogue is still in its infancy, and explanations concerning the neurological manner by which individuals integrate new knowledge with existing knowledge, or how individuals "discover" new conceptual relationships as the brain restructures knowledge, are rare. As the 1990's "Decade of the Brain" continues, one can expect to find more information from the neurosciences that can be related to classroom learning and teaching.

Perhaps one of the greatest benefits still to come from future brain-behavior studies and R & D projects will be a more complete understanding of the neurological basis for the many varied processes that are involved in problem solving. One day soon, large numbers of teachers may learn to recognize instances when a slight neurological impairment or even a normal neurological variation may be influencing the problem solving capacity of a student. Small numbers of teachers currently are being trained to anticipate the possibility of suspected neurodevelopmental weaknesses or variations and to modify their teaching with alternative strategies to accommodate such students, with positive results (Peterson, 1994).

Another area where interesting ideas are still in flux involves the nature of the dependent or independent relationship between declarative and procedural knowledge in problem solving, as discussed by Hauslein & Smith:

> "[The] ability to recognize when certain knowledge is appropriate and to apply that knowledge successfully appears to depend on how well the successful problem solver's declarative knowledge and procedural knowledge is integrated.... In simple terms, certain problem solvers tend to be unsuccessful, not only because they lack certain declarative and/or procedural knowledge, but also because their knowledge is not richly interconnected. They cannot distinguish relevant from irrelevant problem information that facilitates recognition of problem types. "

Among the most extensive and significant studies of the relationship between declarative and procedural knowledge are those case studies reported by Squire, Press, & Amaral (1989) and Squire & MacKinnon (1989) in which they describe amnesia patients whose declarative and procedural memory are found to be independent. While it is generally assumed that declarative and procedural memory interact and are highly dependent in healthy adults, Squire's studies clearly demonstrate that these two kinds of memory can be disconnected as seen in patients who fail to recognize that they have seen and performed a task daily (declarative memory), but can solve the problem in the task more quickly each day (procedural memory)

Squire's studies (1987) suggest the value of, as well as a mechanism for studying both declarative and procedural memory associated with problem solving in normal, non-brain injured students. In this present volume, Roth's studies of healthy high school students who are left to solve problems of their own formulation or choosing in open-inquiry classrooms, suggest a productive environment in which to study or to search for examples of surface procedural knowledge as well as deeply-embedded procedural knowledge that may be operating at subliminal levels.

The Contextual Nature of Problem Solving

A third dominant theme that reflects a cognitive science perspective of problem solving is one which focuses on the contextual nature of problem solving. Looking across chapters, one finds the context of problem solving has been dichotomized as two worlds: *One world is the school classroom;* and *the other is the everyday world outside of school.*

In the classroom where problems are structured by teachers, curriculum developers or researchers, one finds described a full range of problems varying in dimensions such as: (1) formality-informality, (2) familiarity-novelty, and (3) simplicity-complexity. Along the formal-informal continuum of problems and solutions one finds many references to the disparaged textbook problem and "cookbook" laboratory exercise; a few examples of less formal teacher-directed, non-routine problems; and still fewer examples of the least formal open-inquiry science classroom where students generate and solve their own problems (Doran, Helgeson & Kumar; Lavoie; Linn & Clarke; Roth; Wheatley). In contrast to the world of school classrooms, the world of everyday problem solving receives far less attention except for the few examples such as Lave's studies of people solving mathematics problems in supermarkets.

In the ideal future, one might wish for a continuum of compelling examples from both school and non-school settings to soften or ease the sharp contrast between the two worlds of problem solving. Since classroom problem solving has been characterized as having a "brittle or contrived nature," a continuum of authentic examples that might be generated by science educators has the potential to provide new insights into the nature of problem solving (procedural knowledge) transfer. Such a continuum of exemplars ideally would vary in degrees of formality, familiarity, and complexity, as well. Moreover, if one wishes to foster the transfer of classroom problem-solving knowledge to everyday situations, then the conditions for which transfer is targeted must be clearly understood. One way to advance understanding of these conditions is to identify and describe the range of conditions for which transfer is expected, and to study a collection of authentic examples that fulfill those conditions.

To that end the following example is offered; it is based on an authentic account written by the mother of an elementary school child, and published in *The Los Angeles Times* about a year ago.

