Technical Report No. 409

KNOWLEDGE ACQUISITION FOR APPLICATION: COGNITIVE FLEXIBILITY AND TRANSFER IN COMPLEX CONTENT DOMAINS

Rand J. Spiro, W. P. Vispoel, John G. Schmitz,
Ala Samarapungavan, and A. E. Boerger

University of Illinois at Urbana-Champaign

September 1987

Center for the Study of Reading

TECHNICAL REPORTS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
174 Children's Research Center
51 Gerty Drive
Champaign, Illinois 61820

CENTER FOR THE STUDY OF READING

UR.

Technical Report No. 409

KNOWLEDGE ACQUISITION FOR APPLICATION:
COGNITIVE FLEXIBILITY AND TRANSFER
IN COMPLEX CONTENT DOMAINS

Rand J. Spiro, W. P. Vispoel, John G. Schmitz,
Ala Samarapungavan, and A. E. Boerger

University of Illinois at Urbana-Champaign

September 1987

University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

This research was supported in part by Contract No. 400-81-0030 from the National Institute of Education and by Contract No. MDA903-86-K-0443 from the Army Research Institute for the Behavioral and Social Sciences. The senior author would like to express his gratitude to Dr. Paul Feltovich for many helpful discussions related to the topic of this chapter. Correspondence concerning this article should be addressed to Rand J. Spiro, Center for the Study of Reading, 51 Gerty Dr., Champaign, IL 61820. This paper also appears in B. K. Britton and S. M. Glynn (Eds.) (1987), Executive Control Processes. Hillsdale, NJ: Erlbaum.

Abstract

A theoretical orientation to learning and instruction in ill-structured knowledge domains is presented. The theory is especially concerned with the application of knowledge in new situations (knowledge transfer), rather than the mere reproduction of knowledge in the way that it was originally learned. It is argued that knowledge transfer in complex and ill-structured domains is centrally dependent upon "cognitive flexibility." According to the theory, greater flexibility in the representation of domain knowledge will result from approaches that promote highly interconnected rather than neatly compartmentalized or hierarchicalized mental representations; that represent knowledge in terms of multiple, rather than single, prototypes and analogies; that increase the emphasis on learning from cases, while qualifying and restricting the scope of application of abstract principles; and that rely upon situation-dependent schema assembly rather than the retrieval of a rigid, prepackaged schema. A nonlinear system of learning and instruction that promotes these requisite features of cognitive flexibility is presented. In the system, cases or examples in a conceptual "landscape" are criss-crossed in a variety of directions, along multiple dimensions. Empirical paradigms for testing the theory are presented, and positive results of preliminary experiments are discussed. Finally, issues of case selection and sequencing are addressed, and the possible role of visual-perceptual adjunct representations in making the management of complexity more tractable is highlighted.

KNOWLEDGE ACQUISITION FOR APPLICATION: COGNITIVE FLEXIBILITY AND TRANSFER IN COMPLEX CONTENT DOMAINS

A fundamental tenet of all recent theories of comprehension, problem solving, and decision making is that success in such cognitive arenas depends on the activation and appropriate application of relevant preexisting knowledge. Despite the substantial agreement on this general claim, we know very little about the organization of background knowledge and the method of its application to the understanding of new situations when, because of a combination of the breadth, complexity, and irregularity of a content domain, formulating knowledge in that domain to explicitly prescribe its full range of uses is impossible. We call knowledge domains of this type *ill-structured* and contrast them with more routinizable knowledge domains that we refer to as *well-structured*.[1] What does one do when relevant prior knowledge is not already organized to fit a situation (as will frequently be true in ill-structured domains, by definition) and so must be assembled from different knowledge sources in memory? This is a problem of knowledge *transfer*. We address a crucial issue in transfer: How should knowledge be acquired and organized to facilitate a *wide range* of future applications?

The principle contentions developed in this chapter are:

1. that theories in the cognitive sciences have produced a far better understanding of cognitive process in well-structured domains (WSDs) than in ill-structured domains (ISDs);

2. that theories appropriate for WSDs are in many ways inappropriate for ISDs--that, in fact, optimal conditions of learning and instruction in the two kinds of domains are *opposite* in several important respects;

3. that one of the most serious problems with treating ISDs as if they were WSDs is an inability to establish a basis for knowledge transfer, for the application of preexisting knowledge to new situations;

4. that transfer in ISDs is best promoted by knowledge representations that possess the following features: multiple interconnectedness between different aspects of domain knowledge, multidimensional or multiperspectival representation of examples/cases, and allowance for various forms of naturally occurring complexity and irregularity;

5. that in order for knowledge structures to possess the characteristics described in 4, emphasis must shift from the *retrieval* of a precompiled schema to the *assembly* of a situation-sensitive schema from knowledge fragments--the features described in 4 and 5 are characteristics of what we call *cognitive flexibility*;

6. that the best way to learn and instruct in order to attain the goal of cognitive flexibility in knowledge representation for future application is by a method of case-based presentations which treats a content domain as a landscape that is explored by "criss-crossing" it in many directions, by reexamining each case "site" in the varying contexts of different neighboring cases, and by using a variety of abstract dimensions for comparing cases.

Our primary claim is that in situations where complete comprehension or solution paths are neither inherent in the task or its description (i.e., are not implicitly or explicitly prescribed), nor anticipatable in the natural course of earlier learning and concomitant knowledge representation (either by explicit mention or by generative rule), then the key factors affecting the success with which prior knowledge is used to improve performance in a new situation will be the *flexibility* with which the relevant prior knowledge is represented in memory, and the mastery or *control* the individual has over those flexible representations (the ability to recombine elements of the representations, reorder the importance of elements in different contexts, and so on).

When knowledge cannot be routinized, mechanized, or automatized, it must be flexibly controlled. And control is not solely a process issue, independent of content. Executive control *strategies* require flexible knowledge *structures* to operate upon. In turn, control over flexible representations will give an individual more control over a content domain; rather than monolithic prepackaged knowledge structures constraining an individual to apply knowledge in a fixed and limited manner, the individual controls the knowledge; that is, a great variety of nonpredetermined ways to adapt knowledge to the task and content elements involved in the new situation are available. This paper presents a theory of learning and instruction, of knowledge representation and application, for the flexibility-based control that enables transfer. The goal of the program of research that we discuss is the validation of a set of basic principles and related instructional practices that will allow students to better apply the knowledge they acquire from formal schooling to new, real-world cases--knowledge that is built for *use*, not for imitative reproduction in artificial school or laboratory settings.

