Conceptual Knowledge Research Project
Southern Illinois University School of Medicine
Springfield, Illinois

Learning, Teaching and Testing for Complex Conceptual Understanding

Paul J. Feltovich
Southern Illinois University
School of Medicine

Rand J. Spiro
University of Illinois
Urbana-Champaign

Richard L. Coulson
Southern Illinois University
School of Medicine

04 October 1991

Technical Report No. 6

This research was supported in part by the Cognitive Science Program, Office of Naval Research, under Contract No. N00014-88-K-0077, Contract Authority No. NR4422547. This paper does not necessarily reflect the views of the funding agency. Reproduction in whole or part is permitted for any purpose of the United States Government. Approved for public release; distribution unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
04 October 1991

3. REPORT TYPE AND DATES COVERED
Technical Report 01DEC87-30NOV90

4. TITLE AND SUBTITLE
Learning, Teaching and Testing for Complex Conceptual Understanding

5. FUNDING NUMBERS
Contract No.: N00014-88-K0077

6. AUTHOR(S)
Paul J. Feltovich, Rand J. Spiro, and Richard L. Coulson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Conceptual Knowledge Research Project
Southern Illinois University School of Medicine
Springfield, IL

8. PERFORMING ORGANIZATION REPORT NUMBER
C.K.R.P. Technical Report No. 6

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Cognitive Science Program
Office of Naval Research (code 1142CS)
800 N. Quincy St.
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NR 4422547

11. SUPPLEMENTARY NOTES
A version of this report will appear as a chapter in Test theory for a new generation of tests, N. Frederiksen, R. Mislevy and I. Bejar (Eds.), Hillsdale, NJ: Lawrence Erlbaum Associates.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE
CX

13. ABSTRACT *(Maximum 200 words)*
Drawing upon our research which has revealed problems among medical students in their understanding of complex material, we propose in this paper that new visions of instruction and assessment are required if education is to promote deep and usable understanding of complex, difficult subject matter. In the main, we argue that instruction and testing should be congruent with the cognitive goals for students that are desired--that if what is desired is that students obtain accurate understanding, instruction and testing should focus on this; that if what is desired is that students be able to apply knowledge, instruction and testing should focus on knowledge application; that if what is desired is that students acquire a structure of knowledge that they will not easily forget, education should concentrate on building and assessing robust knowledge structures. We propose some desirable characteristics of learning, teaching and testing for achieving this combination of goals. Tied to this, we suggest that educational goals for understanding can be aided by knowing how understanding is likely to break down, and we present numerous examples of ways that students misunderstand.

14. SUBJECT TERMS
Learning; Teaching; Misconceptions; Understanding; Curriculum; Complex Subject Matter; Advanced Knowledge Acquisition.

15. NUMBER OF PAGES
49

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

To appear in N. Frederiksen, R. Mislevy, & I. Bejar (Eds.),
Test theory for a new generation of tests. Hillsdale, NJ:
Lawrence Erlbaum Associates, in press.

Learning, Teaching and Testing for Complex Conceptual Understanding

Paul J. Feltovich' Rand J. Spiro[a] Richard L. Coulson[b]

Cognitive Science Division, Department of Medical Education, Department of Psychiatry, School of Medicine, Southern Illinois University, Springfield, IL 62794.

[a]Department of Educational Psychology, Department of Psychology, Center for the Study of Reading, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Champaign, IL 61820.

[b]Department of Physiology, Department of Medical Education, School of Medicine, Southern Illinois University, Carbondale, IL 62901.

INTRODUCTION

Deficiencies in education and in the capabilities of students are gaining increased attention at all levels of the American educational system--from elementary schools to schools of professional education (GPEP, 1984; National Commission on Excellence in Education, 1983; National Science Foundation, 1982; Porter, 1989). Failures among students in the achievement of sound and useful learning of complex subject matter have also been identified in laboratories of cognitive science concerned with education, and the nature of these shortcomings, as well as their causes and possible remedies for them, is coming under increased investigation (e.g., Feltovich, Spiro, & Coulson, 1989; McCloskey,1983; Spiro, Vispoel, Schmitz, Samarapungavan, & Boerger, 1987; White, in press). Deficiencies in the learning of complex material that are widespread and widely recognized include three major types:

misconceptions and incorrect knowledge (WRONG Knowledge), the inability to flexibly apply knowledge in new situations (what was characterized long ago by Whitehead, 1929, as a problem of INERT Knowledge), and the lack of retention of knowledge that was acquired at an earlier time (LOST Knowledge).

Acquiring and retaining a network of concepts and principles about some domain that accurately represents key phenomena and their interrelationships and that can be engaged flexibly when pertinent to accomplish diverse, sometimes novel objectives, is a reasonable definition of understanding in that domain (cf. Bruner, 1963; Greeno, 1977; Gelman & Greeno, 1989; Feltovich, Spiro, & Coulson, 1989).

Even in educational settings where understanding, in this sense, is a goal, it appears that this is often not accomplished (Coulson, Feltovich, & Spiro, 1989; Feltovich et al., 1989; Perkins & Simmons 1988). Deep and useful understanding of complex educational subject matter is not commonplace and comes at a high price, if at all. This may be because conventional educational practices and methods of testing achievement, while sufficient perhaps for uncomplicated material and for low levels of cognitive processing, are inadequate for difficult material when flexible understanding is a goal. Educating and testing for understanding of complicated material may require special, directed effort that is so resource-consuming that it cannot be applied widely across the many concepts of a curriculum. However, if it is important that some hard topics be learned well, then it would seem that these efforts will have to be made for a subset of the most important conceptual clusters. (In the latter part of the chapter, we propose that one answer to the demanding resource investment apparently required for fostering understanding involves selectivity and the establishment of priorities in curricula--where effort and depth in both teaching and testing are tied to the importance and difficulty of concepts to be taught, as well as to the cognitive objectives desired for the learner.)

Drawing upon our work which has revealed limitations among advanced students in their achievement of understanding of complex material, in this chapter we propose that new visions of instruction and assessment are required if education is to promote deep understanding of complex, difficult subject matter (see also Nickerson, 1989). We advance some guidelines for the forms instruction and testing should take when the achievement of flexible understanding is a goal. In the main, we argue that instruction and testing should be congruent with the cognitive goals for students that are desired (a recommendation that may seem obvious, but that is not routinely honored)--that if what is desired is that students obtain accurate understanding, instruction and testing should focus on this; that if what is desired is that students be able to apply knowledge, instruction and testing should focus on knowledge application; that if what is desired is that students acquire a structure of knowledge that they will not easily forget, education should concentrate on building and assessing this kind of knowledge. Tied to this, we suggest that educational goals for understanding can be aided by knowing how understanding is likely to break down.

The chapter has four main parts. Advanced knowledge acquisition, education where the goals are mastery of complexity and the ability to transfer knowledge to new situations, is discussed in the first part of the chapter. Such learning is different from "introductory" learning, and it is argued that the objectives and practices of introductory learning are often at odds with, and may actually interfere with, those of advanced learning. With reference to research we have conducted on students' learning and understanding of biomedical concepts, the difficulty of achieving understanding of complex material is also discussed, along with characteristics of subject matter that contribute to complexity and proneness to faulty learning. Principles for the design of instruction to promote the goals of advanced knowledge acquisition are presented in the second main section. If testing is to be congruent with this kind of instruction, encouraging and reinforcing the same kinds of goals, it will have to have new foci and characteristics. Tnese are outlined in the third section of the chapter. This is followed, in the fourth section, by a brief discussion of the need for selectivity in curricula, so that the most important and difficult concepts can be given special attention in instruction and testing.

While the chapter may appear at times to focus on instructional practices, this is because instruction, learning, and testing are, effectively, so highly intertwined. Forms of testing that are utilized drive much of learning and instruction, no matter what form the "official" curriculum takes (Frederiksen, 1984). Furthermore, our recommendations for new approaches to learning and instruction will suggest new forms of testing. Given all this, it is necessary to implement systems of testing that are consistent with goals for learning and that, in particular, require for successful performance the kinds of cognitive activities and outcomes valued in the instructional process (cf. Frederiksen & Collins, 1989). Hence, points made about desirable instructional practice are also points about desirable characteristics of assessment, and vice versa. These correspondences are addressed throughout the chapter, but especially in the section which focuses directly on assessment. Thus the sequence the paper follows is cumulative: a section on what goes wrong in learning leads into a section that discusses remedies for the observed patterns of learning failure; the last main section, on testing, is a culminating response to the issues raised earlier.

THE GOALS AND LIMITATIONS OF ADVANCED KNOWLEDGE ACQUISITION

In our work, we have been interested in "advanced knowledge acquisition" (e.g., Spiro et al., 1987; Spiro, Coulson, Feltovich, & Anderson, 1988). This is learning that occurs beyond the introductory stage but before the attainment of expertise (that appears to require long years of practice and experience; Hayes, 1985). This phase of learning has special goals, characteristics, and challenges associated with the attainment of accurate and useful understanding of complex subject matter. As we argue throughout the chapter, these goals make unique demands on the design of effective instruction and testing. The nature of advanced knowledge acquisition and some of the challenges it provides for learning are discussed in this section.

Advanced Knowledge Acquisition and Its Relationship to "Introductory" Learning

Advanced learning and common forms of introductory learning differ in both the instructional goals for students and the forms of assessment used to determine whether these goals have been achieved. In introductory learning the primary educational goal is often exposure to large areas of curricular content ("coverage" of content), without much emphasis on conceptual mastery of knowledge (e.g., Porter, 1989; Spiro et al., 1987, 1988). In particular, students may not be expected to understand concepts deeply or be able to apply them because it is presumed that following exposure heightened understanding and knowledge applicability will be incrementally achieved sometime "later." The demands of assessment, in turn, are often confined to the simple effects of exposure, i.e., recognition and recall of information in roughly the way it was presented in instruction. There is much less attention to testing higher-order skills of thought and knowledge application (Fleming & Chambers, 1983; Morgenstern & Renner, 1984). At some point in the educational process the restrictive goals of introductory learning must be superceded; at some point students must be expected to "get it right." That is, students should be expected to attain an accurate and deeper understanding of content material, be able to reason with it, and to apply it flexibly in diverse, ill-structured, and sometimes novel contexts (Spiro et al., 1987, 1988). This is the stage of advanced knowledge acquisition. The requirements of flexible knowledge use, in particular, place heavy demands on conceptual understanding because of the ill-structured nature of many domains of real-world knowledge application (Feltovich, Coulson, Spiro, & Dawson-Saunders, in press).

By this we mean that numerous concepts are likely to be pertinent in any case of knowledge application within the domain and that the pattern of relevant concepts may differ across instances of application that are classified as being the same. (For example, clinical cases of "hypertension" are individually complex in that they involve multiple biomedical concepts, and the pattern of concept combination can vary substantially across cases.)

In addition to being different, the methods of education and assessment in introductory and advanced learning would seem, in some important ways, to be opposed to each other. For example, common strategies of simplification in introductory learning such as teaching topics in isolation from related ones (compartmentalizing knowledge), presenting only clear instances (and not the many pertinent exceptions), and requiring only reproductive memory in assessment are often in conflict with the realities of advanced learning--where components of knowledge are fundamentally interrelated, where context-dependent exceptions pervade, and where the ability to respond flexibly to "messy" application situations is required. We have found that these discrepancies between introductory and advanced learning often result in situations where the groundwork set down in introductory learning actually interferes with successful advanced learning (Feltovich et al., 1989; Spiro et al., 1987, 1988; Spiro, Feltovich, Coulson, & Anderson, 1989).