> On a trip to a zoo in Orange County, California, a mother and her daughter became so engrossed in explorations that they did not realize they were being locked inside when the zoo closed. Upon discovery that the tall metal entry gates were locked, the mother called out for help. When no one responded, the mother panicked and began shouting at the top of her voice for help. Her daughter realized that her mother was upset and said, "Mother, at school in my science class we learned that, when you have a problem to solve, you should sit down and make a list of all the possible things you could do. Why don't <u>we</u> do that?" The mother dismissed her daughter's suggestion as a distracting tale from school; she now worried that through her negligence she and her daughter would have to spend a long night in the zoo with caged wild animals. In a frenzy, the mother continued to shout for help and again the daughter suggested making a list, only to be ignored. Quietly the girl asked if she could have a piece of paper and a pen. The little girl began making a list that included ideas like: "look for a telephone, try climbing over the gate, look for a night security guard, try to find another way out." Between her mother's shouts, the girl began to read the list and asked if she could try some of the ideas on her list. Her mother stopped for a moment and was astounded by the possibilities that could be tried. She joined her daughter in pursuing the alternatives, willing to add to the list if necessary, until they found a way out of the zoo. Their efforts were successful. The next morning the mother went first to the school to thank her child's teacher for such valuable teaching, and then called the newspaper to recount the tale and her appreciation for her child's education.

This "locked in the zoo" account illustrates several attributes for problem solving exemplars that might be useful for a school/non-school continuum: (1) the student recognized an instance where a problem-solving strategy learned in one setting (in this case, school) could be applied in a new setting; (2) the student applied the knowledge; (3) the student described (spontaneously, in this case) the suitability of the strategy for the new situation; and (4) the student's application of the strategy to a new situation was successful. Exemplars would be useful even if they lacked all four attributes. Recalling Wheatley, a basic heuristic is "Do something." In this spirit, the zoo example is offered.

The Formulation of New Questions and Lines of Inquiry

> H.L.Menchen is credited with the observation:
> **"For every complex problem there is always a simple solution and it's wrong."**

Menchen's credited observation sets an appropriate tone for the task at hand, for the most enduring and interesting questions about learning and teaching science are always complex; and understanding scientific problem solving fits that standard for complexity. Certainly the authors contributing to this volume understand the complexity of scientific problem solving; the cognitive-science perspective in this volume has opened the way to a broader view of problem solving. In that same spirit, the perspective is now enlarged to include a neuro-cognitive view of problem solving.

In this section the themes of Constructivism, the Structure of Knowledge, and the Contextual Nature of Problem Solving are woven together by three broad questions, each accompanied by more specific questions which are formulated to cut across the thematic flow and form new lines of inquiry representative of a neuro-cognitive perspective. The questions focus on (1) the developmental nature of problem solving, (2) the relationship between specific neuro-cognitive processes and problem-solving success, and (3) the natural range of neuro-cognitive variations among typical populations of problem solvers.

The Developmental Nature of Problem Solving

How does problem solving capacity or performance change during childhood and adolescence?

a. Can developmental markers be found throughout childhood and adolescence which represent distinctive stages of problem-solving capacities? If biological clocks control the onset of children's language development, puberty and some other processes, then might biological clocks also trigger the development of problem-solving capacities; and might one be able to recognize developmental markers or mileposts associated with various stages of problem-solving development that would be helpful to those planning instruction?

b. Alternatively, can evidence be found that suggests the development of problem solving capacities is a gradual undifferentiated process throughout childhood and adolescence? If problem solving capacity does not exhibit a stage-like development, then studies of the structure of knowledge from childhood through adolescence might enable one to find domains of knowledge where children's declarative and procedural knowledge is sufficient in quantity and organization to make them <u>experts</u> in solving problems in those domains.

At the heart of *Constructivism* is the unanswered question: To what extent is children's problem-solving performance controlled by biological clocks or timetables which influence children's capacities to organize information, visualize alternatives, and choose among alternative actions to solve problems?

Textbooks and other curricula introduce problem solving in the primary school grades, as opposed to waiting until middle school or high school. Yet informal evidence from classroom teachers suggests that elementary school children generally have considerable difficulty solving word problems once the exemplar problem-solution has been removed; and children have even greater difficulty when word problems requiring diverse algorithms are presented together. Even middle school students have similar difficulties unless they are given clues.