Schema-Theoretic Knowledge Representation and the Problem of Transfer

It is hard to imagine a more valued intellectual ability than that of *independent thought*--the ability to "think for yourself" when applying the necessarily limited knowledge acquired in formal schooling and training to the wide variety of situations in which that knowledge is relevant--the ability to use one's knowledge flexibly and to efficiently adapt it to varying contexts. Of what value is knowledge if its potential for application is much more limited than the range of uses to which it needs to be put? Very little, most would agree.

Nevertheless, a class of very influential contemporary approaches to knowledge representation, schema, frame, and script theories (e.g., Rumelhart & Ortony, 1977; Schank & Abelson, 1977) have had four interrelated shortcomings related to the problem of transfer. (As we see later, many of the following criticisms also apply to another class of approaches to knowledge representation: prototype and exemplar theories; e.g., Rosch & Mervis, 1975).

First, because these kinds of knowledge structures are frequently *prepackaged* (precompiled), they tend to be overly *rigid*. As a result, they provide very little opportunity for adaptation to diverse contexts of use.

Second, these modes of representation tend to isolate or *compartmentalize* aspects of knowledge that, in use, need to be interconnected. Again, the result is a limited potential for transfer.

Third, they have frequently treated complex subject matter as if it were simpler than it really was; complexities that transfer depends on have been *artificially neatened.*

Fourth, they have often implicitly assumed that knowledge domains possess more *regularity* or consistency across cases of application than they actually do.

These four characteristics of the recently predominant modes for representing knowledge, antithetical to the flexibility needed for wide application, are essentially simplifying assumptions: It is assumed that knowledge domains are *simpler* and more *regular* than they in fact are; these assumptions lead to representational approaches that are more rigid and compartmentalized than they ought to be. Historically, the assumptions of simplicity and regularity are strategically understandable and justifiable: Progress in new fields comes much more slowly without simplifying assumptions. And there has been considerable progress in cognitive science. However, when progress has been as slow as it has been on so important a topic as transfer, then it is time to drop those simplifying assumptions that are causing progress to be impeded.[2]

The simplifying assumptions of cognitive science apply as well to dominant modes of education. Simplification of complex subject matter makes it easier for teachers to teach, for students to take notes and prepare for their tests, for test-givers to construct and grade tests, and for authors to write texts. The result is a massive "*conspiracy of convenience.*" To take one example of the consequences of this tendency, Spiro, Feltovich, and Coulson (in preparation) have identified more than a dozen serious errors in the concepts held by a majority of medical students tested. Each of the errors was connected to a different kind of cognitive or educational oversimplification.

The overall effect of the simplifying features of knowledge representation systems and instructional strategies mentioned previously is a leveling tendency, a tendency towards *monolithic* approaches. Understanding is seen as proceeding in essentially the same way across instances of the same topic. Our view is different: The conditions for applying old knowledge are subject to considerable *variability*, and that variability in turn requires *flexibility* of response. Monolithic representations of knowledge will too often leave their holders facing situations for which their rigid "plaster-casts" simply do not fit. The result is the often heard complaint of students: "We weren't taught that." By this, they mean that they weren't taught *exactly* that. They lack the ability to use their knowledge in new ways, the ability to think for themselves. Our research focuses on fostering the kinds of flexible knowledge representations that would free students and trainees from the limitations of having to use the information they receive in instruction in only that way in which it was originally instructed (rigid knowledge structures that need to be applied rigidly).

To achieve this goal we systematically depart from the four interrelated simplifying assumptions of the schema-type theories that we identified earlier (which, again, are also common assumptions in schooling).

- Rigid, monolithic, prepackaged knowledge representations are replaced by flexible representations in which fragments of knowledge are moved about and *assembled* to fit the needs of a given *context* of application. Instead of prepackaged schemata, purpose-sensitive situational schemata are constructed, thus allowing knowledge to be used in different ways on different occasions for different purposes. The emphasis is shifted from prepackaged schemata to the ingredients for many potential schemata; schema selection is devalued in favor of schema assembly; storage of fixed knowledge is devalued in favor of the *mobilization of potential knowledge*. One cannot have a prepackaged schema for everything.

- We replace highly compartmentalized knowledge representations with structures characterized by a high degree of *interconnectedness*. Appropriate compartmentalization of knowledge in one situation may not be appropriate in another. Multiple interconnectedness permits (a) situation-specific categorization, (b) multiple access routes to relevant case precedents in long-term memory from the details of new cases, and (c) the development of a reservoir of potential analogies when case precedents are less literally relevant.

- Instead of inappropriate simplification, we work with complex subject matter (e.g., historical topics, military strategy, biomedical concepts), acknowledging and teaching towards the *complexity* inherent in them. (Later in this chapter we address the issue of how learners can be aided in managing difficult complexities.) And just as subject matter is frequently more complex than is realized, so, too, are the real cases to which knowledge about that subject matter has to be applied. Cases or examples must be studied as they really occur, in their natural contexts, not as stripped down "textbook examples" that conveniently illustrate some principle. The application of knowledge in new situations will be thwarted if the contingencies for application are more complex than the knowledge brought to bear.

- We relax the often unrealistic assumption of regularity, of routinizability, of well-structuredness--the assumption that subsets of individual cases (applications, examples) in a knowledge domain are sufficiently alike that they may be covered in common by a self-consistent system of general principles or rules. Thus, where schema-type theories tend to be overly general and to abstract away from individual cases that are classified in the same way, our approach sticks closer to the specific details and characteristics of individual cases. The consequences of treating ill-structured material as if it were well-structured are knowledge representations that are inapplicable in transfer situations or that produce inaccuracy. (Note that although complexity is highly related to irregularity, they are not the same thing. For example, the facade of the Old Executive Office Building next to the White House has a highly complex but regular structure. We work with domains that are complex and ill-structured.)