How have we arrived at these contentions about the possible inhibitory relationship between the goals and tactics of introductory learning and the requirements of successful advanced knowledge acquisition? In our laboratory, we have been studying medical students' learning, understanding, and application of biomedical science concepts that are centrally important, by consensus of medical school teaching faculty across the North American continent that we surveyed (Dawson-Saunders, Feltovich, Coulson, & Steward, 1990). Medical school (as well as other schools of professional education) would seem to be a prototype of an advanced knowledge acquisition setting. Students have generally had some prior exposure to what they are learning, and the expectations for advanced mastery are high. Nonetheless, our studies have revealed a substantial incidence of misconception of central concepts. These misconceptions often involve oversimplification, and many have an impact upon knowledge application (Coulson, Feltovich, & Spiro, 1989; Feltovich et al., 1989; Myers, Feltovich, Coulson, Adami, & Spiro, 1990; Spiro et al., 1989). The development of these misconceptions seems at least partially traceable to cognitive and instructional strategies of the sort found in introductory learning. Yet they often persist despite students' having eventually been exposed in some fashion to appropriate information. [The existence of strongly held misconceptions, despite usual classroom efforts at instruction, has been found for difficult concepts in other subject matter areas as well, e.g., physics (cf. White, 1984)]. Besides the persistence of specific oversimplifications, it appears that simplificational 'habits' of thought and learning acquired in introductory learning are carried over to advanced learning--a tendency that is reinforced by instruction which likewise continues to oversimplify.

The Problem of Oversimplification: Misconceptions Resulting from Reductions of Complexity

It is instructive to examine the nature of misconceptions students acquire because these are seen to have a direct bearing on our recommendations for testing and instruction. Previewing the kinds of claims that will be made later in this chapter, a detailed understanding of the ways learning can go wrong should provide a guide for how instruction should be done (to avoid those problems) and for what should be tested and how. Likewise, knowing what it is about the nature of subject matter that causes difficulty for students can provide focus for instruction--both for what should receive emphasis and for how this should be taught if students are to be successful.

As noted earlier, many of the deficiencies in understanding we have observed in students appear to result from a cognitive inclination (a disposition in thinking) to simplify complex material--an inclination that is sometimes supported by similar simplificational practices of education (including, as we will discuss later, testing). We have termed the general tendency towards oversimplification in learning and understanding the "reductive bias," and individual instances "reductive biases." Numerous forms of the bias have been identified (for more detailed treatments, see Coulson et al.,1989; Feltovich et al., 1989; Myers et al.,1990; Spiro et al. 1989). Examples of reductive biases and associated misconceptions are given next. (Note: In each entry, a reductive bias is described first, followed by an example misconception from one of the areas of biomedical science and medicine that has been investigated in our laboratory. Each reductive bias is characterized using a few descriptive statements. These descriptive statements are intended to represent variations on a theme that runs through each reductive bias. This list is only a subset of the reductive biases that we have identified in our studies of conceptual understanding.)

- In general, this first bias is a disposition towards seeing entities as more similar than they actually are **(Similarity bias):** treating new examples as exact replicas of prototype examples that have been presented; treating partial analogies between concepts as exact correspondences (Analogy treated as isomorphism); failures to discriminate among similar but subtly different concepts.

  Example: The physical density of blood (its mass property) is frequently treated erroneously by students as being the same as, and as having the same hemodynamic effects as, the viscosity of blood (its 'thickness' or 'stickiness').

  Example: A superficial similarity in the qualitative relationship between force production in an individual muscle fiber (which involves the length to which a fiber is stretched and the contractile force it can generate at that length) and force production in an intact ventricle of the heart (which involves the volume to which a ventricle is filled with blood and the force of ejection it can then produce) is overly reified, leading to the erroneous belief that the same fundamental mechanisms of force production are operative in the two situations.

- Treating dynamic, changing processes more statically; assuming that a 'snapshot' or temporal slice of a dynamic process is representative of its nature; treating rates and derivatives as though they were equivalent to their mathematical integrations. **(Static bias)**

Example: Changes in cardiac output, which is a rate of blood flow (change of position of volume/minute), are often treated by students as though they were changes of blood volume, leading students to believe, for example, that increases in cardiac output would propagate increases of blood volume, and consequently blood pressure, to the veins (when, in fact, increases in cardiac output lead to decreases in venous pressure).

- Assuming that a schema or general principle accounts for all of a phenomenon when, in fact, it accounts only for a small part; the whole of a system is like some known part of a system; a functional relationship has a single causal basis throughout its range. **(Uniformity of explanation)**

  Example: The "length-tension relationship," a function which describes the relationship between the length to which a muscle fiber is stretched and the force it can generate at that length, is assumed by students to have the same causal basis across all the lengths a fiber can achieve. In fact, the causal mechanisms responsible for this function are different at different parts of the possible range of lengths.

- Treating multidimensional phenomena as unidimensional or according to only a small subset of their dimensions. **(Reduction of simultaneously considered dimensions)**

  Example: The degree of contractile force that a muscle can generate is a product of several factors, including both physical structural factors and others associated with the degree with which the muscle is metabolically activated. In a widely-held misunderstanding of the underlying causal basis of congestive heart failure, students reflect their earlier instructional emphases by attending to the structural factors affecting contraction, to the detriment of appropriate consideration of the activational influences.

- Understanding phenomena from the point of view of a single theory, schema, or conceptual perspective, when multiple sources of explanation are actually required. **(Restricted perspective)**

  Example: Opposition to blood flow in the cardiovascular system (cardiovascular impedance) is interpreted by students from the perspective of obstruction; that is, in terms of factors which provide physical hindrance to the movement of blood through the vessels. This is a perspective which gives prominence to the hemodynamic concept of resistance. A richer understanding of cardiovascular impedance can be gained by adopting a perspective on impedance which focuses on energy production and depletion in the cardiovascular system. From this perspective, factors other than resistance which are not obstructional can be recognized as contributing opposition to blood flow by depleting energy produced by the heart, and thus making less of it available to produce flow forward through the circulation. These added factors include the need to accelerate and decelerate the blood mass (because the heart produces pulses of pressure) and the need to move blood into and out of the bulging of stretchy blood vessel walls.

- Treating continuous attributes and processes as though they were discrete **(Discreteness bias):** bifurcation of continuous attribute dimensions to their poles (Bipolarization); segmentation of continuous processes into discrete steps, with associated agents and acts (Step-wise bias).

  Example: When considering the maintenance of acid-base balance in the human body, acid states and base states are treated inappropriately as polar opposites, rather than as reflecting a single continuum regulated by multiple factors.

  Example: Continuous blood flow in the cardiovascular system is decomposed in thinking to a set of sequences and steps, causing students, for instance, to misunderstand relationships between output from and input to the heart.

- Treating concepts separately (and as separable) that are, in fact, highly interconnected. **(Compartmentalization)**

  Example: Pressure-volume relationships in blood vessels--relationships between the size of vessels and the pressure they contain--and pressure-flow relationships--relationships between blood pressure and the blood flow through vessels--are often addressed separately in instruction (for example, in different chapters in textbooks) to emphasize different pedagogical points: for instance, differences in blood 'storage' capacity between arteries and veins in the case of pressure-volume relationships, and the circulation of blood in the case of pressure-flow relationships. This lack of integration carries over to students' (mis)understanding of the physical opposition to blood flow (cardiovascular impedance), where understanding requires conceptual integration of the two kinds of relationships.

- Assuming that the same elements of knowledge combine in the same routinizable way for all instances of conceptual application that are of the same nominal type when, in fact, the pattern of pertinence and combination changes in different situations. **(Precompiled schema retrieval)**

  Example: Medical conditions that are all instances of "hypertension," high blood pressure, can vary in their etiologies, contributing factors, and, most importantly for the present discussion, in the concepts from the biomedical sciences necessary for their understanding. Depending on particular circumstances, pertinent concepts can range over those associated with physical resistance to blood flow, the stretchability of blood vessels, the physical mass associated with blood, the volume of blood in the cardiovascular system, and various kinds of mechanisms of hormonal balance and imbalance (and their effects on such things as heart rate). Students' understanding of hypertension suffers from an overly restricted view of the variability of concepts and principles germane across instances of hypertension and, correspondingly, an overly homogeneous perception of the factors relevant across different instances (largely overemphasizing the applicability and uniformity of application of the concept of physical resistance).

- Assuming that a concept always applies in the same way when, in fact, its uses are often linked by only a general family resemblance. **(Uniformity of application)**

  Example: The relationships that describe the (interlocking) regulation of heart function by the vasculature and the regulation of vascular function by the action of the heart (the so-called "Starling/Guyton" relationships) do not apply in any universal manner; they differ in their application, depending on numerous conditions, including the operative blood volume in the cardiovascular system, the contractility (strength of pumping) of the heart, and the degree of stiffness of the vasculature. Students have difficulty accounting for this variability of application, and this contributes to their difficulties in understanding the relationships between outflow from and inflow to the heart.

- Assuming that if a concept is applicable in idealized conditions, separated from its natural situation of occurrence, it will be applicable in more natural and realistic contexts **(Extirpation)**. Treating different aspects of a topic of understanding as independent, as able to be treated separately and then 'additively' reassembled (thereby, in actuality, missing conceptual interactions). The conceptual whole is equal to the sum of its component parts **(Atomization/Insulation from synergism).**

  Example: Many of the concepts that apply to the activation and contraction of a single muscle fiber when it is isolated in the laboratory from its natural context as part of a complex of fibers in the heart, are erroneously extended by students to apply in the same way to the function of the intact heart. This kind of thinking is fundamental in supporting a widely held misconception of the basis of congestive heart failure.

- Ignoring causal dynamics: comprehension as description rather than explanation; causal mechanisms underlying the covariation of phenomena are glossed over. **(Superficiality/Insufficiency of causal-explanatory understanding)**

  Example: Because the heart becomes enlarged in congestive heart failure, this reinforces the commonly held misconception that the heart fails because individual muscle fibers within the heart become stretched to lengths at which they cannot generate adequate contractile force--when, in fact, the heart enlarges as a function of a complex set of consequences of its failing.

Patterns of Incidence and Acquisition of Oversimplification-Based Misconceptions

We have just described some examples of ways that students respond to complexity in subject matter and some of the ways this results in conceptual deficiency. However, it would be misleading to suggest that the consequences of the reductive bias amount simply to an isolated misunderstanding here and there. In fact, the reductive bias appears to participate in a larger scheme that makes the nature of what is acquired in learning more complicated and the task for education and testing more challenging.

Levels of misconception. First, the misconceptions that students acquire are of different types, which, in aggregate, affect all aspects of cognition. Misconceptions that have been identified exist at several levels of knowledge representation and reasoning, including the treatment of subject matter content, the mental representation of knowledge for use in thinking, and epistemological presuppositions about the structure and function of physical systems ('world views'). We have found examples at all three levels. Contentive errors often involve overgeneralization (or sometimes overdiscrimination). Areas of subject matter are seen as being more similar or more different than they really are (e.g., the Similarity bias--as when the effects of blood density and blood viscosity in the cardiovascular system are treated as the same). Errors in mental representation of subject matter also occur. For example, dynamic (constantly changing) processes are often represented more statically (Static bias--as when changes in blood flow, which is the rate of change of position of blood volume, are treated as changes in the simple magnitude of blood volume). Prefigurative 'world views,' the assumptions a learner makes about the nature of understanding in general (Feltovich et al., 1989; Pepper, 1942), also cause problems. An example is the common presupposition that parts of systems "add up" to wholes or that components of systems can be isolated from their naturally occurring contexts and still retain their essential characteristics (e.g., Atomization and Extirpation).