From a neuro-cognitive perspective, one must ask if children's difficulties in solving problems suggest a lack of neuro-cognitive readiness. It is a question that has been considered during recent discussions of school readiness sponsored by the Maternal and Child Health Bureau in Washington, D.C. (Klivington, 1992; Peterson, 1992, 1994). But it is a question that was addressed much earlier by Piaget (1971) in his original formulations of what has become known as a constructivist perspective. Piaget assumed that children's cognitive development is influenced by their biological development, by socialization, and by the environment they experience. But to what extent were "stages" an indication of neuro-developmental readiness?

Scientists have been tracking the brain's development for many years and have provided brain images throughout the human life span, beginning with the prenatal period, as soon as brain imaging technology was available. The first report of a positive relationship between actual brain growth spurts during childhood and early adolescence, and Piaget's data on the development of logical thinking, appeared twenty years ago (Epstein, 1974). Epstein discovered stages of rapid and slow brain growth in humans (and in all mammal thus far studied). He concluded then and now that, "The human brain growth stages occur at the onsets of all the major Piaget stages of reasoning development and, therefore, are probably the biological bases of those stages" (Epstein, 1993).

Soon after Epstein's first reports, other scientists reported indirect evidence of differences in brain growth or development that were related to performance on cognitive tasks. Kraft and others (1980), for example, found evidence of age-related differences in brain maturation in a part of the brain that connects the right and left hemispheres, the corpus callosum, which allows the left (language) and right (spatial) hemispheres of the brain to talk to each other. When children aged 5 and 6 years (Pre-Operational) performed Piagetian tasks involving the conservation of substance, their left and right hemispheres did not communicate during three phases of the task: (1) listening to the task being explained, (2) performing the task, and (3) explaining the outcome of their actions. In contrast, children aged 7 to 9 years (Concrete Operational) exhibited interhemispheric communication during all three phases of the task. Kraft interpreted children's EEG data as evidence of age-related differences in corpus collosum development.

Other scientists have contributed greater detail ((Goldman-Rakic,1986, 1989; Picton, 1986; Thatcher, Walker & Guidice, 1987) regarding the relationship between problem solving and the brain. An important growth spurt is known to occur during early adolescence in an area of the brain known as the prefrontal cortex. Many who have studied this area of the brain know that it is involved in a significance way with problem solving. The prefrontal cortex often is described as the control center of cognitive activity. Just as the conductor of an orchestra controls each musician's contribution, and pilots control the flights of aircraft, the prefrontal cortex is responsible for pulling together information from various sources and evaluating it for actions like problem solving.

As one reads of brain-injured adults who, with prefrontal cortex damage, have lost their ability to connect categories of familiar information to solve problems that used to be simple for them, one is increasingly reminded of the struggles of children faced with similar problem-solving tasks. One interpretation of this research is that the substantial evidence of difficulty experienced by elementary school children when they try to solve word problems is this: in a neuro-developmental sense, they may not yet have hardwiring in the prefrontal cortex that allows them to pull together all of the relevant information needed to solve the problems they are given. While most children seem able to recognize, recall, name, and use symbols, they are less able to manipulate symbolic and other information in their heads -- that is, to "hold information out at arm's length" to evaluate as they attempt to select among alternative procedures or relevant information for solving problems.

Taking these and other brain-behavior studies into account (Gordon, 1988, 1989; Spelke, 1991; Tsunoda, 1989), one conclusion is that the developmental nature of problem solving warrants a place in future research agendas.

The Relationship Between Specific Neuro-Cognitive Processes and Problem-Solving Success

How are variations in capacities for attention, memory, language, visual-spatial reasoning, and temporal-sequential reasoning related to success in scientific problem solving?

a. How do various neuro-cognitive processes influence the acquisition and recall of declarative and procedural knowledge? How are strengths or weaknesses in attention, memory, language, visual-spatial reasoning or temporal-sequencing reasoning related to the acquisition or recall of declarative knowledge or procedural knowledge?

b. Which neuro-cognitive processes are critical for success in problem solving? Which neuro-cognitive processes are absolutely essential for success in solving specific kinds of problems? For solving a wide range of problems?

c. Do different kinds of problems require different neuro-cognitive processes? Can most kinds of problems in science be modified so that students with neuro-cognitive weaknesses can be successful working within a group or working alone?