A Note on Related Research

Our approach has many connections to prominent thematic strands of current theoretical interest. However, the manner in which we instantiate these strands, especially the way they are configured in our overall approach, is novel. Therefore, because the main thing in common between what we do and the research of others involves overlap in subsets of the constituent themes emphasized rather than in the specifics of how those themes are applied and combined, we simply point to those thematic areas of overlap here: analogy (Carbonell, 1983; Gentner & Gentner, 1983; Gick & Holyoak, 1980, 1983; Rumelhart & Norman, 1981; Winston, 1983); learning from examples (Anderson, Kline, & Beasley, 1979; Carnine, 1980; Medin & Schaffer, 1978); reminding (Kolodner, 1980; Ross, 1984; Schank, 1982); connectionism, parallel distributed systems, "society of mind" (McClelland & Rumelhart, 1985; Minsky, 1979; Waltz & Pollack, 1985); tutorial guiding (Collins & Stevens, 1983); constructivity (J. Anderson, 1983; R. Anderson, 1977; Barsalou, in press; Bartlett, 1932; Spiro, 1977, 1980a, 1983); cognitive complexity (Scott, Osgood, & Peterson, 1979; Streufert & Streufert, 1978); automatic versus controlled processing (Shiffrin & Schneider, 1977; Schneider & Shiffrin, 1977); efficiency and cognitive economies in mental processes (Spiro, 1980b; Spiro & Esposito, 1981; Spiro, Esposito, & Vondruska, 1978); metacognition and learning to learn (Brown, Campione, & Day, 1981); perceptual and experiential aspects of conceptual memory (Bartlett, 1932; Spiro, 1983; Spiro, Crismore, & Turner, 1982); contextual encoding variability (Smith, Glenberg, & Bjork, 1978); linear separability (Medin & Schwanenfluegel, 1981); fluid task environments (Lenat, Hayes-Roth, & Klahr, 1983); structural analysis of content-area text (Britton & Black, 1985; Spiro & Taylor, in press).

A good example of a researcher whose work incorporates many of the same themes that we do is Barsalou. As we are, he is centrally concerned with contextual variability in conceptual structure, cognitive flexibility, and temporary concept construction under contextual constraint (Barsalou, in press). Even here though, the similarities are not as instructive as they might at first appear because of the difference in the kind of domains studied. Barsalou works with relatively low-level concepts and categories. There are many fundamental differences in both internal structure and the cognitive processes that must operate on that structure between concepts like "Birds" and "Fruit" (or even "Places to Go On A Vacation"), on the one hand, and complex *topics* such as "20th-Century History" or "Military Battles," on the other hand. There are also big differences between semantic memory tasks and the application of knowledge to real cases. All the problems with concepts and categories addressed by researchers like Barsalou (e.g., graded structure; contextual variability; see also Medin & Smith, 1984; Smith & Medin, 1981) are so greatly exacerbated by dealing with knowledge application in real-world content domains that are made up of *many concepts*, that new issues of learning, representation, and transfer are inevitably introduced. The theoretical and empirical solutions corresponding to these issues produce a picture that overlaps very little between simple concepts and complex topical domains.

So, as promised earlier, because of the novel way we apply the preceding headlined themes, we move on to further discussion of the nature of our own approach. However, before proceeding we should explicitly acknowledge the most pervasive influence on our theoretical orientation and empirical procedures, the later work of Wittgenstein (1953). Our treatment of complex topical knowledge was inspired by prominent Wittgensteinian metaphors for knowledge organization and learning, especially the metaphor of the "criss-crossed landscape." These and other debts to Wittgenstein (e.g., the reliance on approximate processes of family resemblance and the role of visual-perceptual forms of representation in attaining *synoptic* conceptual understandings) will be obvious in the following section and throughout the paper.

Ill-structuredness, Learning from Cases, and the Instructional Metaphor of the "Criss-Crossed Landscape": A Prescription for Transfer

Ill-structuredness

In many domains, if one compares the features of large numbers of cases, a subset of the following conclusions may be drawn:

- There are no rules or principles of sufficient generality to cover most of the cases, nor defining characteristics for determining the actions appropriate for a given case.
- Hierarchical relations of dominance and subsumption are inverted from case to case.
- Prototypes tend to often be misleading.
- The same features assume different patterns of significance when placed in different contexts.
- An explosion of higher order interactions among many relevant features introduces aspects of case novelty.

It is such failures of general principles as these that we designate by the family label *ill-structuredness*. As will be seen later, this failure of general principles in ill-structured domains is directly related to the most sweeping recommendation that will emerge from our theoretical orientation: an emphasis in learning and instruction upon *multiplicity*. Instead of using a single knowledge structure, prototype, analogy, and so on, multiple knowledge precedents will need to be applied to new situations (multiple schemas, several past cases, overlapping analogies). Under conditions of ill-structured complexity, single approaches provide insufficient coverage.

"Ill-structured Complexity in the Domain 20th-Century History." How would the events and phenomena of the 20th century be classified and presented in typical instruction? Some abstract system for organizing material would be developed. One likely approach would be to identify several themes of the 20th century to serve as organizing compartments (chapters, subheads), and then individual examples would be slotted where they seemed to best fit. "Moral Relativism" and "Knowledge Specialization" might both be discussed in a chapter on "Fragmentation as a Theme of 20th-Century Life"; a chapter about "Irrationality and Uncertainty in 20th-Century Ideas" might refer to "Freud's Psychological Theory of Unconscious Control of Motivation and Behavior" and "The Uncertainty Principle in Physics"; there might be a chapter on "Alienation" that would cite "Massive Bureaucracies Dwarfing the Individual" and "Mass Media Brainwashing"; and so on. Notice the implicit assumption of well-structuredness: Cases are taken as (often interchangeable) instances or illustrations of abstract themes. "The Advent of Nuclear Weapons" and "Existentialist Philosophy" both illustrate irrational aspects of the era, and it is the latter, more abstract, point that is taken as the important lesson.