Networks of reciprocally supportive interactions. Second, misconceptions at all these levels may interact in reciprocally supportive ways (e.g., misconstruing one idea makes it easier to misconstrue another, and vice versa). Beyond this, sets of misconceptions can combine to produce yet other, higher order misconceptions. An example of this mutual bolstering of component misconceptions is the widely held misconception regarding the ultimate causal basis for congestive heat failure which has been used in this section as one source for examples of reductive biases. The primary misconception is that in heart failure the heart loses its capacity to pump an adequate supply of blood because individual muscle fibers of the heart become stretched to lengths at which they have a decreased ability to generate contractile force. This primary misconception can be seen to be a composite of four different, but related, component misconceptions that bolster each other and, in addition, provide multiple ostensibly reasonable paths to the same erroneous conclusion (Coulson et al., 1989).

Misconceptions have multiple sources. Third, multiple sources of influence appear to contribute to the development and maintenance of simplification-based misconception. In particular, cognitive biases toward simplification of complexity on the part of the learner seem to be reinforced by various instructional practices involving simplification (in textbooks, lectures, and so on) that extend beyond introductory learning into more advanced stages of learning, and also by some similar orientations and practices of biomedical science research. For instance, it is not uncommon instructional practice to use simpler (because they ignore the pulsatile pressure produced by the heart) constant-pressure hemodynamic systems to introduce basic properties of the cardiovascular system. This kind of focus neglects properties of the cardiovascular system that are due to the constantly changing pressure produced by the heart and thus contributes to misunderstanding of the concept of opposition to blood flow (where change in pressure and flow are particularly important). In another example, it is common to teach the topic of cardiac muscle function by focusing (at least initially) on skeletal muscle, which is more familiar to students and less complex than cardiac muscle. However, skeletal muscle is different from cardiac muscle on some of the very dimensions of muscle function that are important in the cause of congestive heart failure, the difference being such that an orientation toward skeletal muscle contributes to the misunderstanding of heart failure (Coulson et al., 1989; Feltovich et al., 1989). Simplification in instruction is sometimes mirrored in practices of biomedical science research (cf. Wimsatt, 1980), as when, for reasons having largely to do with experimental tractability, a particular form of cardiac muscle (cardiac papillary muscle) that is most like skeletal muscle is used in studies of cardiac muscle function, contributing to the perception of similarity of skeletal and cardiac muscle (and, hence, to the misconception of heart failure we have observed). Observation of ostensibly defensible reduction of complexity in the practices of authorities, such as teachers and researchers, probably lends justification and credence to simplification as a means for students to achieve understanding and to the understandings thus acquired.

Misconception affects knowledge application. Finally, in addition to their effects on fundamental accuracy of understanding, reductive biases of the sort we have described can carry over to inadequacies in the application of knowledge. This is illustrated in the following protocol excerpt from a student who, in one of our studies, is addressing a clinical case and demonstrates a form of the Bipolarization bias with regard to acid-base balance. The bipolarization leads to a most inappropriate clinical interpretation:

> "Well, first of all, (the patient has) severe vomiting and diarrhea. Vomiting you lose stomach acid; you're losing acid and that can cause alkalosis. However, severe diarrhea, you're losing bicarbonate, which can cause acidosis. So, if you're going to vomit or have diarrhea, it's better to do both of them at the same time, because you keep your pH balance in the middle."
> (From Myers et al., 1990, p. 157)

Among other flaws, this inappropriate 'prescription' results from viewing acid and base as bipolar opposites and from the mental extirpation of acid-base balance from its complex context of regulation, including the interaction of acid-base regulation with fluid regulation (e.g., the inappropriate prescription would almost surely add to problems of dehydration which are already likely to exist in the patient).

In this section we have presented examples of the ways students simplify complex subject matter and of the kinds of misconceptions that can result. We have also discussed some of the ramifications and extensions of this type of approach to learning, including its multilayered influence on cognition, the multifaceted reinforcement of such thinking from various practices of instruction and laboratory science, and the repercussions of oversimplification for knowledge application. What has been proposed is a complex system of knowledge acquisition, error, and cognitive limitation emanating from the artificial simplification of the real complexity of subject matter. In the next section, we address what makes concepts difficult and complex, so that they are likely to induce the reductive orientations to learning that we have discussed.

Notes on the Nature of Conceptual Difficulty: Structure and Process Issues

So far, we have been discussing reductions of complexity and the potential cognitive consequences of this kind of simplification. In this section we address characteristics of concepts that make them complex and difficult, so that they are liable to induce simplification and error, and so that misconceptions, once they develop, may be particularly stable (i.e., strongly held and difficult to emend). This can serve as a guide for the kinds of topics that are likely to require special handling in testing and instruction. A framework for conceptual difficulty and the stability of misconceptions is outlined below. The framework focuses on characteristics of the appropriate understanding to be achieved (e.g., the nature of cognitive processes required for understanding), characteristics of misconceptions associated with the correct idea, differences between the right and wrong ideas, and various kinds of external (outside the individual) sources of support that might exist for a misconception. The broad categories of the framework are given first, followed by examples and illustrations:

Characteristics of the concept as correctly understood and of its related network of component concepts. Because any complex concept is likely to be related to others in a conceptual network, this aspect of the framework pertains to the nature of the correct concepts to be attained, such as their individual difficulty, and to the structure of relations among them. The internal structure of some concepts is intrinsically more difficult than others (e.g., some concepts must be differentiated into more components than others). Similarly, some patterns of relationship among concepts pose more difficulty (e.g., recursive embedding).

Characteristics of the network of component misconceptions that make up an overall misconception; that is, characteristics of the faulty mental representation of the correct idea. This has to do with the nature of the component misconceptions that make up a misconception and, again, with the relationships among them. For example, the degree of reciprocal support among the misconceived components--the extent to which belief in one component eases belief in others (and vice versa)--will influence the stability of a misconception. Also pertinent is the relationship between the erroneous mental representation and the correct ideas. For instance, because of the disposition we have observed for individuals to adopt simple mental models, it would seem that simple misrepresentations of complex correct ideas will tend to be easily adopted and difficult to change (cf. Dember, 1991). In other words, if aspects of the correct idea to be acquired are quite difficult, and if their corresponding elements in the misconception are cognitively easier to handle, this should add greatly to the adoptability and stability of the misconception (because what is "understood" is satisfying and not cognitively strenuous, compared to the correct alternative).

Characteristics of the concept's typical treatment by authorities. This involves the extent to which popular media, scientific literature, and people who are presumed to know (e.g., teachers) promote the misconception or aspects of it: If important people are saying it, it must be true.

We propose in this scheme that a conception derives its relative propensity to be adopted and maintained partly from internal, cognitive supports--having to do with such things as the nature of the cognitive processes required for understanding and the characteristics of related knowledge structures--and from external supports, such as the way the concept is treated by textbooks and other authorities. Among other things, internal support depends on where a represented misconception stands on various dimensions of cognitive processing complexity, in comparison to where the correct conception falls on these same dimensions. A misconception will be more readily adopted and stable to the extent that it falls on the simpler ends of these dimensions of complexity, relative to the correct idea to be attained--for example, linear relationships in the misconception versus nonlinear ones in the correct idea. Some pertinent dimensions of difficulty and complexity are listed below:

- **Concreteness/Abstractness.** Are processes concrete and visualizable or abstract?
- **Discreteness/Continuity.** Are attributes and processes discrete or continuous?
- **Static/Dynamic.** Do properties or processes depend on fixed entities or values, or do they depend on change? Are characteristics of a process well represented by a fixed 'snapshot, ' or are characteristics of the process inherently entwined with change in the process from snapshot to snapshot?
- **Sequentiality/Simultaneity.** Do processes occur in a sequential, stepwise fashion, or are there aspects of simultaneity?
- **Mechanism/Organicism.** Are effects tractably traceable to the actions of agents (mechanistic), or are they the product of more holistic, organic functions?
- **Separability/Interactiveness.** Do different processes run independently of each other (or with only weak interaction), or are processes strongly interactive?
- **Universality/Conditionality.** Are there principles of function or relationships among entities that are universal in their application or validity, or are regularities much more local and context-dependent?
- **Linearity/Nonlinearity.** Are functional relationships among processes or entities linear or nonlinear?

Three additional sources of internal support involve the structure of existing or prior knowledge in its relationship to the correct and incorrect ideas. The first might be termed "p-prim congruence," after the construct of "p-prim" proposed by diSessa (1983). A p-prim is a fundamental belief about how the world works and is similar to what we, in our own work, have called a "prefigurative" conceptual scheme (Feltovich et al., 1989). To the extent components of a misconception are congruent with such p-prims and those of the correct interpretation are not, the misconception will be more readily adopted and more stably held; the misconception will seem intuitively right, while the correct idea will not. In addition, there will be a kind of 'mind-twister' characteristic to the correct notion because it will require a way of thinking that is discrepant with existing interpretive schemes. A second additional source of internal support is the existence of salient examples or analogies that appear to conform to the misconception. The more instances there are in memory of seemingly related kinds of phenomena that appear to be in concert with the misconception (or to its components), the more readily adopted and stable the misconception will be. The third knowledge-related source of support involves internal consistency or congruence among the components of the misconception. Important in this regard is the degree of reciprocation among components, the degree to which belief in some components makes belief in others easier.

In addition to the internal sources of support, another set of supports for a belief is 'external' to the individual, involving credence provided by authorities. Is a misconception commonly taught or suggested in textbooks, or implied by various aspects of biomedical science research? For example, one of the factors that contributes to the widely held belief in the misconception of heart failure we have used as an example in earlier parts of the chapter is that it is often proferred in medical textbooks and in clinical teaching.

A brief sketch of our framework for the analysis of conceptual difficulty and likely misconception stability has been presented in this section. The analysis, when applied to a set of concepts (those in a curriculum, for instance), can be used to help identify which among them are likely to need special attention in instruction and testing because they are difficult, likely to be misunderstood, and apt to be hard to change in students. However, for such areas of a curriculum, what is required of instructional practice if accurate, and, in addition, usable understanding is to be achieved? What is required of testing, if the soundness of understanding of complex subject matter is to be assessed? A proposal for the nature of the necessary instruction is taken up next. The instructional principles that are developed within the next section will be seen to provide dual service, forming the basis not only for instruction but also for the kind of testing proposed in the section that follows the treatment of instruction.

PRINCIPLES FOR INSTRUCTION IN ADVANCED KNOWLEDGE ACQUISITION

Taking into account deficiencies of understanding of complex material that we have observed in advanced students, we outline in this section principles for the design of instruction to achieve the goals of advanced knowledge acquisition, that is, the attainment of accurate, useful, and well retained understanding of complex material (cf. Cognitive Flexibility Theory--Spiro et al., 1988; Spiro, Feltovich, Jacobson, & Coulson, 1991). Desirable outcomes of learning and the common failures associated with each are taken up next, along with characteristics of cognition, educational practice, and subject matter that appear to contribute to the deficiencies and, hence, are to be avoided or surmounted in instruction and testing when sound and useful understanding is an objective.

Promoting Accurate Understanding and Overcoming Misunderstanding

As has been discussed, in learning complex and difficult subject matter students frequently acquire misconceptions that can be difficult to dislodge. In addition to our own work in biomedical understanding, such misconceptions on the part of learners have been found in subjects as diverse as arithmetic (Brown & Burton, 1978), physics (Caramazza, McCloskey, & Green, 1981), electricity (Gentner & Gentner, 1983), and climatology (Collins & Gentner, 1983). Often these misconceptions exist despite students' having been exposed in some fashion to accurate materials in instruction (although, as we have noted, some common practices of instruction appear to aid in the development of misunderstanding). Methods of addressing the important sources contributing to the development of misconception are presented next.