Everyone enjoys teaching "good" students how to solve problems, and the failure of good students immediately galvanizes teachers into revisionist actions. If "average" students fail in their attempts to solve problems, teachers typically spend extra time with the entire class, trying to understand the reasons for students' difficulties. When only a few students fail consistently to solve problems presented in class, teachers respond in a very different manner. Elementary and middle school teachers typically "move on" and think of referring slow students to a counselor or resource teacher for special help. In contrast, few secondary teachers make such referrals or spend very much time analyzing the problem-solving performance of "poor" students, nor do they often make many modifications for students who are failing in their classes. Many secondary teachers report as many as 50% of their students in some classes are unsuccessful at solving problems. So common is the expectation that some students are excellent problem solvers, others are average, and still others are poor problem solvers within every cultural and linguistic group, that most secondary teachers cease to wonder about the phenomenon. (Though harsh, these generalizations are based on conversations between the author and hundreds of teachers who were asked to describe their response to the conditions described above, while enrolled in a mainstreaming course taught by the author.)

From a pedagogical perspective, teachers' lack of understanding of the reasons for students' success, partial success and failure, deserves serious attention. Further, from a neuro-cognitive perspective, the study of students' success and failure in scientific problem solving continues to be an important area for extensive and careful study, since many of the underlying neuro-cognitive processes that are thought to be involved in scientific problem solving (e.g., students' capacities for attention, memory, language, visual-spatial reasoning, and temporal-sequential reasoning) are poorly understood by most teachers.

It is often useful to consider the various different reasons for students' failures to solve problems. One kind of failure was described by Linn & Clarke, and Doran, Helgeson & Kumar, and seems related to the need for *situational assistance.* Another kind of problem solving failure appears related to *mismatches between* children's neuro-cognitive *readiness and the instructional demands* associated with problem solving (Levine, 1990, 1992a, 1992b; Peterson, 1993). A third and perhaps most troubling kind of student failure may be related to mild but *undetected impairments* in one or more neuro-cognitive functions that are required for solving problems. Consider the following examples.

<u>Situational Failure</u> is illustrated by a classic example in a study described by in this volume by Doran, Helgeson, & Kumar in which 70% of college students failed to solve problems until they were guided toward embedded clues; and thereafter, 75% of the students were successful in solving test problems. The need for coaching described by Doran, Helgeson & Kumar and by Linn & Clarke can be thought of as a need for an auxiliary expert system, an attentional selector to alert one to look for clues, or an external hard drive for extra memory to remind one to consider specific overlooked possibilities.

<u>Failure Due to Neuro-Cognitive Un-Readiness</u> is suggested by the research of Kraft, et. al. (1980) in the example of a kindergarten teacher who expects her 5-year-old students to solve problems which require Concrete Operations in a Piagetian task. A second example is

suggested by the research of Thatcher, et. al. (1987); that is, solving science or math word problems in novel contexts may be neuro-cognitively premature for most elementary school children, if such demands are dependent upon a major growth spurt in the prefrontal lobes. Obviously, brain-behavior studies are needed to clarify the hypothesized relationship between brain development and children's ability to consider all of the relevant alternatives for solving problems.

Failure Due to Neuro-Cognitive Incompatibility may be suggested by students who consistently fail in classes. They may be the most likely individuals to have one or more mild neuro-cognitive weaknesses (e.g., in attention, memory, language, visual-spatial reasoning, or temporal-sequential reasoning) that interfere with their learning, given the instructional constraints created by the teacher and the curriculum. Such students might be successful if their teachers were aware of the weaknesses and modified their approach; but if not, students must adopt disguises to prevent humiliation from peers and teachers when they continuously fail to meet the demands of the teacher or curriculum, and fail to understand their own neuro-cognitive weaknesses (Levine, 1990). Students' disguises are relatively easy to recognize in elementary school, but by high school, adolescents have had several years to perfect their disguises and they typically are perceived by their teachers and parents as bored, unmotivated, indifferent, confrontational, or hostile. When they find their disguises too difficult to "fake" (maintain), they are absent. Studies of these causes of failure are rare.

Pediatrician Melvin Levine is an exception. His research with early adolescents who are failing in several school subjects suggests the presence of one or more specific neuro-developmental weaknesses (dysfunctions too minor to qualify for special education) in areas required for success in classrooms; these include: attention, memory, language, visual-spatial processing, temporal-sequential processing, fine motor control, and social cognition (Levine, 1990, 1992, 1993).