However, think some more about the samples. Is not existentialist philosophy an example of alienation as well as a demonstration of the recent trend away from Cartesian views of rational man? Are there not advantages to grouping Freud's theories with the subliminal influences of mass media to bring out a point about how influences we are not aware of have come to increasingly control 20th-century life? In fact, any single organizational scheme for presenting examples from 20th-century history will suffer from two important shortcomings: (a) It will not be possible to present together all the cases that it would be *instructive* to present together, and cases that are not in the same physical vicinity in text will tend not to be closely connected in memory; and (b) much of the multifacetedness of the individual case will be lost as its significance is *narrowed* to the abstract point that the case is presented as an illustration of. Similar problems would arise if some other system were used to represent the 20th-century cases. For example, if events were presented in chronological order, connections between temporally distant events would be likely to be missed. Similar consequences would ensue from a division along lines like politics, economics, and culture. *Any* single system for organizing material would not establish enough connections along enough dimensions to prepare one to deal with the great varieties of discourse about the 20th century that a learner is likely to encounter.

The Importance of Cases

It is partly because of such ill-structuredness that classroom instruction in professional fields (e.g., medicine, business) is so often augmented by considerable *case experience* (and that instruction in basic, nonapplied fields of knowledge ought to be). Lectures and textbooks tend to stress generalizations, commonalities, and abstractions over cases. Such approaches are clearly very effective when cases tend to be pretty much alike in how they have to be analyzed and responded to (i.e., in well-structured domains). Unfortunately, it is too often true that such assumptions are unwarranted, and the greater convenience associated with the traditional pedagogies is negated if they are ineffective.

How is successful performance possible in the absence of generalizability across cases? The answer is by focusing analysis and knowledge representation more at the level of the individual case (example, occasion of use, event, or other cognate terms) and by guided experience with large numbers of cases. If cases come in many forms, one needs to see many cases in order to represent their varieties of contextual influences and configurations of features. In a well-structured domain, cases (examples, etc.) are luxuries, helpful in illustrating general principles and then discardable; the principles can stand for all their subsumed, interchangeable cases. In ill-structured domains, crucial information tends to be uniquely contained in individual cases--examples are not just nice, they are necessary. Our approach to knowledge acquisition is highly case-based.

The problem is compounded, however, by the fact that ill-structuredness has a limiting effect on case-based training as well as on the more abstract approaches: Just as there are no homogeneous systems of principles or rules that can be generally applied, there are *no generally applicable prototype cases*. Ill-structuredness means that there cannot be *any* recourse to homogeneity, to any single course of action across instances, whether it involves a single guiding principle, a single organizational scheme, or a single prototype case. (Of course, in relatively well-structured domains like "Trips to Fancy Restaurants," a single conceptual macrosystem can be very effective.) Real-world cases tend to possess a multifaceted complexity and thus need to be represented in lots of different ways in order to bring out those multiple facets. Then, instead of a single case being the basis for case-based cognitive processing, *aspects* of different cases need to be *combined*, and it is the resulting assemblages, made up of fragments of different cases, that underlie an important part of case-based reasoning. The *re*construction of knowledge requires that it first be *de*constructed--flexibility in applying knowledge depends on both schemata (theories) and cases first being disassembled so that they may later be adaptively reassembled. (Of course, some integral, nondecomposed case information must also be retained for guidance in ecologically realisitic case assembly.)

Thus, two kinds of flexibility are needed in knowledge structures for ill-structured domains, each of which is a central feature of the mental representations we attempt to build:

1. Each complex real-world case needs to be decomposed and represented along many partially overlapping dimensions (i.e., the same information must be represented in lots of different ways).

2. Many connections must be drawn across the decomposed aspectual fragments of the cases in 1, thus establishing many possible routes for future assembly and creating many potential analogies useful for understanding new cases or making new applications; it is for this reason that our instructional system emphasizes connections between apparently dissimilar cases (and aspects of cases)--connections among obviously similar cases are much more likely to be noticed without special training, and thus not cause transfer problems.

Nonlinear and Multidimensional Arrangement of Cases

An example of how the two preceding theoretical commitments imply instructional methodologies is the use in our program of *rearrangements of case presentation sequences* so that the same case occurs in the context of various other cases. This enables different features of the individual multifaceted cases to be highlighted, depending on the characteristics of the other cases they are juxtaposed to, while simultaneously establishing multiple connections, including distant ones.

The discussion of case rearrangement brings us to the dominant metaphor employed in our theory of flexible knowledge representation and case-based instruction. Following the lead of Wittgenstein (see, for example, the preface to the *Philosophical Investigations*), we think of an ill-structured knowledge domain as akin to a *landscape*. Landscapes are often complex and ill-structured. No two sites are exactly alike, yet all sites possess many (but not all) of the salient features of the total landscape. The same could be said of landscape regions made up of several sites. The best way to come to understand a given landscape is to explore it from many directions, to traverse it first this way and then that (preferably with a guide to highlight significant features). Our instructional system for presenting a complexly ill-structured "topical landscape" is analogous to physical landscape exploration, with different routes of traversing study-sites (cases) that are each analyzed from a number of thematic perspectives.

The notion of "criss-crossing" from case to case in many directions, with many thematic dimensions serving as routes of traversal, is central to our theory. The treatment of an irregular and complex topic *cannot be forced in any single direction* without curtailing the potential for transfer. If the topic can be applied in many different ways, none of which follow in rule-bound manner from the others, then limiting oneself in acquisition to, say, a single point of view or a single system of classification, will produce a relatively *closed* system, instead of one that is open to context-dependent variability. By criss-crossing the complex topical landscape, the twin goals of highlighting multifacetedness and establishing multiple connections are attained. Also, awareness of variability and irregularity is heightened, alternative routes of traversal of the topic's complexities are illustrated, multiple entry routes for later information retrieval are established, and the general skill of working around that particular landscape (domain-dependent processing skill) is developed. *Information that will need to be used in a lot of different ways needs to be taught in lots of different ways.* Real cases (events, uses, etc.) have multiple slants, and, because the goal of widespread transfer must be to be ready for anything that realistically is likely to arise, learning and instruction must anticipate using many of these many slants in the ways they tend to occur.[3] Criss-crossing a topic in many directions serves this purpose. It builds flexible knowledge. Accordingly, we construct *nonlinear* and *multidimensional* acquisition texts.