Discordant Prior Knowledge and Belief

Among sources of misconception are the prior models and beliefs related to a concept that students bring to the learning situation. New learning does not occur in a vacuum; it is tailored and influenced by what has been learned before--either in formal schooling or in ordinary experience. Such prior knowledge schemes can clash with those necessary for understanding new ideas, such that, for example: existing knowledge accentuates only a subset of facets of a concept to be learned, causing certain dimensions to be missed; or the new material is inappropriately subsumed to the prior knowledge; or prior knowledge causes new concepts to be seen as counterintuitive.

Counterintuitiveness is a major source of misconception. For example, many concepts from formal physics are difficult because they clash with ordinary experience. Many students erroneously believe that for an object to be moving, an applied force must be acting on it at all times because in a frictional world the need for such force is often accentuated. In an example from the cardiovascular domain we have studied, one of the things that makes it so difficult for students to gain a sound understanding of opposition to blood flow (the concept of cardiovascular impedance) is that it is very difficult for them to conceive that making blood vessels more easily stretchable (more compliant) could ever lead to greater opposition to the flow of blood by the vessels (which it can). To reiterate an earlier discussion, dysfunctional prior mental models and beliefs can exist at many levels of cognition, from beliefs about low-level subject matter content, to models of phenomena, to fundamental epistemological models about how the world works and is structured (diSessa, 1983; Feltovich et al., 1989).

Compounding the potential obstructive effects of prior wrong beliefs is the fact that they can be relatively opaque to challenge. By this we mean, first, that these beliefs can be implicit and tacit even to the student, hence hard to detect, and, second, that it is sometimes difficult to create circumstances that constitute challenge to them in such a precise way that the challenge feeds back to the appropriate sources of error. An instance of the latter was alluded to in the physics example given above, where much of the common feedback of everyday experience does not impinge directly enough on the misconception to be effective.

Instruction that assumes accurate, passive reception on the part of the student--that assumes, for example, that because the right information is presented to the student, correct ideas will be developed--is vulnerable to the effects of ingrained prior belief. Ideally, instruction should have a diagnostic component, in which students' preconceptions are made clear, as well as a more prescriptive component, providing directed and interpretable challenges to areas of knowledge that are potential hindrances to the achievement of appropriate understanding (cf. Green, McCloskey, & Caramazza, 1985; White, 1984).

Such considerations lead to the first of a set of design principles for effective instruction:

> Principle #1: Know what beliefs and interpretive models, germane to the concepts to be learned, students are likely to hold and the kinds of misconceptions they are likely to develop as a result.
>
> Principle #2: Provide directed challenges to misconceptions likely to be held or acquired.

Conceptual Isolation

Probably in any domain of rich complexity, but certainly in the biomedical sciences, concepts are highly intertwined and interdependent, such that the nature of any one depends on its interactions with many others. For example, the concept of cardiovascular impedance, opposition to blood flow, cannot be understood in isolation from concepts associated with cardiac muscle activation and contraction, regulation of cardiovascular flow, energetic metabolism, and several others. Furthermore, in a situation where concepts are in reality highly interdependent, students' misconceptions can likewise take on interdependencies, so that inadequate or overly compartmentalized understanding in one part of a cluster of pertinent concepts can have repercussions for the understanding of others (Coulson et al., 1989; Feltovich et al., 1989, on "spreading misconception").

Instruction that focuses on concepts in isolation promotes the idea that concepts are more independent and regular (i.e., less subject to variation in the context of other concepts) than they really are. In addition, because the true dependencies are not appreciated, isolated treatments restrict the richer understanding of related concepts and can even undermine their understanding more directly.

Principle #3: Focus on clusters of related concepts, not individual concepts.

Principle #4: Employ anticompartmentalization measures. Emphasize connection and combination, conceptual dependency, and conceptual variation across contexts.

<u>Singular Conceptual Aids</u>

Various kind of devices are used in instruction in an attempt to aid conceptual understanding. Common among these is the use of analogy. Employing an analogy in instruction allows the learner to import an intact cognitive structure to help in interpreting a new concept, rather than having to construct one more fundamentally. Analogies have been shown to have a powerful effect on learning (e.g., Collins & Gentner, 1983). However, when any single analogy is used to convey a complex, multifaceted concept, it is likely that the analogy will not cover all aspects of the concept (i.e., it will miss some aspects) and may actually be misleading with regard to others (Halasz & Moran, 1982). In fact, we have shown that analogies that help learners achieve the modest goals of introductory learning can interfere with the later mastery of more complete treatments (Spiro et al., 1989).

Instruction that utilizes singular adjunct representations as an aid to conceptual understanding runs the risk of restricting understanding of a new concept to only those aspects emphasized in the representation. Furthermore, the representation itself may induce some misunderstandings.

The use of multiple analogies and mental representations has been proposed as a means for alleviating the potential hazards to learning produced by single representations and as a means for enhancing understanding (Burstein, 1985; Burstein & Adelson, 1990; Spiro et al., 1989). Representations can be linked and meshed, such that conceptual aspects missed by one are addressed by others and so that misleading aspects are emended by others (Spiro et al., 1989). Similar approaches using multiple representations could be used to counter the maladaptively reductive effects of a single schema, prototype example, line of argument, and so on.

Principle #5: Multiple representations should be used for complex concepts.

The Reductive Bias

We have already discussed what appears to be a pervasive cognitive tendency on the part of students (and others) toward oversimplification of complex conceptual material. A number of particular cognitive processes involving simplification of complexity have been described, along with some of the kinds of misconceptions that can result (see section, "The problem of oversimplification: Misconceptions resulting from reductions of complexity"). This tendency toward simplification extends from the understanding of subject matter content itself, to the cognitive representational processes by which subject matter is coded for use in thought, to basic presuppositions of epistemology. It appears to be a powerful source of misconception.

Similar cognitive tendencies toward oversimplification have been established in the domain of probabilistic reasoning, where there, too, they often lead to error (e.g., Kahneman, Slovic, & Tversky, 1982). Our research has extended such findings into the domain of complex conceptual processing. It appears that human cognitive processing is such that there is a natural tendency to try to understand things in fundamentally simple ways (cf. Smolensky, 1986; Dember, 1991). This may be adequate when what is being learned is well structured and well-defined. It can lead to error and misconception when, in fact, material is complex and complexly structured, as it often is in the biomedical sciences and medicine, and in many other domains (especially those that involve real-world knowledge application).

The bias toward oversimplification also appears to extend to many educational, instructional practices (and even, as we have noted, to some practices of biomedical science research and reporting). For instance, it is fairly common educational practice to simplify complex material, at least in introductory learning, with the hope that complexity can later be introduced incrementally. Our investigations suggest that this strategy often 'backfires,' that initial, simplistic, and cognitively satisfying conceptualizations form obstacles to the progress of students. The basic reductive disposition on the part of the learner, when it interacts with simplificational strategies from instruction, can result in a powerful potion for misconception.

Principle #6: Do not oversimplify. Instead, utilize means to help students deal with the real complexity of things more tractably.

The reductive bias is an important influence that appears to pervade the learning process. It has a part in many of the other factors that contribute to ineffective learning and misconception, for example, the reductive use of analogy discussed above, in which a topic of instruction is overly identified with just those features accentuated by a powerful instructional analogy. It also appears to play a role in the development of knowledge that cannot be adaptively used. This topic is taken up next.

Promoting Flexibly Useful Knowledge in Contrast to Inert Knowledge

One of the goals of advanced knowledge acquisition is the development of knowledge that students can use in novel ways to address substantive problems and tasks. All too often, however, students may in some sense possess knowledge but not be able to use it in any way other than that in which it was originally learned. This problem of 'inert' knowledge is generally seen as one of the inadequacies of the educational process (e.g., GPEP, 1984; National Commission on Excellence in Education, 1983). The problem of inert knowledge is addressed here, along with design principles for instruction aimed at combatting this problem.

Passive Versus Active Learning

Good learning is not a passive process of reception of information; rather, it is an active process, on the part of the learner, of constructing new knowledge and incorporating it into what is already known (e.g., Bartlett, 1932; Bransford & Franks, 1972; Chiesi, Spilich, & Voss, 1979; Spiro, 1980). The more actively students process material, the more they embellish it with their own ideas, and the more that they question their own understanding, the more they learn and the more they retain from what they learn (e.g., Anderson, & Reder, 1979; Farr, 1987; Gates, 1917; Markman, 1981). Studies of good and poor learners (e.g., those who are able to solve problems versus. those who are not) reinforce the efficacy of activeness in learning; Better students extend what they learn, fill in gaps in what is presented, challenge what is presented, and question and test their own understanding of what they are learning (Bransford, Stein, Shelton, & Owings, 1982; Chi, Bassok, Lewis, Reimann, & Glaser, 1989).

Instruction that engages students in a passive manner is less likely to engender sound learning and knowledge that will be usable by the students than instruction that engages students actively.

Principle #7: Engage students in an active way. Students should be encouraged to manipulate and use the knowledge they are acquiring.

Purely Abstracted Knowledge Versus the Centrality of Cases

A major goal of the kind of instruction we propose is to promote understanding of important conceptual knowledge in such a way that it can be used in analyzing and working with realistic problems, for example, medical patient cases in the domain of biomedicine. It has been claimed that knowledge and skills of reasoning are highly intertwined. Reasoning is not a process separate from knowledge; instead, it is highly dependent on and intertwined with knowledge. Students' development of reasoning skills in a domain is most successful when these skills are tightly coupled with the manipulation and use of subject-matter content from that domain (Anderson, 1982; Barrows, 1983; Glaser, 1984; Wason & Johnson-Laird, 1972). In addition, there is a growing body of evidence that knowledge is in many ways bound to the contexts in which it is learned; that is, knowledge is more readily available for later use if the settings, cognitive processes, and goals active at the time knowledge needs to be used resemble those that were active when knowledge was acquired (e.g, Anderson, Farrell, & Sauers, 1984; Baddeley, 1982; Tulving & Thomson, 1973; Lave, 1988; Ross, 1987, 1989). Considerations such as these argue that conceptual knowledge should be acquired in close coupling with application and use, in the kinds of situations (cases of application) that will ultimately demand attention (cf. Barrows, 1983; Brown, Collins, & Duguid, 1989). This issue goes to the heart of what knowledge is--whether knowledge is something external from and engaged in use or whether it is something most appropriately thought of as a tool, constructed in the interaction between a mind and situations calling for action (Brown et al., 1989).

However, beyond such considerations, there is another reason why the use of cases should be central in learning complex material when the goal is knowledge application. This has to do with the ill-structuredness of conceptual knowledge in its relationship to instances of realistic application (Feltovich et al., in press; Spiro et al., 1988). As discussed earlier, an ill-structured domain is one in which many concepts, in interaction, are pertinent to an instance or case of knowledge application, and different patterns of concepts might be relevant across cases that appear to be alike or that are categorized as being alike (as in cases of, for instance, hypertension--see section, "The problem of oversimplification: Misconceptions resulting from reductions of complexity"). In an ill-structured domain, the guidance that abstractions and principles can provide for facilitating understanding and determining appropriate action is reduced. This is partly because there is likely to be great variability from case to case in the conceptual elements that will be relevant and in the patterns of combination among these. Single, general principles or concepts will not be sufficient to capture the workings of a case, and there will be variability across cases in the ways that any concept will be used and applied. Concepts whose uses vary greatly across contexts must be tailored to the particulars of application environments. Furthermore, it will be difficult to recognize from the apparent features of a case which elements of conceptual knowledge will be germane, because cases that appear similar may embody different elements of conceptual knowledge, and cases that appear different may embody common conceptual elements. When applications of a concept have such a complex and irregular distribution, it becomes impossible to specify, a priori, case features that should trigger the use of that concept. In such a situation, greater weight must be placed on examining the family resemblance (Wittgenstein, 1953) between a new case that is encountered and past cases that embody various sets of concepts.