Memory impairment is a good example because the effectiveness of a student's memory affects how much of the appropriate stored knowledge (declarative or procedural) can be "called up one the screen" at one time, thus affecting the integration of new with old information. Similarly, the vividness of a student's *episodic memory* (i.e., that vivid recall of autobiographical events) may interfere with a teacher's effort to replace a student's naive science conception with a new conception of a natural causal event. Visual-spatial processing is another example; when a student's capacity to interpret visual or spatial information is limited, the student's ability to integrate visual or spatial information into existing knowledge is compromised.

Neuro-cognitive functions may also exist as strengths rather than weaknesses, and contribute to successful scientific problem solving. In the present volume, Glynn, Duit, & Britton make a valuable contribution about exceptionalites through their discussion of the use of analogy in problem solving. An additional example of exceptional visual-spatial reasoning was offered by John Maynard Keynes:

> His [Newton's] peculiar gift was the power of holding continuously in his mind a purely mental problem until he had seen it through. *(Essays in Biography*, 1933)

A Natural Range of Neuro-Cognitive Variations Among Typical Populations of Problem Solvers.

How can instruction accommodate the neuro-cognitive strengths and weaknesses of students if one wishes to advance scientific problem solving?

a. What is the natural range of variation among typical populations of students in their capacities for attention, memory, language, visual-spatial reasoning, and temporal-sequential reasoning? How are neuro-cognitive variations related to students' general success or failure in the classroom -- or in scientific problem solving?

b. Can neuro-cognitive strengths compensate for neuro-cognitive weaknesses? If students have weaknesses in attention, memory, language, visual-spatial reasoning, temporal-sequential reasoning, or specific combinations of neuro-cognitive weaknesses that affect their problem solving success, can teachers plan instruction that builds on students' strengths to compensate for their weaknesses in problem solving?

Visiting a middle school campus at lunch time provides an immediate impression of the enormous variation and diversity that are characteristic of early adolescence: in every direction one sees classic "normal distributions" in terms of students' height, weight, sexual development, and social-emotional development. With such obvious variations in physical development and behavior, it would be naive to think that students' brain-mind or neuro-cognitive development is uniform. But one discovers very quickly that teachers treat most of their students as though students had relatively similar capacities for academic success in their class -- or "doing passing work," and as a consequence, teachers commonly interpret differences in students' responses to instruction as a reflection of differences in students' effort, attitudes, or willingness to study.

In defense of teachers, it is easy to understand why they attempt to treat all students alike; the underlying message -- both ethically and legally -- must be that students are treated as equals in spite of differences in their physical development, appearance, cultural heritage, or native language. Gradually, the concept of "equal treatment" is being replaced with the concept that pluralism requires equal opportunities to learn and demonstrate knowledge and skill. After sitting in middle school classrooms for a period of time, one realizes that many good teachers have not been trained to recognize students' mild neuro-cognitive weaknesses, and mistake them for lack of motivation, lack of effort, or indifference; and then teachers act on their assumptions.

Gardner's (1983, 1993) theory of multiple intelligences offers some promise for change. Gardner described seven intelligences which are recognizable in most school populations: linguistic intelligence, logical-mathematical intelligence, spatial intelligence, musical intelligence, bodily-kinesthetic intelligence, interpersonal intelligence, and intrapersonal intelligence. It is clear that Gardner views scientific problem solving as requiring logical-mathematical intelligence: "Jean Piaget, the great developmental psychologist, thought he was studying *all* intelligence, but I believe he was studying the development of logical-mathematical intelligence" (1993, p.8). Perhaps if Piaget were able to discuss his view of scientific problem solving with Gardner today, Piaget might very well argue that scientific problem solving also involves *spatial intelligence*, as Gardner himself defines it: "Spatial intelligence is the ability to form a mental model of a spatial world and to be able to maneuver and operate using that model" (1993, p. 9). And if Michael Roth were to join the dialog between Piaget and Gardner, he might argue that *interpersonal intelligence* also plays an important role in scientific problem solving, as reflected in Roth's chapter in this publication.

Gardner's (1993) theory leads one to hypothesize correspondences between students' neuro-cognitive strengths and their success in academic subjects where those strengths are demanded, and correspondences between students' neuro-cognitive weaknesses and their failure in academic subjects where their weaknesses become a disadvantage -- or dysfunctional.