Empirical Paradigms for Testing the Theory

The acquisition phase of a typical experiment involves presenting cases in an experimental condition that, in general, involves manipulations of acquisition texts that, by using the kind of landscape criss-crossing orientation we have discussed, promote awareness and representation of complexities, ill-structuredness, and multiple potential connections across cases. Our control conditions tend to simplify, to present the same case material according to a single well-structured system ("textbookization"), and to minimize case interconnectedness across the compartments of classification employed in the acquisition text.

Subjects always take at least two kinds of test, one that involves fairly literal understanding and reproductive recall of information from the acquisition text, and the other involving some sort of transfer (see following). Differences between experimental and control conditions are measured in terms of how much transfer they produce, and also in terms of the efficiency with which information stored in memory produces the potential for transfer. We call the amount of transfer enabled by equivalent amounts of old knowledge in long-term memory a measure of *transfer punch*.

There is a basic set of predictions in experiments testing the "Landscape Criss-Crossing" theory. Control conditions are expected to have mnemonic advantages that result in the rapid and accurate memorizability of the material presented in the acquisition phase of the experiments (compared to experimental conditions)--clear-cut "scaffoldings" are most effective for supporting reproductive, fact-retention-type memory (Anderson, Spiro, & Anderson, 1978; Ausubel, 1968). However, it is also expected that control conditions will produce overly inflexible, closed down representations that will result in less transfer than the experimental conditions. (This should especially be so at later stages of practice, because the highly interconnected representations produced by the experimental manipulations probably require some "critical mass" of information before the assembly processes they depend on can begin to operate.)

Results of preliminary tests of the theory. Data from two experiments using high school subjects have been analyzed. Both experiments produced results that conformed to our theoretical predictions. In the first experiment, 24 prominent and characteristic examples of 20th-century events and phenomena were selected. These cases included such developments as the advent of nuclear weaponry, the development of rapidly transmitted mass communications, and the loss of individuality caused by dealing with impersonal and massive bureaucracies. A paragraph was written about each case/example. In the control condition, each of the paragraph-long case descriptions was placed in a chapter corresponding to one of three abstract themes it *best illustrated* (e.g., "Chaos, Uncertainty, and Irrationality," "Fragmentation of Old Unities," and "New Freedoms Mix With Powerful New Controls on the Individual"). This seductive nesting of examples under an abstract point or general principle was intended to parallel the unidimensional treatment of cases and the abstraction-centered organization typical of textbooks and training manuals. In the experimental condition, after subjects read the control text, the same case-paragraphs (with minor modifications) were re-presented in a completely different context: Each case was paired with one from a *different chapter*. The intent was to bring out the multifacetedness of the individual case and to demonstrate that the deceptively simple and neat structure of the control text is not an accurate reflection of the domain's actual organization. Both groups had the same amount of study time.

The second experiment paralleled the first one, with the exception that the 20th-century cases were not re-paired in the experimental condition. Instead, the cases were presented in the same order as in the control condition but with individualized case-to-case linkages substituting for more abstract superordinate linkages of several cases. In other words, an experimental *case-centered* presentation scheme was contrasted with an *abstraction-centered* approach to the exact same material.

These early studies did not provide ideal tests of the theory. It would have been desirable to allow longer study time to process the complex materials and, especially, to permit more re-pairings of the cases, more criss-crosses of the topical landscape. Yet, even these first attempts, under far from satisfactory conditions, produced the predicted results. In the first experiment, subjects in the "textbookized" control condition outscored those in the "criss-crossed" experimental condition on a reproductive memory test for the gist of the material as it was presented. The neat, well-structured scaffolding of the control condition provided better support for encoding and retrieving details than did the more complex scheme of the experimental condition. However, when the knowledge from the acquisition text had to be applied in some new way, performance in the experimental conditions of both experiments exceeded that of controls on six out of six transfer measures employed.

The following six transfer tests were used in the two experiments (three in each):

1. and 2. Two different tests of integrative comprehension of new 20th-century texts that were only minimally related to the acquisition materials and that were drawn from books about the 20th century (i.e., the kind of naturally occurring new texts that you would like instruction to transfer to).

3. Comprehension of new texts constructed to link information on the topic of three of the cases from the acquisition text.

4. An essay on a topic (art) that was not mentioned at all in the acquisition text (a picture of an abstract expressionist composition by Kandinsky was presented with the question "How is this painting typical or not of trends of the 20th century?").

5. A remote associations test requiring subjects to draw meaningful connections between randomly paired 20th-century events and phenomena that had not been previously paired in the acquisition phase.

6. A test requiring the selection of a correct description of a prominent (but unfamiliar to high schoolers) 20th-century icon or symbol. For example: "Giacometti's sculptures of walking men have been taken by many to be symbols of the 20th-century's '*Existential Man*.' Were these walking men portrayed as (a) excruciatingly thin or (b) grotesquely fat?" Notice how this question has a correct answer (they were thin; and these sculptures have actually been chosen to represent existentialist aspects of the modern condition, for example, on the cover of Barrett's famous book on that topic, *Irrational Man*); note as well that it is possible to develop chains of reasoning that allow that answer to be figured out, but that there is no single, determinate path to that answer.

These preliminary results suggest that there is a fundamental choice in methods of learning and instruction. Conventional methods seem to produce superiority when measured by conventional tests that stress reproductive, fact-retention types of memory. The methods developed from Landscape Criss-Crossing Theory are not as successful at producing mindless, imitative recall. However, if one agrees that the goal of learning and instruction should be the acquisition of generative knowledge with wide application in novel but partially related contexts, then it would seem that methods like ours are far preferable to the conventional ones.

Toward a Science of Instructional Sequencing, Case Selection, and Case Arrangement

There is no science of instruction from cases. The best we have are occasional useful recommendations (e.g., Collins & Stevens, 1983). Where case-based instruction is practiced (mainly in professional schools and in some training programs), procedures tend to be haphazardly determined by the intuitions of curriculum planners. When there is a systematic basis for case selection, that basis is probably misguided: Cases tend to be picked when they neatly illustrate some instructional point that is being covered. The same features that make such "textbook cases" desirable for formal teaching (the way they form a clear-cut illustration of a topic that is being covered) are also likely to impede transfer, because in ill-structured domains true textbook cases are rare, and the intermediate, across-topic character of typical cases will be missed.