This discussion is not intended to suggest that conceptual, abstracted knowledge is unimportant in ill-structured domains. Rather, it is to emphasize that when the linkage between conceptual knowledge and cases of application is complex and irregular, the role of abstract knowledge must be increasingly supplemented with and intertwined with elements of case-centered reasoning. The need for embedding conceptual learning within cases of application, while always important, is accentuated.

Principle #8: Couple cases of knowledge application to the learning of conceptual knowledge. Accentuate patterns of concepts involved in cases and their interactions, the variability of cases involving similar concepts, the similarity of cases embodying different concepts, and so forth.

Precompiled Schemas Versus Knowledge Assembly and the Noncompartmentalization of Knowledge

The irregularity of conceptual patterns across cases in an ill-structured domain makes it less likely that large-scale, preassembled structures of knowledge can be usefully imported for understanding and working with cases. Such intact structures assume routinizability in the domains where they apply. By definition, there is no routine in an ill-structured domain. Therefore, prepackaged "common denominator" structures will miss too much of the variability across cases. Ill-structuredness implies kinds of complexity and irregularity that are not compatible with gross prepackaging of knowledge and interpretive structures. Large, rigid structures are less useful than the ability to assemble and flexibly recombine smaller units of knowledge. In understanding a new case, many previous cases, parts of these cases (a new case may be "kind of like this old one, kind of like that old one"), and diverse pieces of conceptual knowledge may be helpful.

To support this ability for flexible and adaptive assembly, both conceptual knowledge and cases should not be compartmentalized. Concepts should be interrelated in diverse ways and not isolated in separate 'chapters,' and cases (as well as parts of cases) should be addressed in relation to other cases. When it cannot be clear in advance what knowledge will be needed or helpful--and in what patterns of combination--flexibility in response is aided by multiple systems of connections among cases and concepts, cast along multiple conceptual and clinical dimensions (Spiro et al., 1988). The larger the repertoire of this kind of connected case and conceptual knowledge available, the greater the support for adaptive knowledge assembly it should provide.

Principle #9: Utilize numerous cases in instruction, including small cases and parts of cases that convey important lessons.

Principle #10: Emphasize relationships among cases and between cases and concepts. Model for students the ways that conceptual knowledge is assembled from different sources to fit the needs of a new case. Contrast this with the inappropriate retrieval of prepackaged conceptual prescriptions for thought and action.

Surface Versus Rich Structural Indexing and Categorization: The Role of Multipurposing and Multiple Perspectives

Having to use knowledge to accomplish goals, to carry out tasks, appears to elicit and accentuate structural relationships and lead to indexing and categorization schemes characteristic of expertise. In this regard, a phenomenon has been observed in such diverse areas as maturation (Carey, 1985), the use of analogy (Gentner, 1988), "reminding," that is, the circumstance in which one thing reminds a person of something else (Ross, 1989), problem solving (Chi, Feltovich, & Glaser, 1981), and medical diagnosis (Feltovich, Johnson, Moller, & Swanson, 1984). What has been observed is that people with little experience working within a content domain notice, classify by, and have their actions driven by, apparent and superficial ("surface") features of situations. In contrast, with greater experience, noticing, categorization, and the basis for action come progressively to be driven by more covert, relational characteristics of situations ("deep" structures), including operative principles and concepts. For example, novice physics problem solvers see and solve problems alike that contain the same kind of objects (e.g., pulleys), while experts classify and solve problems alike that embody the same principles of physics (e.g., principles of energy), even though the problems may appear very different on the surface (Chi et al., 1981). This change from overt, surface feature orientations to a focus on complex embedded relational structures has been termed the "relational shift" (Gentner, 1988;1989). A growing interpretation of the relational shift is that surface features are in some sense easier to 'see' in a situation (e.g., a problem) than relational structure, and that it is in working with material to achieve purposes, to do something with it, that relations and relational structure are made more salient and important; in addition, different perspectives and purposes accentuate different aspects of situations as germane (cf Anderson & Pichert, 1978; Brown, 1989; Bransford, Franks, Vye, & Sherwood, 1989).

Rich, relational indexing schemes (for accessing knowledge), categorization schemes, and patterns of action are constructed from a perspective, in the course of use, tailored to purposes (e.g., the need to accomplish a task, or to evaluate a situation, or to teach). In ill-structured domains, and especially in biomedicine, where the linkage between surface features of cases and applicable concepts is irregular and rich, relational indexing and categorizations are not only particularly important, but also particularly difficult for the learner to construct. Surface orientations will often fail not only because, for example, cases that appear similar may involve different patterns of concepts and concepts applied in different ways, but also because many concepts are likely to be pertinent to any case of application. This is different from a more well-structured domain like physics (at least classroom physics), where indexing of relevant knowledge, classification of problems, and problem solutions can, for the expert, all be organized usefully around a handful of concepts (Chi et al., 1981).

Biomedicine, as an ill-structured domain, is particularly sensitive to changes in perspective and goal. The concepts that are relevant and the uses of these concepts may change, for instance, depending on whether a medical case is being addressed to understand its current state or to project its future course. The adaptability of indexing and categorization necessary in biomedicine is reflected in the indexing systems and categorical structures used by medical experts. These are characterized by multiple, redundant, and nonhierarchical (overlapping, latticelike) systems of knowledge organization, tailored to situations of use (e.g., Clancey, 1989; Feltovich et al., 1984).

Instruction that emphasizes limited perspectives on material and that encourages students to represent, classify, and use materials in only a small number of ways will not prepare them for the flexible application of knowledge across diverse and irregular new situations. "Knowledge that will be used in many ways has to be learned, represented, and tried out (in application) in many ways" (Spiro et al., 1988).

> Principle #11: Encourage the adoption of multiple perspectives, and the use and representation of knowledge in multiple ways. 'Revisit' cases and concepts, from different useful points of view, and for the purpose of achieving different useful kinds of goals.

Promoting Robust Knowledge and Overcoming Lack of Retention

The problem of forgetting or lack of retention of knowledge, if not more serious in biomedicine and medical education than in other fields, has probably had more attention directed to it in these areas than it has had elsewhere. This is because medical education has traditionally had separated "basic science" and clinical parts, with the basic science, conceptual part occurring in the early years of the curriculum before clinical training begins. The question of what from the basic sciences is retained by students for use in clinical training, and, in fact, what of it is useful and how, has been a continuing concern of medical education. More often than not, studies of the retention of basic science material by students in the later years of medical school have shown rather dismal results (e.g., Levine & Forman, 1973).

This issue of retention is not a simple one, as it involves consideration of the different ways in which having "retained" knowledge is measured, differences in the kinds of knowledge under consideration in different studies, how long after "acquisition" (learning) retention is measured, and fundamental notions about what the nature of learning is (for example, whether it makes any sense to say that at some point a body of knowledge was "acquired," as though this were some kind of discrete and finalized event, and that this acquired body of knowledge is at some later time retained, or recollectable). Issues such as these involved in thinking about long-term retention are not addressed in detail in this chapter. Instead, we point to the conclusions of a major study of the long-term retention of knowledge and skills, generalizing somewhat over the qualifications and conditionalities expressed there (Farr, 1987).

Knowledge and skills, especially complex knowledge and skills, appear to be better retained when conditions of the following sort hold (Farr, 1987);

(1) When material at the time of acquisition is processed deeply, embellished, and connected to and integrated with other knowledge--in general, when knowledge is not compartmentalized, but is richly structured and indexed.

(2) When complex material is "understood," rather than learned by rote: "When concepts, principles, and rules complement or supplement teaching of rote knowledge or facts" (Farr, 1987).

(3) When, and to the extent that, conditions of training and learning resemble those in which the learned knowledge and skill will be applied. (Notice that this condition for retention is made more complicated in ill-structured domains, where there are special demands for knowledge transfer that result from the likely variability between conditions of

initial learning and later use. This accentuates the need for using multiple cases, multiple perspectives, and multiple goals for knowledge use in the course of learning in ill-structured domains. This will help to promote both retention and transfer.)

Considerations such as those just listed can be recognized as being interwoven throughout the design principles discussed earlier in regard to promoting correct and usable knowledge. Hence, in addition to the functions of facilitating correct and usable knowledge, we would expect instruction that is in conformance with our principles of design to lead also to respectable retention of complex knowledge and skills.

GENERAL GUIDELINES FOR TESTING WHEN DEEP CONCEPTUAL UNDERSTANDING AND KNOWLEDGE APPLICATION ARE GOALS

So far, we have discussed a paradigm for conceptual understanding that: (a) embodies a reconceptualization of the goals of understanding for advanced students; (b) recognizes that those goals are often not accomplished; and (c) suggests that their attainment requires ways of thinking that are in large measure antithetical to cognitive approaches that are appropriate for the earlier, less demanding educational goals. Drawing upon the kinds of deficiencies in understanding that we have observed among advanced students, we have proposed a set of guidelines for instruction when sound and flexibly useful understanding of complex material is a goal. Given the acknowledged influence of the testing process on instructional practices and on students' approaches to learning, the goals and practices of instruction and testing must be in conformance if educational objectives are to be attained (Frederiksen, 1984).

In this section we outline some principles for testing in support of the goals of advanced knowledge acquisition. What we propose for testing has two major motivations. The first stems from the goals of advanced knowledge acquisition themselves--that students should understand complex material accurately and deeply, including the ability to use this understanding in the accomplishment of substantial tasks. If we want students to understand deeply, we must test for deep understanding. If we want students to be able to apply knowledge, we must test for substantial knowledge application. The second motivation comes from the deficiencies in understanding and knowledge application we have observed in advanced students. Testing for understanding of complex material will benefit from knowing where understanding is likely to break down. In addition to outlining the kind of testing we believe is needed, we also give some examples of the kinds of forms such assessment might take.

Testing That Is Focused Both Toward Failures of Understanding and the Desired Goals of Advanced Knowledge Acquisition

Tests of understanding should have a substantial focus on likely points of comprehension failure, of the sort we have discussed earlier in the chapter (see section on "The problem of oversimplification: Misconceptions resulting from reductions of complexity"). Tests should be constructed that specifically target elements of complexity and concomitant potential failure, in the same way that these provide focus for instruction. At the same time, testing should promote desired cognitive goals for students, that is, sound and useful understanding. Analogues in testing of the principles of instruction discussed in the last section include the following:

- **Misconception Undermining** (Instructional principles #1, 2): Test items should be crafted to address known misconceptions with regard to a body of subject matter. This pertains not only to subject matter content itself but also to known reductive ways of thinking related to subject matter--e.g., treating continuous attributes and processes as discrete. Such items can have both diagnostic value, eliciting the existence of misconceptions, and pedagogical value for correcting them.

  Example: Requiring students to discuss or predict the mechanisms of action of beta-blocking drugs, drugs which actually raise vascular resistance but are used therapeutically to reduce blood pressure, can serve both to diagnose and emend the widely held misconception that opposition to blood flow (cardiovascular impedance) is entirely resistance-based.

- **Noticing Differences** (Instructional principle #1): Items should allow for detection of comprehension failures that result from presuming too much similarity among entities (e.g., recognizing the limitations of analogies and prototype examples).

  Example: Since blood viscosity and blood density are different and contribute differently to opposition to blood flow (but are treated by some students as being the same), items which require students to predict or discuss the effects on blood pressure of changes to density and/or viscosity can test for lack of discrimination and can force the desired discrimination.