Neuro-Cognitive Framework for Scientific Problem Solving

The dominant themes throughout the chapters have been considered, and questions have been posed which cut across those themes. The challenge that remains is to integrate the several key ideas into a broader theoretical framework that reflects a neuro-cognitive foundation.

Ultimately, those who propose theoretical frameworks are required to justify their endeavors; and such justifications are expected to meet certain scientific standards including explanatory power, predictability, and usefulness. With noticeably less profound qualifications, the theoretical framework that follows simply houses two main ideas that involve natural timetables and natural variations; but the "housing" does have the advantage of being consistent with recent research and thinking in the neurosciences. Beyond that, no claims are made.

One last point deserves mention before turning to the theoretical framework, and that concerns the origins of this endeavor. Disturbed for some time by the disappointment expressed by excellent teachers, when their efforts failed to achieve what they expected, it gradually dawned on me that such occurrences might constitute an anomaly in the field of teaching. Thus, I returned to something I had read nearly 25 years earlier:

> [It] commences with the awareness of anomaly, i.e., with the recognition that nature has somehow violated the paradigm-induced expectations that govern normal science. It then continues with a more or less extended exploration of the area of anomaly. And it closes only when the paradigm theory has been adjusted so that the anomalous has become the expected. (Thomas Kuhn, *The Structure of Scientific Revolution*, 1968, pp. 52-53)

What will follow is a description of the two main elements of a neuro-cognitive framework, supported by several easily documented observations.

Natural Timetables and Natural Variation: Major Elements of a Neuro-Cognitive Framework

Two common phenomena underlie much of what is thought about students and classroom instruction; these phenomena guide much of what educators do. At an abstract level, the phenomena are related to *natural timetables* and *natural variations* . While *natural* is intended to suggest *biological* origins, familiar school contexts are the place to begin.

Natural timetables. Academic achievement, on the surface, appears to be incremental and hierarchical. In principle, teachers and other educators generally agree that some things should come before others, instructionally speaking, and that readiness for academic tasks is influenced by student age as well as previous experience. This agreement is so common among educators that they rarely discuss the question -- even though they may hold quite different views about what actually comes first and last or exactly when a particular concept or skill should be introduced. One can see the massive strength of consensus among educators when an "outsider" suggests introducing a high school level or university level curriculum to students at the middle or elementary school level. One also sees confirming evidence when an experienced teacher explains to a beginning student teacher that he or she "talked over the heads" of most of the class. A final example is seen in the gradual loss of faith -- or outrage --among educators that occurs when textbooks or standardized tests are selected by "outsiders" who have never taught the targeted students or subjects themselves.

It is not argued that consensus among educators is a good predictor of students' neuro-cognitive readiness for academic content. Rather, educators operate on the assumptions about age-appropriateness until they have reason to question their assumptions. The point is that one must re-visit assumptions about the age-appropriateness of any curriculum or instruction, and closely examine exceptions to the rule or the weight of negative evidence.

Negative evidence is slow to capture one's attention. This is illustrated by numerous reports that elementary school students have considerable difficulty applying what they know when they are asked to solve problems in novel situations. Normally, when well-trained teachers instruct, students learn. Yet in the case of elementary school students' problem-solving capacities, in spite of the substantial efforts of very talented elementary school teachers (including those who teach GATE or gifted and talented children) and in spite of continuous major investments by funding agencies in projects designed to improve children's problem-solving capacities or performance, the expected payoff is not evident. Since this age-related phenomenon is easily recognized in one segment of the population (generally grades 1 through 5 or 6) and not found in the rest of the population (typically from junior high school onward), it is reasonable to suspect a biological source.

During the next decade, educators are likely to become increasingly aware of neuro-developmental timetables which characterize the development of the brain in relation to the unfolding of cognition during childhood and adolescence. Scientists now know that different regions of the brain reach maturity in a complexly orchestrated manner, that the process of brain maturation takes much longer than was previously thought, and that there is no basis for thinking of neurological maturation as a linear or undifferentiated process. According to Bernstein and Waber (1991), the functional consequences of developmental (or natural) timetables are largely uncharted. Yet, with the availability of non-invasive technology for brain imaging, educators may anticipate the day when they will see neuro-developmental timetables which allow them to re-examine questions about norms for school-aged children, much like the developmental timetables that now exist depicting the developmental path for eye-hand coordination in relation to focal length and large/small muscle development, or the chronological timing of the onset of puberty with the development of secondary sex characteristics during early adolescence.