As important, there is currently no systematic basis for case sequencing. Whatever effects there are of learning individual cases, carry-over learning from case to case can vary in its quality. Clearly, not all case arrangements are equivalent in the intercase learning they promote. Why not just allow for a random order of cases, as naturally occurs in the accrual of real-world experience? The answer is that you do not want to have to wait so long for experience to accrue. It is very possible that with the right system of case arrangement, learning from examples can be made far more rapid, accurate, and efficient than with haphazard presentation orders. The development of basic principles of case sequencing could substantially shorten the acquisition time and number of cases needed to support widespread transfer.

Well- to ill-structured versus ill- to well-structured sequences. Part of conventional wisdom is that you start from the simple and then work toward the more complex (Glaser, 1984). In the terms of this paper, this would translate into a recommendation to first present material according to some well-structured abstract organizations, followed by the later introduction of complexities and irregularities. There are two dangers in this approach. First, it may inculcate a false impression that the domain is fundamentally simple and well-structured. Second, it may result in initial representations that are so neatly self-contained that they rigidly resist the complicated restructurings necessary to mirror the irregular contours of the domain. The problem with starting with an ill-structured presentation, on the other hand, is that it may be unrealistic to expect that subjects will be able to deal with complexity in material they are not sufficiently familiar with--the ingredients for flexible idea combination have to be established in memory before such combinatorial play can operate.

These two instructional orders for teaching the same material can be contrasted with a third approach, an intermediate degree of well-structuredness of initial presentations. We expect this to be the most effective in promoting transfer, because it is not subject in extreme degree to the problems that result from having one of the two approaches dominate early instruction. A mixture of well- and ill-structuredness in the early stages of learning should allow sufficient prerequisite material to become well learned prior to operating in a complicated fashion on that material, and at the same time avoid the establishment of an overly rigid representation that would be difficult to dislodge.

Before proceeding, it is very important to note that although we entertain different options for instructional sequencing, our position on the *goals* of learning is unequivocal. An *advanced stage of knowledge acquisition* will always be reached where learning criteria should no longer involve the demonstration of a superficial familiarity with subject matter and the memorization of some definitions and facts. Instead, at these more advanced stages, a learner should be required to get ideas *right* (even if that is hard to do) and to be able to appropriately apply those ideas (even if there are no simple formulas that they can memorize for doing so). If extra difficulty and confusion are the price of this shift in criteria, that is unfortunate but not an excuse for oversimplifying instruction. Wyas must be found to reduce the difficulties of the learner as much as possible without sacrificing the integrity of the subject matter being learned.

Altorithms for case-to-case sequencing for transfer. What is the most efficacious ordering of cases to produce knowledge representations that maximize transfer (i.e., structures that are multiply interconnected along multiple dimensions)? It should first be noted that the many available discourse or text analysis schemes provide little guidance here. Successive cases are unlike successive parts of a text in that each case is an integral entity, rather than a continuation of the preceding case. In this sense, each case is like a new text. Thus, structural analysis models employing, say, linear cohesion principles are inappropriate for analyzing intercase structure. Instead, we argue that several strands of thematic interconnectedness between cases must be considered in evaluating across-case structure. The metaphor of the "line" of development is then superceded by such others as "multistranded weaves," by the degree of overlap of many thematic elements (Wittgenstein, 1953).

Our starting point for identifying optimal case-arrangement algorithms involves the notion of intermediately related cases, of *partial overlap* across cases. We believe that it is important to neither group the cases that are most alike (and thereby promote the formation of spurious generalizations), nor to group cases in a way that maximally highlights their differences (thereby causing those regularities that do exist to be missed). Rather, we believe once again, as with case selection, that an intermediate course is to be preferred. We first seek a balance between continuity and discontinuity; some thematic features would overlap across adjacent cases and some would not; and which features overlap across cases would vary from one pair of cases to the next. Because ill-structured domains do not, by definition, permit single (or even small numbers of) connecting threads to run continuously through large numbers of successive cases, the notion of intermediateness of connection is intended as a first approximation to an alternative metric of strength of connection: "woven" interconnectedness. In this view, strength of connection derives from the partial overlapping of many different strands of connectedness across cases, rather than from any single strand running through large numbers of the cases (Wittgenstein, 1953).

By comparing the overlap of thematic features for each adjacently presented case, an index can be calculated of the extent to which cases will appear to be relatively similar (and thus expendable--abstraction away from the cases would be thought to be possible) or dissimilar (in which case attempts at abstractive generalization would be perceived as futile). This is the first part of a "Partial Overlap Index." According to our theory, this Index should optimally have a value between its mathematically derivable maximum and minimum. The second part of the Partial Overlap Index is calculated at a higher level of abstraction by looking at the relationships between the adjacent overlap relationships; successive case pairs should have an intermediate degree of overlap in the specific thematic features that they share (e.g., perhaps half of the overlapping themes for Cases 1 and 2 should be overlapping themes for Cases 2 and 3). The highest values of the Index will be attained when both individual case information and abstract perspectives retain their viability, but without either achieving dominance--they both add something, but neither can supplant the other.

Intermediate values on these two components of the Partial Overlap Index will allow important information about individual cases to be maintained and some partial abstractions to be formed. By maximizing the *joint function of case uniqueness and abstractability away from cases* (that is by treating cases as separate but overlapping), the best of both worlds is attained and the characteristics of complex real-world domains is most accurately mirrored.

Perhaps more importantly, high values of the Index (i.e., intermediate degrees of overlap) should produce mental representations that maximize the potential paths for going from case to case in trying to assemble an appropriate set of precedents from prior knowledge to most closely fit the needs for processing some new case in the future. Because transfer/application in domains that lack rules or general principles of wide application is dependent on such situationally dependent adaptive assembly processes, intermediate degrees of adjacent case overlap should be ideal.