- **Conceptual and Instructional Integration** (Instructional principles #3, 4): Test items should have answers that require the integration of several component concepts or theories, especially when they are likely to have been treated separately in acquisition. Test items should have answers that require the integration of information on the same concept or theory that was likely to have been presented in distantly nonadjacent sections or 'chapters' of instruction at the time of learning.

  Example: The processes of hypertrophic adaptation, by which muscle responds to abnormal levels of stress, and the processes of diminished pumping ability characteristic of congestive heart failure are often treated separately in instruction and are not well integrated by students, partly contributing to a widely held misconception about the fundamental basis of heart failure (Coulson et al., 1989). Test items that focused on the time-course development of heart failure and its stages (rather than, say, any particular state of it)--items that involve changes in the disease, hypertrophic adaptations to these, further changes, further adaptations, and so forth--would require conceptual integration of the two kinds of processes for successful completion and could assess the extent of this integration in students.

- **Integration of Analogies, Conceptual Aids, and Prototype Examples** (Instructional principle #5): Test items should require assembling from acquisition materials an appropriate subset of presented analogies, examples, or conceptual aids that are relevant to a test question, and further should require using only the relevant parts of those analogies and examples.

  Example: A group of analogies relevant to muscle cell operation, such as oarsmen in a scull (related to the "cross bridge theory" of muscle fiber function, which describes the means by which contractile force is produced in a fiber), a turnbuckle (related to the "sliding filament theory," which describes the movement of structures of the fiber in relation to each other), fingercuffs (which relates to the action of collagen in providing resistance to stretch of muscle fibers and which, because of a set of anatomical and functional differences between skeletal and cardiac muscle, is more pertinent in cardiac muscle) could be presented and the student asked to choose which of two examples, e.g., a cardiac papillary muscle and a skeletal sartorius muscle, is best represented by the set and to explain why.

- **Reorganization** (Instructional principles #5, 11): Information that was presented in one way should be able to be appropriately reassembled in other ways. For instance, the same example may fit into different organizational compartments (multiple sortings of examples). Information that was grouped together for purposes of exposition in acquisition should be able to be regrouped by the student for newly prescribed purposes.

Example: The following four case descriptions are all instances where implications involving the Law of LaPlace are important but where each represents a problem of an entirely different origin and course of development:

1] A relatively benign mitral valve prolapse (partial failure of the input valve of the left ventricle) in a young woman.

2] A pediatric case of ventricular septal defect (an abnormal hole connecting the lower ventricular chambers of the heart).

3] A ventricular aneurysm (a bulging of the heart wall due to weakened or diseased heart muscle) in a postinfarction (after a "heart attack") old man.

4] A dissecting aortic aneurysm (a bulging of the wall of the body's main artery rather like a bicycle innertube bulging through a split in the tire) in a middle-aged man.

A test item could require students to detect the unifying principle of the Law of LaPlace in the group of different cardiac pathologies when the principle itself was probably learned in the context of the physiology of the lung (or, perhaps, in the context of volume overload hypertrophy and heart failure, contexts rather different from those addressed by the test item). This would require the student to decontextualize the principle from its original acquisition setting and to reconstrue it for new purposes.

- **Multiple Perspectives and Context Sensitivity** (Instructional principle #11): Tests should allow for different correct "main idea" answers depending on different perspectives (where the student is required to adopt different perspectives from which to answer).

  Example: The same (medical) clinical case, when addressed from different perspectives such as diagnosis (determination of what is wrong), treatment (trying to ameliorate the condition), prognosis (predicting how the medical condition will progress--untreated and under different courses of treatment), or the monitoring of intermediate efficacy of treatment (determination of whether treatment is 'on course' toward the desired end) can implicate different combinations of biomedical concepts as important for each purpose, or the same concepts used in a somewhat different way. Students can be required to address a case from different points of view such as these, identifying the biomedical concepts relevant in each instance, and describing how they are applicable and what they imply.

- **Underlying Causal Mechanisms** (Instructional principles #7, 11): Besides being able to answer questions about what is happening in a presented situation, the student should also be able to determine why it is happening. For example, for natural science phenomena, students who have comprehended material well should be able to go beyond mere description to answer questions involving prediction (what would happen next), postdiction (specifying what came before), determination of inconsistencies in descriptions (critique), experimentation (determining what information is missing and how conditions might be arranged to acquire it), and, especially, causal reasoning about the mechanisms by which a part of a situation is affecting other parts (Forbus, 1985).

Example: Two patients in heart failure could be described with identical current hemodynamics, for example, the same blood pressures, the same cardiac outputs, the same peripheral pulses, and so forth. One is an instance of old heart failure, reemerging, where the patient had previously been controlled with digitalis and diuretics. The other is an instance of new heart failure appearing in a hypertensive patient whose condition had been managed previously with the drug propranalol (a beta blocker). The student would be required to compare and contrast the current physiological states of the patients, the likely natural histories that eventuated in these states, and the projections of future course (prognoses), giving causal justifications based on the physiological, pharmacological, and pathological concepts and principles pertinent to these situations.

- **Problem Solving, Decision Making, and Educated Guessing** (Instructional principles 7, 8, 9): Test the ability to apply conceptual knowledge to actual cases or examples that are relevant to the concept or set of concepts being tested (knowledge into practice).

  Example: A set of clinical cases of "hypertension," embodying different ultimate physiological sources for the hypertension and, hence, engaging different patterns of concepts related to cardiovascular impedance (e.g., heart rate, blood volume, and inertial reactance in one case; hormonal constriction of the vasculature and resistance in another) can test students' ability to recognize the applicability of and to apply concepts associated with impedance. Attention is paid to students' prescriptions for treatment in the different cases, since for individuals with correct understanding of impedance treatments should be differentially sensitive to the different patterns of applicable concepts across the cases.

- **Scaffolding for Subsequent Comprehension and Transfer** (Instructional principles 7, 11): Examine the extent to which current understanding provides a scaffolding for comprehension of new material on the same topic ("new" knowledge as "prior" knowledge for new learning--see the last subsection of this section). Examine the extent to which current understanding supports the analysis of situations and the solution of problems not addressed in instruction.

  Example: Students' understanding of the principles governing the cardiovascular system can be assessed by requiring them to construct a critique of a journal article advancing a new theory of the pumping action of the heart in the propulsion of blood. Attention is directed at the accuracy and creativity of students' use of cardiovascular principles and concepts in this critique.

Example: Understanding of the principles of cardiovascular impedance can be tested by requiring students to apply them to airflow in the lungs, where they are directly applicable (but where the pattern of importance of concepts is somewhat different), or to any system in which anything flows, for example, information flow in an electronic mail system.

Some General Characteristics of Tests of Complex Conceptual Understanding

The implementation of these principles of testing requires that test instruments have certain general characteristics and be used and interpreted certain ways. Some of these are discussed in this section.

First, assessing comprehension of complex material will require long enough "stems" to build the requisite dimensions of complexity (in congruence with instructional Principle 6). Tests that will allow for targeted assessments of the kind just discussed (e.g., integration across nonadjacent sections/chapters of instructional material; interaction of multiple concepts relevant to the same topic; and so on) will unavoidably have to be much longer than is typical of common tests. [In medical education, for example, test 'stems' might take the form of simulations of entire medical patient cases (Norman, Muzzin, Williams, & Swanson, 1985) , and students may be required to deal appropriately with such cases in order to become physicians (Barrows, Williams, & Moy, 1987).] Complexity takes time to develop in a test "item." The price of shorter test items is an absence of just those properties in tests that are most likely to cause (and hence reveal) comprehension failure and that may, therefore, be most diagnostic of success in advanced knowledge acquisition.

Second, testing for understanding should be viewed as a montage of partially overlapping 'snapshots.' An important implication of the view of the goals of understanding as complex, multiply interconnected, multiperspectival, and so on is that testing must involve multiple approaches--a single pass of evaluation will elicit only a subset of the facets of understanding that are to be demonstrated. Thus, with a limited testing scope, it is possible for students who have only partial comprehension to have their understanding greatly overestimated. More importantly, a restricted scope of assessment will miss altogether the complex interrelationships that are at the heart of advanced understanding. Adequacy of understanding must be revealed by assembling a perspicuous montage from a series of partially overlapping assessment snapshots for the same material--no single "picture" (from a single assessment pass) will be comprehensive enough.

It should be restated that singular approaches are also likely to be inadequate, or even misleading, in instruction. For complex material, in both testing and instruction, it seems prudent not to do anything one way. Singular approaches are likely to be detrimental because they: (a) do not provide a wide enough 'lens' on the numerous aspects of the material to be taught or understood (b) are likely to miss the interconnectedness of the target material with other related material, and (c) reinforce a misleading orientation toward complex material, by suggesting that it is simpler than it really is.

Finally, testing for understanding partly requires testing the readiness of current understanding to serve as input to new understanding and to knowledge application. A somewhat paradoxical aspect of the paradigm we are describing is that the adequacy of the outcome of understanding must be determined in part by how well this output can be used subsequently as input (background knowledge) for later comprehension. A student usefully understands a concept or set of concepts when he or she is able to apply that understanding to support comprehension of new material on the same, or perhaps even a distantly related, topic (far transfer). In other words, our earlier points concerning knowledge output (the goals of understanding for advanced learners) and knowledge input (e.g., schema assembly for subsequent comprehension) cannot be fully separated. And testing must reflect this interdependency.

So, to have adequately understood, a learner must have done more than form a coherent representation of the material and reproduce it for a test: The student must have learned from the material in such a manner that what was learned will be flexibly usable in future comprehension situations.

A NOTE ON SELECTIVITY IN INSTRUCTION AND TESTING

It is often said that a typical medical school curriculum includes thousands of concepts (and this is likely true of other substantial curricula as well). The curriculum is dense and the pace is fast. There are, then, considerable pressures on both teachers and learners to move through material quickly, and simplification in both testing and instruction may just be part of a survival strategy. A curricular stance that values "coverage" is prevalent: "What will some future medical residency director think if one of our students has not even heard of (some topic)?" Yet, there are fundamental concepts that should be understood, and from our investigations it appears that achieving this takes considerable time and focused effort. One thing that would be helpful is a method for prioritizing curricular content and for linking educational methods to the cognitive goals held for students. If instruction and testing for deep understanding require intense, directed effort, as appears to be the case, then it is clear that achieving and assessing the achievement of solid and flexibly applicable understanding of thousands of concepts is an impossible goal.

First of all, it should be said that not everything in a curriculum needs to be understood at the same level of depth. In some areas it is sufficient to meet the challenge that "the resident should have at least heard" about some item of curricular content, in which case simple forms of instruction and testing (e.g., overview lectures, hand-out lists, and tests of recall or recognition) should suffice for the purpose. It is also true that not all concepts are difficult to understand. They may, for example, be well structured, relatively self-contained, congruent with intuition, and/or amenable to the reductive tendencies in cognition that we have described.

In such cases, standard means of learning and testing are, again, likely to be adequate. But, how is curricular focus to be achieved for those concepts that are truly important, difficult, and hence worthy of substantial investment in testing and instruction?

Our laboratory has employed a method, designed originally to guide us in the selection of topics for our investigations of conceptual understanding, that has also yielded benefit for curricular design and focus. This method has involved three major parts, as follows.

Consultation with people actively involved with the topics. In our case, interviews were conducted with medical school teachers, medical students, and medical practitioners from within our medical school and the local community to gain a first set of ideas about biomedical concepts that are important for practicing medicine and that are also chronically difficult for students to learn, understand, and apply. The key idea was to gain guidance for topics to be studied in laboratory investigation from those who routinely address these topics in their teaching or practice.