What segment of society could benefit more from neuro-developmental charts depicting natural timetables for brain-mind development than the 25% of the population engaged in schooling?

Natural variations: Individual differences revisited. Move now from the context of natural timetables which are thought to characterize entire populations, to consideration of differences in the talent or performance among individuals within each segment of the population. This second phenomena concerns the *natural variations* in human talent and performance that are commonly observed in classrooms and other school settings.

Scientists during the 1980's and early 1990's have found evidence of variations in brain structure and function which are associated with students' academic failure. As researchers learn more about the natural range of variations in students' capacities for attention, perception, memory, language, visual-spatial reasoning, temporal-sequential reasoning, and other neuro-cognitive processes, it will be important to re-evaluate long-held assumptions about bell-shaped curves and the students who seem habitually distracted, indifferent, unmotivated, bored, lazy, hostile, tardy or absent. Educators will be challenged to redesign instruction to accommodate natural variations that will be understood.

Finally, the two proposed biological foundations of the theoretical framework can be stated as follows:

All individuals share a common ontogenetic path of neuro-cognitive development. Capacities for attention, memory, and language follow natural and predictable timetables for development among children throughout the world. Unusual circumstances such as deprivations of food or sensory input may alter developmental timetables, but under normal child-rearing conditions throughout the world, the natural timetables for development of these and other basic neuro-cognitive functions are expected to be the same.

All populations are characterized by a natural range of variations among individuals. Variations in physical appearance such as in height, weight, facial features, hair color and texture, physical stamina, and perceptual sensitivities are documented for populations throughout the world. All of these examples have in common the fact that they are readily visible; one finds very little disagreement among observers. What are less well documented or recognized are neuro-cognitive variations among individuals in their capacities for attention, memory, language, and so forth, -- capacities that influence success in school.

It is thought that these two biological forces, *natural timetables* and *natural variations* can, but need not, lead to continuous mismatches between the instructional demands in science and other classrooms, on the one hand, and students' capacities to succeed, on the other hand. Clearly, these two biological forces have the potential to influence students' success in scientific problem solving. When either or both biological forces -- students' biological readiness for the instructional challenge, or each student's particular mind-brain architecture and chemistry -- are poorly matched to the instructional demand, the most predictable result is student academic failure to understand the instruction, and of special interest at this time, failure at scientific problem solving. It is in this larger sense that a neuro-cognitive framework may allow science educators to predict continuous mismatches of the kind that could lead to student failure -- at problems, in classes, in school.

Conclusions

To sum, the goal has been to advance understanding of scientific problem solving. This monograph has offered a comprehensive picture of current research and thought about scientific problem solving; and this final chapter has attempted to pull together a few main themes and several thought-provoking ideas, to raise questions that cut across chapters, and to propose a theoretical framework that might house the empirical questions raised.

What intellectual processes are essential for solving specific kinds of problems; and which other process - while not essential - might be advantageous for solving problems of various kinds or even all kinds? H. L. Menchen (if he was correctly quoted) was correct: there are no simple solutions to complex problems.

When one thinks about the level of successful problem solving competence that has been proposed for all Americans, the goal seems attainable from one perspective: so much is now known about scientific problem solving. But from a neuro-cognitive perspective, the goal seems impossible unless efforts are begun to understand and plan for the neuro-cognitive nature of problem solving from childhood through adolescence. Science educators could wait for someone else to chart the course, but who would begin, if not science educators? In a context quite different from the present one, a noted neuroscientist said, "Sometimes we have to pour our own foundations" (Squire, 1989). This monograph is a step in that process.

References

Bernstein, J. H. & Waber, D. P. (1990). Developmental neuropsychological assessment: the systemic approach. In Boulton, A. A., Baker, G. B., & Hisscock, M., (Eds.) *Neuromethods: 17, Neuropsychology.* Clifton, N.J.: Humana Press.

Epstein, H. T. (1974a). Phrenoblysis: Special brain and mind growth periods. I Human brain and skull development. *Developmental Psychobiology, 7*, 207-216.

Epstein, H. T. (1974b). Phrenoblysis: Special brain and mind growth periods. II Human mental development. *Developmental Psychobiology, 7*, 217-224.