Another model of case sequencing is similar in many respects to the Partial Overlap model. It seeks primarily to highlight points of difficulty and confusion in a knowledge domain (which is, of course, where transfer is hardest to effect), especially those points where inappropriate abstraction is likely to mask useful case information. In this model, cases are juxtaposed in such a manner as to achieve a balance between calling attention to (a) differences in superficially similar cases (to avoid having the complexities of individual cases lost due to a reductive assimilation to their apparently common features), and (b) similarities between apparently common features), and (b) similarities between apparently dissimilar cases (to avoid overly separating the mental representations of cases that may jointly figure as precedents for the analysis of some future case or that may be usable as analogies, one for the other).

Learning as a cycle of successive "mutual bootstraps" between case knowledge and abstract domain knowledge. Implicit in the ideal, discussed earlier, of intermediate degrees of overlap between cases is the value of maintaining a balance between cases considered as unique and to be subjected to close analysis, on the one hand, and systems for representing the partial abstractions that characterize the across-case domain, on the other hand. This suggests another approach to sequencing: A cyclical alternation between abstraction-centered presentations, in which cases illustrate or concretize the abstractions, and case-centered presentations, in which the same abstractions are now used in combined form to describe the cases. So, for example, the abstract theme of Fragmentation in the 20th century, illustrated by such cases as "Increases in Specialization" and "Multiple Influences from Media Exposure" in an abstraction-centered approach, would alternately be treated as an attribute, as part of the adjectival descriptor vocabulary for characterizing individual cases in a case-centered approach (i.e., "fragmentation" is a feature of the phenomenon of "Increasing Specialization," rather than the latter merely illustrating the former).

Each of these two kinds of systems should be able to help develop the other (hence the use of the expression, "mutual bootstraps"). As knowledge grows about how the complexities of individual cases are structured, that information should be useful in forming better abstract representational schemes for the domain that is the union of those cases. In other words, the more you know about cases in a domain, the better should be your systems for representing information about the domain that is itself, of course, constituted by those cases. At the same time, knowledge advances at the level of the entire domain should provide information useful in deriving insights for the development of better systems for representing the individual case (i.e., building a more adequate *mathesis singularis).* For example, the more you know about trends in the relationships between cases in a domain, the more ideas you will have for how to study and represent an individual case. Therefore, in this instructional approach, new information about the overall, macrostructure of a domain feeds into a next cycle of studying how best to represent the complexities of individual cases already studied, and so on, continuing the cycle. Again, the conjecture is that the two systems may be able to iteratively serve as "mutual bootstraps," each improving the other. Because both systems contribute to novel transfer/application, a self-perpetuating cycle of increasingly finer tuned improvements in the two systems would have obvious learning value.

The Role of Visual-Perceptual Representations and Adjunct Aids in Transfer

As is well known, the human perceptual system is very successful at representing large amounts of highly complex multivariate information at a glance (imagine all the information contained in an image of, say, a face), and storing that information in memory with great accuracy and durability. Studies have shown that hundreds of visual scenes presented in a short amount of time are recognized at very high levels of accuracy (Shepard, 1967; Standing, Conezio, & Haber, 1970). Furthermore, representations of visual information are rapidly comparable using an approximate, "family resemblance" type recognition process. So, the perceptual system is ideally suited to facilitating the representation of multidimensional complexity in individual cases, the subsequent recall of those cases, and the recognition of approximate family resemblances across cases. Because of these advantages, it may be beneficial to recode ill-structured representations into inputs acceptable for the perceptual system to operate upon, thereby reducing mnemonic overloads and fostering recognition-based assembly processes (Spiro & Myers, 1984).

A means for effecting this perceptual recoding is available in the work on integral visual displays for observation of complex data. For example, we are exploring the use of one of these display systems, Chernoff Faces (Chernoff, 1973), in our own work. Case information is coded along multiple thematic dimensions, and a partly arbitrary code maps thematic features to different features of the visual display (i.e., different features of emotionally neutral Chernoff Faces; more important conceptual features are assigned to more salient facial features). Thus, large amounts of case information are presented in a single integrated picture that can be grasped at a glance.

Furthermore, the rapidity of perceptual processing means that large amounts of case information should be learnable in a much shorter time than would otherwise be possible. And a variety of resemblances across cases can be detected using the common human ability to recognize family resemblances across faces.

Discovery Learning in Well-structured Domains

The knowledge gained from studying ill-structured domains should be extendable to discovery learning in well-structured domains. From the perspective of the learner, well-structured domains are ill-structured until the principles of well-structuredness are discovered. Therefore, our principles of flexibly interconnected knowledge representation should increase the chances of *noticing* likely candidate systems of organization for the domain and should permit sufficient old information to still be available in memory to develop and test those candidate systems.

Concluding Remarks

We have discussed various issues in an area of cognition and instruction that are currently poorly understood: how to get people to independently go beyond their specifically instructed knowledge. In response to this gap in knowledge, we offered a theory of case-based learning for transfer in ill-structured knowledge domains and suggested methods of case-based instruction to produce flexible knowledge representations.

From a practical point of view, the approaches we propose are not easy, and they may result in some increases in the time and effort required in initial instruction in a domain. However, we also expect that that investment will be more than justified by the fact that it will not be wasted instruction, limited in application to situations that happen to fit some narrow range of explicitly established preparedness. Wasted either because the instruction was too narrow in the case prototypes used, thus limiting applicability to resemblant cases; or wasted because oversimplified general principles and rules were taught, accounting for too little of the relevant variability in the knowledge domain. Systems of instruction must be developed that produce knowledge that can be flexibly adapted to the wide variety of new situations to which it will need to be applied, even at some additional early cost.

We know of no area of human endeavor that lacks an ill-structured aspect. Success in ill-structured areas tends to come only with a considerable accumulation of actual case experience. Application of the learning principles we have proposed has the potential to take material that is either taught poorly or not taught at all (and thus left to the vagaries of haphazard acquisition from "experience" over long periods of time) and, for the first time, make that material directly instructable.

References

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

Anderson, J. R., Kline, P. G., & Beasley, C. M. (1979). Schemata as scaffolding for the representation of information in discourse. *American Educational Research Journal, 15*, 433-440.

Anderson, R. C. (1977). The notion of schema and the educational enterprise. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Erlbaum.