Broad follow-up surveying. Based upon these interviews, a more formal survey was developed and distributed to a sample of medical teaching faculty at all medical schools in the United States and Canada. The survey addressed the same issues that were addressed in the interviews and yielded a target list of biomedical concepts that medical school teachers (who were in most instances also medical practitioners) considered both important to practicing medicine and difficult for students to learn well (Dawson-Saunders et al., 1990).

Concept analysis to identify sources of difficulty, and associated laboratory studies. When teachers claim that an idea is difficult to master, this does not necessarily make clear why the material is difficult or what might be done to help students understand better. Furthermore, it may be that concepts teachers perceive to be relatively tractable are, in fact, rather difficult. Our research has involved an ongoing attempt to identify sources of conceptual difficulty and, in addition, to predict the circumstances under which misconceptions will be strongly held. A scheme of analysis for this purpose (called the "Conceptual Stability Scheme," which shares many of the features of concept analysis outlined earlier in this chapter in the section titled "Notes on the Nature of Conceptual Difficulty: Structure and Process Issues") has been under development and is yet another tool for providing curricular focus, by providing greater precision in identifying topics likely to require special attention. Development of this scheme has been a cyclic endeavor, in which the scheme for analyzing and predicting sources of conceptual difficulty is revised, based on the results of empirical studies of students' understanding of selected concepts, and the revised scheme, in turn, is used to guide further laboratory investigation (and so on).

Laboratory work involving medical students, medical practitioners, and using selected concepts from the list obtained from the survey has been conducted to determine why particular concepts are hard, and to determine particular impediments to understanding and the systematic misconceptions individuals acquire. This work involves detailed conceptual screening instruments associated with a target concept (and its set of highly interrelated concepts), as well as directed laboratory tasks used as more precise follow-ups to these initial conceptual probes (Feltovich et al., 1989).

Such an overall program for addressing curricular selectivity, which is sensitive to guidance from practicing professionals in the domain of their everyday work, but which is augmented by laboratory investigation, yields focus for educational effort. It also provides insights into impediments to the understanding of concepts judged to be important.

CONCLUSION

In this chapter we have proposed a set of guidelines for learning and testing in advanced knowledge acquisition, emphasizing particularly those subject matter areas where material is complex and difficult and deep understanding is valued. The methods we propose are motivated in part by research we have conducted on advanced students' understanding of complex material which has revealed consistent patterns of deficiency in understanding, and in part by failures of the general educational system that are more widely recognized. We have argued that many of the problems we perceive stem from not adequately addressing the real complexity of material in testing and instruction, or from not instructing and testing in ways commensurate with complex learning goals (e.g., promoting the ability for flexible knowledge use and transfer). Indeed, we have argued that many common practices of education that involve simplification, especially as these interact with a cognitive tendency to simplify, may go beyond not addressing complexity in material to actually undermining the development of the ability to master complicated (difficult, ill-structured, etc.) topics. Hence, the unifying theme for the methods of instruction and testing we propose is that they confront complexity 'head-on'--in complicated subject matter and in sophisticated goals for the learner.

In turn, the approaches we advocate are resource intensive and probably cannot be applied uniformly or universally across the many concepts and topics of a curriculum. Hence, we have discussed the need for setting priorities in curricula and have outlined a method for achieving this kind of selectivity.

Even with selective focus, it may be infeasible (or at least impractical) to implement within the day-to-day educational process alone the kind of program for instruction and assessment that we have outlined in this chapter (involving such things as polling to identify likely areas of curricular difficulty, extensive diagnostic evaluation to identify impediments to understanding among these topics, intensive work with students to identify systematic types of error and misconception that they acquire, and so forth). What is possible, especially in schools with a relatively restricted and coherent focus (such as schools of professional education), is the development of a system in which pertinent basic cognitive research and the ongoing educational process are tightly coupled. The research program, which can tolerate the resource investment, can carry out the background investigation for instruction and testing, creating information and materials for classroom use. The classroom, in turn, provides direction and feedback for the research endeavor. On another plane, communities of researchers and educators can share the agenda of developing such a system, with different medical education research teams, for instance, addressing the particular problems associated with students' learning and understanding of different biomedical science concepts important to the practice of medicine (e.g., Patel, Kaufman, & Magder, 1991).

ACKNOWLEDGMENTS

The research reported in this chapter was supported in part by the Office of Naval Research, Cognitive Science Division (N00014-88-K-0077), the Basic Research Office of the Army Research Institute (MDA903-86-K-0443), and the Office of Educational Research and Improvement (OEG0087-C1001). The chapter does not necessarily reflect the views of these agencies. Parts of the chapter were presented at the American College Testing Program (RS), Jan., 1988 and at the Annual Meeting of the American Education Research Association (PF), April, 1988. Joan Feltovich read several drafts of the paper and made many useful suggestions.

REFERENCES

Anderson, J.R. (1982). Acquisition of cognitive skill. *Psychological Review, 18,* 396-406.

Anderson, J.R., Farrell, R., & Sauers, R. (1984). Learning to program in LISP. *Cognitive Science, 8,* 87-129.

Anderson, J.R., & Reder, L.M. (1979). An elaborative processing explanation of depth of processing. In L.S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 385-403). Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, R.C., & Pichert, J.W. (1978). Recall of previously unrecallable information following a shift of perspective. *Journal of Verbal Learning and Verbal Behavior, 17,* 1-12.

Baddeley, A.D. (1982). Domains of recollection. *Psychological Review, 89,* 708-729.

Barrows, H.S. (1983). Problem-based, self-directed learning. *JAMA, 250,* 3077-3080.

Barrows, H.S., Williams, R.G., & Moy, R.H. (1978). A comprehensive performance-based assessment of fourth-year students' clinical skills. *Journal of Medical Education, 62,* 805-809.

Bartlett, F.C. (1932). *Remembering.* Cambridge, UK: Cambridge Univ. Press.

Bransford, J.D., & Franks, J.J. (1972). The abstraction of linguistic ideas. *Cognitive Psychology, 2,* 331-350.

Bransford, J.D., 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). Cambridge, UK: Cambridge University Press.

Bransford, J.D., Stein, B.S., Shelton, T.S., & Owings, R.A. (1982). Cognition and adaptation: The importance of learning to learn. In J. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 93-110). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A.L. (1989). Analogical learning and transfer: What develops? In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369-412). Cambridge, UK: Cambridge University Press.

Brown, J.S., & Burton, R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 2,* 155-192.

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

Bruner, J.S. (1963). *The process of education.* Cambridge, MA: Harvard University Press.

Burstein, M.H. (1985). *Learning by reasoning from multiple analogies. Unpublished doctoral dissertation.* Yale University, New Haven, CT.

Burstein, M.H., & Adelson, B. (1990). Issues for a theory of analogical learning. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 137-172). Hillsdale, NJ: Lawrence Erlbaum Associates.

Caramazza, A., McCloskey, M., & Green, B.F. (1981). Naive beliefs in sophisticated subjects: Misconceptions about trajectories of objects. *Cognition, 9,* 193-215.

Carey, S. (1985). Are children fundamentally different kinds of thinkers than adults? In S.F. Chipman, J.W. Segal, & R.Glaser (Eds.), *Thinking and learning skills: Research and open questions* (Vol. 2; pp. 485-517). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M.T.H., Bassok, M., Lewis, R., Reimann, P., & Glaser, R. (1989). self-explanation. How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145-182.

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.

Chiesi, H.L., Spilich, G.J., & Voss, J.F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18,* 257-273.

Clancey, W.J. (1989). Acquiring, representing, and evaluating a competence model of diagnostic strategy. In M.T.H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise* (pp. 343-418). Hillsdale, NJ: Lawrence Erlbaum Associates.

Collins, A. M., & Gentner, D. (1983). Multiple models of evaporation processes. In *Proceedings of The Fifth Annual Conference of the Cognitive Science Society.* Rochester, NY.

Coulson, R.L., Feltovich, P.J., & Spiro, R.J. (1989). Foundations of a misunderstanding of the ultrastructural basis of myocardial failure: A reciprocation network of oversimplifications. *The Journal of Medicine and Philosophy, 14,* 109-146.

Dawson-Saunders, B., Feltovich, P.J., Coulson, R.L., & Steward, D. (1990). A survey of medical school teachers to identify basic biomedical concepts medical students should understand. *Academic Medicine, 7,* 448-454.

Dember, W.N. (1991). Cognition, motivation, and emotion: Ideology revisited. In R.R. Hoffman & D. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 153-162). Hillsdale NJ: Lawrence Erlbaum Associates.

diSessa, A. A. (1983). Phenomenology and the evolution of intuition. In D. Gentner & A.L. Stevens (Eds.), *Mental models* (pp. 15-34). Hillsdale, NJ: Lawrence Erlbaum Associates

Farr, M.J. (1987). *The long-term retention of knowledge and skills.* New York: Springer-Verlag.

Feltovich, P.J., Coulson, R.L., Spiro, R.J., & Dawson-Saunders, B.K. (in press). Knowledge application and transfer for complex tasks in ill-structured domains: Implications for instruction and testing in biomedicine. In D. Evans & V. Patel (Eds.), *Advanced models of cognition for medical training and practice.* New York: Springer-Verlag.

Feltovich, P.J., Spiro, R.J., & Coulson, R.L. (1989). The nature of conceptual understanding in biomedicine: The deep structure of complex ideas and the development of misconceptions. In D. Evans & V. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling* (pp. 111-172). Cambridge, MA: MIT (Bradford) Press.

Feltovich, P.J., Johnson, P.E., Moller, J.H., & Swanson, D.B. (1984). LCS: The role and development of medical knowledge in diagnostic expertise. In W. J. Clancey & E. H. Shortliffe (Eds.), *Readings in medical artificial intelligence: The first decade* (pp. 275-319). Reading, MA: Addison Wesley.

Flemming, M. & Chambers, B. (1983). Teacher-made tests: Windows on the classroom. In W.F. Hathaway (Ed.), *New directions for testing and measurement* (pp. 29-38). San Francisco, CA: Jossey-Bass.

Forbus, K.D. (1985). Qualitative process theory. In D.G. Bobrow (Ed.), *Qualitative reasoning about physical systems.* Cambridge, MA: MIT Press.

Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher, 18*, 27-32.

Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist, 39*, 193-202.

Gates, A.I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, 40, 104.

Gelman, R., & Greeno, J.G. (1989). On the nature of competence: Principles for understanding in a domain. In L.B. Resnick (Ed.), *Knowing, learning, and Instruction: Essays in honor of Robert Glaser* (pp. 125-186). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gentner, D. (1988). Metaphor as structure-mapping: The relational shift. *Child Development, 59*, 47-59.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). Cambridge, UK: Cambridge University Press.

Gentner, D., & Gentner D.R. (1983). Flowing waters and teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 99-130). Hillsdale, NJ: Lawrence Erlbaum Associates.

Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist, 39*, 93-104.

[GPEP] Panel on the General Professional Education of the Physician. (1984). *Physicians for the twenty-first century: The GPEP report.* Washington, DC: American Association of

Medical Colleges.]

Greeno, J.G. (1977). Process of understanding in problem solving. In N.J. Castellan, D.B. Pisoni, & G.R. Potts (Eds.), *Cognitive theory* (Vol. 2; pp. 43-83). Hillsdale, NJ: Lawrence Erlbaum Associates.

Halasz, F., & Moran, T.P. (1982). Analogy considered harmful. In *Proceedings of the Human Factors in Computer Systems Conference.* Gathersburg, MD.

Hayes, J.R. (1985). Three problems in teaching general skills. In S.F. Chipman, J.W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions* (Vol. 2; pp. 391-405). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kahneman, D., Slovic, P., & Tversky, A. (Eds., 1982). *Judgment under uncertainty.* New York: Cambridge University Press.