Epstein, H. T. (1993). Personal correspondence regarding his new publication: *Cognition, the brain, and education.*

Friedman, S. L., Klivington, K. A., Peterson, R. W. (1986). *The brain, cognition, and education.* Orlando, FL.: Academic Press.

Gardner, H. (1993). *Multiple intelligences: The theory in practice.* New York: Basic Books.

Gardner, H. (1983). *Frames of mind: the theory of multiple intelligences.* New York: Basic Books.

Goldman-Rakic, P. S. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task: evidence for dependence on dorsolateral prefrontal cortex. *Experimental brain research, 74* (1), 24-40.

Goldman-Rakic, P. (1986). Setting the stage: Neural Development before birth. In Friedman, S.L., Klivington, K.A, Peterson, R.W., (Eds.) *The brain, cognition, and education.* Orlando, FL: Academic Press

Gordon, H. W. (1988). The effect of right brain/left brain cognitive profiles on school achievement. In Molfese, D.L.; Segalowitz, S.J. , (Eds.) *Brain lateralization in children.* New York: The Guilford Press.

Gordon, H. W. (1989). Hemisphericity and its validity for education, vocation, and cognitive style. *International Journal of Neuroscience, 47*, 47-55.

Keynes, J. M. (1933). *Essays in Biography.*

Klivington, K. A. (1992). "What do we know about natural timetables for school readiness? Emerging evidence from the neural and cognitive sciences." Conference for school readiness sponsored by the U.S. Department of Health and Human Services, Washington, D.C.

Kraft, R. H., Mitchell, R. O., Languis, M. L., Wheatly, G. H. (1980). Hemispheric asymmetries during six- to eight-year-olds' performance of Piagetian conservation and reading tasks. *Neuropsychologia, 18*, 637-643.

Kuhn, T. (1968). *The Structure of Scientific Revolution.* The University of Chicage Press. pp. 52-53.

Levine, M. D. (1993). *All kinds of minds: A young student's book about learning abilities and learning disorders.* Cambridge, MA: Educators Publishing Service.

Levine, M. D. (1992a). *Developmental demands over time: some evolving and cumulative school-related expectations.* Cambridge, MA: Educators Publishing Service.

Levine, M. D. (1992b). Readiness as an on-going process: what have we learned from "unready" older children? Conference on school readiness sponsored by the U.S. Department of Health and Human Services, Washington, D.C.

Levine, M. D. (1990). *Keeping a head in school: a students' book about learning abilities and learning disorders.* Cambridge, MA: Educators Publishing Service.

Peterson, R. W. (1993). *The search for neuro-cognitive models to guide research in science education.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta.

Peterson, R. W. (1994). *Are neurodevelopmental variations related to academic success in middle school and science?* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Anaheim, CA.

Peterson, R. W. (1994). School readiness considered from a neuro-cognitive perspective. *Early Education & Development, 5*,121-140.

Piaget, J. & Inhelder, B. (1958). *The growth of logical thinking from childhood to adolescence.* Translated by A. Parsons and S. Milgram. New York: Basic Books.

Piaget, J. (1971). *Biology and knowledge: An essay on the relations between organic regulations and the cognitive processes.* Chicago, Il.: The University of Chicago Press.

Picton, T. W. (1986). Attention and the Brain. In Friedman, S.L., Klivington, K.A., Peterson, R.W. (Eds.) *The Brain, Cognition, and Education.* Academic Press.

Spelke, E. S. (1991). Pioneering research challenges accepted notions concerning the cognitive abilities of infants. In Raymond, C., *The chronicle of higher education,* Jan. 23.

Squire, L. R., Press, G. A., Amaral, D. G. (1989). Hippocampal abnormalities in amnesic patients revealed by high-resolution magnetic resonance imaging. *Nature, 341*, 54-57.

Squire, L. R., MacKinnon, D. F. (1989). Autobiographical memory and amnesia. *Psychobiology, 17*, 247-256.

Squire, L. R. (1987). *Memory and brain.* New York: Oxford University Press.

Thatcher, R. W., Walker, R. A., Guidice, S. (1987). Human cerebral hemispheres develop at different rates and ages. *Science, 236*, 1110-1113.

Tsunoda, T. (1989). Hemispheric dominance in Japan and the West, in *The science of mind*, K. Klivington, (Ed.). Cambridge, MA: The MIT Press.

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").