Anderson, R. C., Spiro, R. J., & Anderson, M. C. (1978). Schemata as scaffolding for the representation of information in discourse. *American Educational Research Journal, 15*, 433-440.

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

Barsalou, L. W. (in press). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts reconsidered: The ecological and intellectual bases of categories*. Cambridge, MA: Cambridge University Press.

Bartlett, F. C. (1932). *Remembering*. Cambridge, MA: Cambridge University Press.

Britton, B. K., & Black, J. B. (1985). *Understanding expository text*. Hillsdale, NJ: Erlbaum.

Brown, A. L., Campione, J. C., & Day, J. (1981). Learning to learn: On training students to learn from texts. *Educational Researcher, 10*, 14-21.

Carbonell, J. G. (1983). Derivational analogy in problem solving and knowledge acquisition. *Proceedings of the International Machine Learning Workshop*, Champaign, IL.

Carnine, D. (1980). Three procedures for presenting minimally different positive and negative instances. *Journal of Educational Psychology, 7*, 452-456.

Chernoff, I. (1973). The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association, 68*, 361-368.

Collins, A., & Stevens, A. L. (1983). Goals and strategies of inquiry teachers. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 2). Hillsdale, NJ: Erlbaum.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models*. Hillsdale, NJ: Erlbaum.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 14*, 1-38.

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

Kolodner, J. L. (1980). *Retrieval and organizational strategies in conceptual memory: A computer model* (Research Rep. 187). New Haven, CT: Department of computer Science, Yale University, Lenat, D. B., Hayes-Roth, F., & Klahr, P. (1983). Cognitive economy in fluid task environment. *Proceedings of the International Machine Learning Workshop*, Champaign, IL.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*, 159-188.

Medin, D. L., & Schaffer, M. H. (1978). A context theory of classification learning. *Psychological Review, 85*, 207-238.

Medin, D. L., & Schwanenfluegel, P. J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 355-368.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35*, 113-138.

Minsky, M. (1979). The society theory. In P. Winston (Ed.), *Artificial intelligence: An MIT perspective*. Cambridge, MA: MIT Press.

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology, 16*, 371-416.

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

Rumelhart, D. E., & Norman, D. A. (1981). Analogical processes in learning. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

Schank, R. C. (1982). *Dynamic memory*. Cambridge, MA: Cambridge University Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1-66.

Scott, W. A., Osgood, D. W., & Peterson, C. (1979). *Cognitive structure: Theory and measurement of individual differences*. Washington, DC: Winston.

Shepard R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Memory, 6*, 156-163.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127-190.

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

Smith, S., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory and Cognition, 6*, 342-353.

Spiro, R. J. (1977). Remembering information from text: The "State of Schema" approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague, (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, NJ: Erlbaum.

Spiro, R. J. (1980a). Accommodative reconstruction in prose recall. *Journal of Verbal Learning and Verbal Behavior, 19*, 84-95.

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

Spiro, R. J. (1983). Subjectivity and memory. In W. Kintsch & J. Le Ny (Eds.), *Language comprehension.* The Hague: North-Holland.

Spiro, R. J., Crismore, A., & Turner, T. J. (1982). On the role of pervasive experiential coloration in memory. *Text, 2*, 253-262.

Spiro, R. J., & Esposito, J. (1981). Superficial processing of explicit inferences in text. *Discourse Processes, 4*, 313-322.

Spiro, R. J., Esposito, J., & Vondruska, R. (1978). The representation of derivable information in memory: When what might have been left unsaid is said. In D. L. Waltz (Ed.), *Theoretical issues in natural language processing* (TINLAP-2). New York: Association for Computational Linguistics and Association for Computing Machinery.

Spiro, R. J., & Myers, A. (1984). Individual differences and underlying cognitive processes in reading. In P. D. Pearson (Ed.), *Handbook of research in reading.* New York: Longman.

Spiro, R. J., & Taylor, B. M. (in press). On investigating children's transition from narrative to expository discourse: The multidimensional nature of psychological text classification. In R. Tierney, J. Mitchell, & P. Anders (Eds.), *Understanding readers' understanding.* Hillsdale, NJ: Erlbaum.

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures. Single-trial learning of 2500 visual stimuli. *Psychonomic Science, 19*, 73-74.

Streufert, S., & Streufert, S. C. (1978). *Behavior in the complex environment.* Washington, DC: Winston.

Waltz, D. L., & Pollack, J. (1985). Massively parallel parsing. *Cognitive Science, 9*, 51-74.

Winston, P. H. (1983). Learning by augmenting rules and accumulating censors. *Proceedings of the International Machine Learning Workshop*, Champaign, IL.

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

Footnotes

[1]These terms receive further explication later in the chapter and are clarified in a more general manner by their *use* throughout the chapter--as "family resemblance" concepts, part of their definition can be no more than implicit in a complicated network of similarities and differences across uses (Wittgenstein, 1953).

[2]It should be noted that the relatively harsh tone taken toward schema-type theories throughout this chapter should be understood in the qualified sense in which it is intended. Rather than a blanket condemnation of these theories, we intend only to point out their shortcomings in enabling transfer in certain fairly common situations characterized by irregular complexity. Much of our earlier work has been in the schema theory tradition, and we are well aware of the importance of schema-type approaches. Again, our claims against those approaches are limited, namely to aspects of knowledge domains that are "ill-structured" (keeping in mind that even when well-structured knowledge is involved, the context of its *application* is frequently ill-structured). The next step after the kind of work that we describe in this chapter will be to *combine* schema-type representations with the more flexible kind we are developing, because the two approaches seem to be natural complements, possessing compensating strengths and weaknesses. Also, the fact that we emphasize the less studied ill-structured domains does not imply that we discount the existence of domains with substantial regularity--they are just not the focus of our research.

[3]It is not necessary that criss-crosses of the landscape be exhaustive to produce generative structures. Rather, if enough of the topical landscape is portrayed, accurate anticipations of the structure of nonportrayed aspects will be possible. This can be seen by analogy to a situation in which one is presented with pictures of a man's face only when he has expressions of anger or happiness; one would be able to fairly accurately anticipate what that person would look like if he were sad, grateful, sleepy, and so on.

This page is intentionally blank.