Lave, J. (1988). *Cognition in practice.* Boston, MA:Cambridge.

Levine, H.G., & Forman, P.M. (1973). A study of retention of knowledge of neurosciences information. *Jnl. Medical Education, 48*, 867-869.

Markman, E.M. (1981). Comprehension monitoring. In W.P. Dickson (Ed.), *Childrens' oral communication skills* (pp. 61-84). New York: Academic Press.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A.L. Stevens (Eds.), *Mental models* (299-324). Hillsdale, NJ: Lawrence Erlbaum Associates.

Morgenstern, C.F. & Renner, J.W. (1984). Measuring thinking with standardized tests. *Journal of Research in Science Teaching, 21*, 639-648).

Myers, A.C., Feltovich, P.J., Coulson, R.L., Adami, J.F., & Spiro, R.J. (1990). Reductive biases in the reasoning of medical students: An investigation in the domain of acid-base balance. In B. Bender, R.J. Hiemstra, A.J.J.A. Scherbier, & R.P. Zwierstra (Eds.), *Teaching and assessing clinical competence* (pp. 155-160). Groningen, the Netherlands: BoekWerk Publications.

National Commission on Excellence in Education. (1983). *A nation at risk.* Washington, DC: U.S. Government Printing Office.

National Science Foundation. (1982). *Science and engineering education: Data and information.* Washington, DC: National Science Foundation.

Nickerson, R.S. (1989). New directions in educational assessment. *Educational Researcher, 18*, 3-7.

Norman,G.R., Muzzin, L.J., Williams, R.G., & Swanson, D.B. (1985). Simulations in health sciences education. *Journal of Instructional Development, 8*, 11-17.

Patel, V.L., Kaufman, D.R. & Magder, S. (1991). Causal explanation of complex physiological concepts by medical students. *International Journal of Science Education, 13(2)*, 171-185.

Pepper, S. (1942). *World hypotheses.* Berkeley, CA: University of California Press.

Perkins, D.N., & Simmons, R. (1989). Patterns of misunderstanding: An integrative model for science, math, and programming. *Review of Educational Research, 58*, 303-326.

Porter, A. (1989). A curriculum out of balance: The case of elementary school mathematics. *Educational Researcher, 18*, 9-15.

Ross, B.H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 629-639.

Ross, B.H. (1989). Remindings in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 438-496). Cambridge, UK: Cambridge Univ. Press.

Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In J.L. McClelland, D.E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing* (Vol. 2; pp. 194-281). Cambridge, MA: MIT (Bradford) Press.

Spiro, R.J. (1980). Constructive processes in prose comprehension and recall. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 245-278). Hillsdale, NJ: Lawrence Erlbaum Associates.

Spiro, R.J., Coulson, R.L., Feltovich, P.J., & Anderson, D.K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In *The tenth annual conference of the cognitive science society* (pp. 375-383). Hillsdale, NJ: Lawrence Erlbaum Associates.

Spiro, R.J, Feltovich, P.J., Coulson, R.L., & Anderson, D.K. (1989). Multiple analogies for complex concepts: Antidotes for analogy-induced misconception in advanced knowledge acquisition. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 498-531). Cambridge, MA: Cambridge University Press.

Spiro, R.J., Feltovich, P.J., Jacobson, M., & Coulson, R.L. (1991). Cognitive flexibility, constructivism, and hypertext: Advanced knowledge acquisition in ill-structured domains. *Educational Technology, 31(9)*, 22-25.

Spiro, R.J., Vispoel, W., Schmitz, J., Samarapungavan, A., & Boerger, A. (1987). Knowledge acquisition for application: Cognitive flexibility and transfer in complex content domains. In B.C. Britton (Ed.), *Executive control processes* (pp. 177-200). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tulving, E, & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373.

Wason, P.C., & Johnson-Laird, P.N. (1972). *The psychology of deductive reasoning: Structure and content.* Cambridge MA: Harvard Univ. Press.

White, B.Y. (1984). Designing computer activities to help physics students understand Newton's laws of motion. *Cognition and Instruction, 1*, 69-108.

White, B.Y. (in press). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction.*

Whitehead, A.N. (1929). *The aims of education.* New York: Macmillan.

Wimsatt, W.C. (1980). Reductionistic research strategies and their biases in the units of selection controversy. In T. Nickles (Ed.), *Scientific discovery: Case studies* (pp. 213-259). Dordrecht, Holland: D. Reidel Publishing.

Wittgenstein, L. (1953). *Philosophical investigations.* New York: Macmillan.

Dr. Phillip L. Ackerman
University of Minnesota
Department of Psychology
75 East River Road
N218 Elliott Hall
Minneapolis, MN 55455

Dr. Beth Adelson
Department of Psychology
Rutgers University
Camden, NJ 08102

Dr. James A. Anderson
Department of Cognitive
and Linguistic Sciences
Brown University
Box 1978
Providence, RI 02912

Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Thomas G. Bever
Department of Psychology
University of Rochester
River Station
Rochester, NY 14627

Dr. John Black
Teachers College, Box 8
Columbia University
525 West 120th Street
New York, NY 10027

Dr. Ann Brown
Graduate School of Education
University of California
EMST-4533 Tolman Hall
Berkeley, CA 94720

Dr. John T. Bruer
James S. McDonnell Foundation
Suite 1610
1034 South Brentwood Blvd.
St. Louis, MO 63117

Dr. Michelene Chi
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Susan Chipman
Cognitive Science Program
Office of Naval Research
800 North Quincy St.
Arlington, VA 22217-5000

Dr. David E. Clement
Department of Psychology
University of South Carolina
Columbia, SC 29208

Chief of Naval Education
and Training (N-5)
NAS
Pensacola, FL 32508

Director
Life Sciences, Code 1142
Office of Naval Research
Arlington, VA 22217-5000

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Geory Delacote
Exploratorium
3601 Lyon Street
San Francisco, CA 94123

Defense Technical Information Center
Cameron Station, Bldg. 5
Alexandria, VA 22314
(2 copies)

Dr. Martha Evens
Department of Computer Science
Illinois Institute of Technology
10 West 31st Street
Chicago, IL 60616

Dr. P.A. Federico
Code 51
NPRDC
San Diego, CA 92152-6800

Mr. Wallace Feurzeig
Educational Technology
Bolt Beranek & Newman
10 Moulton St.
Cambridge, MA 02238

Dr. Kenneth D. Forbus
Institute for the Learning Sciences
Northwestern University
1890 Maple Avenue
Evanston, IL 60201

Dr. Carl H. Frederiksen
Dept. of Educational Psychology
McGill University
3700 McTavish Street
Montreal, Quebec
CANADA H3A 1Y2

Dr. Norman Frederiksen
Educational Testing Service
(05-R)
Princeton, NJ 08541

Dr. Dedre Gentner
Northwestern University
Department of Psychology
2029 Sheridan Road
Swift Hall, Rm 102
Evanston, IL 60208-2710

Dr. Herbert Ginsburg
Box 184
Teachers College
Columbia University
525 West 121st Street
New York, NJ 10027

Dr. Drew Gitomer
Educational Testing Service
Princeton, NJ 08541

Dr. Robert Glaser
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Susan R. Goldman
Peabody College, Box 45
Vanderbilt University
Nashville, TN 37203

Dr. Sherrie Gott
AFHRL/MOMJ
Brooks AFB, TX 78235-5601

Dr. T. Govindaraj
Georgia Institute of Technology
School of Industrial and
Systems Engineering
Atlanta, GA 30332-0205

Dr. Arthur C. Graesser
Department of Psychology
Memphis State University
Memphis, TN 38152

Dr. James G. Greeno
School of Education
Stanford University
Room 311
Stanford, CA 94305

Dr. Kristian J. Hammond
Department of Computer Science
University of Chicago
1100 E. 58th St.
Chicago, IL 60637

Dr. James Hiebert
Department of Educational
Development
University of Delaware
Newark, DE 19716

Dr. Keith Holyoak
Department of Psychology
University of California
Los Angeles, Ca 90024

Dr. Claude Janvier
CIRADE, X-7120
UQAM
P.O. Box 8888, Succ. A
Montreal, Quebec
CANADA, H3C 3P8

Dr. Ruth Kanfer
University of Minnesota
Department of Psychology
Elliott Hall
75 E. River Road
Minneapolis, MN 55455

Dr. Michael Kaplan
Office of Basic Research
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Janet L. Kolodner
Georgia Institute of Technology
College of Computing
Atlanta, GA 30332-0280

Dr. Jill Larkin
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

Dr. Marcia C. Linn
Graduate School of Education
EMST
Tolman Hall
University of California
Berkeley, CA 94720

Dr. Sandra P. Marshall
Department of Psychology
San Diego State University
San Diego, CA 92182

Dr. Richard E. Mayer
Department of Psychology
University of California
Santa Barbara, CA 93106

Dr. Allen Munro
Behavioral Technology Laboratories
USC
250 N. Harbor Dr., Suite 309
Redondo Beach, CA 90277

Deputy Director Manpower,
Personnel & Training Div.
Naval Sea Systems Command
ATTN: Code 04MP 511
Washington, DC 20362

Dr. Donald A. Norman
Department of Cognitive Sciences
University of California
La Jolla, CA 92093-0515

Director
Training Systems Department
NPRDC (Code 14)
San Diego, CA 92152-6800

Director
Training Technology Department
NPRDC (Code 15)
San Diego, CA 92152-6800

Dr. Stellan Ohlsson
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Judith Orasanu
Mail Stop 239-1
NASA Ames Research Center
Moffett Field, CA 94035

Dr. Roy Pea
Institute for the Learning Sciences
Northwestern University
1890 maple Avenue
Evanston, IL 60201

Dr. Peter Pirolli
School of Education
University of California
Berkeley, CA 94720

Dr. Joseph Psotka
ATTN: PERI-IC
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Brian Reiser
Cognitive Science Lab
221 Nassau Street
Princeton University
Princeton, NJ 08542

Dr. Lauren Resnick
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Dr. Alan H. Schoenfeld
University of California
Department of Education
Berkeley, CA 94720

Dr. Judith W. Segal
OERI
555 New Jersey Ave., NW
Washington, DC 20208-5648

Mr. Richard J. Shavelson
Graduate School of Education
University of California
Santa Barbara, CA 93106

Dr. Edward Silver
LRDC
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Jerome E. Singer
Department of Medical Psychology
Uniformed Services
University of the Health Sciences
4301 Jones Bridge Road
Bethesda, MD 20814-4799

Dr. Derek Sleeman
Computing Science Department
The University
Aberdeen AB9 2FX
Scotland
UNITED KINGDOM

Dr. Richard E. Snow
School of Education
Stanford University
Stanford, CA 94305

Dr. James J. Staszewski
Department of Psychology
University of South Carolina
Columbia, SC 29210

Dr. Douglas Towne
Behavioral Technology Labs
University of Southern California
1228 Spring Street
St. Helena, CA 94574

Dr. Kurt VanLehn
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. David Wiley
School of Education and Social Policy
Northwestern Univesity
Evanston, IL 60208

Dr. David C. Wilkins
University of Illinois
Department of Computer Science
405 North Mathews Avenue
Urbana, IL 61801

Dr. Merlin C. Wittrock
Graduate School of Education
University of California-Los Angeles
Los Angeles, CA 90024

Dr. Joseph L. Young
National Science Foundation
Room 320
1800 G Street, NW
Washington, DC 20550

Dr. Harriet Zucheruea
Vice President
Andrew W. Mellon Foundation
140 E. 62nd St.
New York, NY 10021