ADVANCES IN THE PSYCHOLOGY OF HUMAN INTELLIGENCE

Volume 5

Edited by

Robert J. Sternberg
Yale University

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1989 Hillsdale, New Jersey Hove and London

natic human information processing:
l theory. *Psychological Review, 84,*

ice. In C. Sophian (Ed.), *Origins of*
Erlbaum Associates.
in solving physics problems. In R.
325–348). Hillsdale, NJ.: Lawrence

ambridge, MA.: Harvard University

al and the paradox of interference.

plied Psychology, 10, 1–36.
[. Fluid and crystallized abilities and
er (Eds.), *Intelligence and learning*

g, and analogical reasoning: The
Lawrence Erlbaum Associates.
intelligence. New York: Cambridge

the puzzler. *Psychology Today, 16,*

.evel theory of intelligent thinking.
ational Psychology].
nd cognition. In M. Eysenck (Ed.),
chester, England: Wiley.
reactions. *Journal of Experimental*

logical Review, 64, 49–60.
-structured problems. In M. T. H.
. Hillsdale, NJ.: Lawrence Erlbaum

l intelligence in the everyday world.
elligence (pp. 51–83). New York:

presentation in novice and expert
es, 14, 391–396.
nd: Methuen.

6 The Specific Character of Abstract Thought: Categorization, Problem Solving, and Induction

Douglas L. Medin and
Brian H. Ross
University of Illinois at Urbana-Champaign

I. INTRODUCTION

Although the typical layperson's view of intelligence includes a role for practical knowledge and common sense, he or she would probably consider the epitome of intelligence to be abstract, analytic thought. This ability is manifest in both deduction and induction. The truly analytic mind is as at home in evaluating the conclusion that "Some A's are C's" from the premises "Some A's are B's and some B's are C's" as it is in evaluating "Some dogs are feathered" from the premises "Some dogs are pets, and some pets are feathered." Similarly, on the basis of one or two concrete examples of a type of problem one might expect a very bright person to abstract out a general solution that could be applied to superficially different problems of the same type, while the rest of us may fail to see the connection.

The layperson's tendency to equate intelligence with abstract thought is shared by cognitive psychologists. We have repeatedly demonstrated the limitations of people as abstract, deductive reasoners and noted with chagrin the difficulty of producing transfer of training or generalized problem-solving skills. Pure, domain-independent, abstract thought remains as our standard. Thus in both categorization and problem-solving research, people are presented with numerous examples with the idea that abstractions (prototypes or schemata) will be made that will facilitate performance on new examples. The general idea is that mental representations take the form of abstract information bound up with concrete details and that transfer or generalization takes place when the irrelevant detail is discarded.

This appreciation of abstract thought carries over into the theories of categorization and problem solving. The abstract schema or prototype is taken to be the desirable end-state, and the concern is with how the learning of specific examples facilitates the transition to generalized expertise. Overall, then, there is a clear convergence of cognitive psychologists' and lay people's views of intelligence as abstract thought.

In the present chapter we argue for a different view. Although we do not deny that abstraction occurs, our thesis is that less abstract information, which we will refer to as the specific, also needs to be given its due. The idea that learning consists of discarding specific, superficial information in favor of general, abstract information may represent a dangerous oversimplification. We believe that the specific and the abstract are more tightly linked than the modal theory of classification learning or problem solving would imply. These linkages need to be spelled out in careful detail in order for psychologists to develop adequate theories of inductive learning, we would argue.

We are not the first to suggest that specific information is important in cognition. Lee Brooks and his associates (e.g., Brooks, 1978, 1987; Jacoby & Brooks, 1984; Vokey & Brooks, in press) have shown that nonanalytic reasoning by analogy to specific examples may be as powerful as, or more powerful than, more abstract, analytical reasoning. He would claim, for example, that one might classify some novel animal as a mammal not by reference to some abstract definition of mammalhood but because the unfamiliar animal reminds the observer of some familiar mammal such as a rabbit. A few years ago Brooks went so far as to organize a conference on "The Primacy of the Specific," paralleling Hajek's (1969) paper, "The Primacy of the Abstract."

Our thesis is not that abstraction does not occur, but rather that abstraction is more closely tied to specific examples than is commonly thought. This chapter is organized around three claims:

1. Reasoning is often case-based and often relies on specific examples rather than on more abstract knowledge.
2. Abstraction is not necessarily an autonomous process that operates on or performs computations on examples but is otherwise independent of them. Instead, abstraction often appears to derive naturally from exemplar comparison and use.
3. In general, induction is conservative in the sense that it preserves more information associated with examples than many theories of problem solving and categorization imply.

If these three claims concerning categorization and problem solving can be supported, then they suggest that context-free abstract representations are not the epitome of normal intellectual functioning. At first thought, the idea of having knowledge tied to particular contexts and examples appears to be a limitation. We do not deny that there are situations where transfer of training is limited by context-bound knowledge. On the other hand, specificity has its virtues. In particular, we argue that specificity may make access to and application of relevant knowledge easier, may permit graceful updating of knowledge, may protect the cognitive system from incorrect or inappropriate inferences, and may provide just the sort of context sensitivity that much of our knowledge should, in fact, have.

The rest of this chapter is organized as follows. We first review evidence from the literature on categorization and problem solving bearing on our claims of case-based or context-dependent knowledge, nonautonomous abstraction, and conservative induction. This forms the main body of our paper. We next consider some research in artifical intelligence and psychology that makes use of related ideas. Finally, we summarize the issues and implications of this specificity.

CONTEXT-RICH, CASE-BASED REASONING: PRIORITY OF THE SPECIFIC

Introduction

Imagine a clinical psychologist interviewing a new client who reports becoming extremely anxious whenever he gets on an elevator. How might the clinician bring to bear her knowledge that is relevant to treating this phobia? One possibility is that she has a mental representation of phobia that is a summary of typical characteristics or symptoms, usual treatments and prognosis. The clinician may then evaluate whether the new client displays or describes enough of these characteristic features to be classified as phobic. A second idea is that the clinician has a set of defining properties in mind that are shared by all phobics and not seen in nonphobics. The client will either satisfy or not satisfy this definition and accordingly be classified as either having or not having a phobia. Still a third possibility is that this new client might remind the clinician of a previous client who also suffered from claustrophobia. In this instance the associated mental representation might include a variety of relevant and irrelevant details concerning the course of the earlier treatment. For example, she might remember that the claustrophobia in the earlier instance turned out to be a side effect of some medication the client was taking. If this were so, the clinician might ask the new client about any medications being taken even if, in general, it were rare for phobias to be attributed to side effects of drugs.

The above three possibilities correspond to distinct theories concerning the nature of learning associated with experience with examples of categories. One view, the so-called prototype view, is that a natural part of the cognitive system performs a sort of "mental averaging" to derive summary representations of what is typical or characteristic of a category. The second view, known as the classical view, is more stringent in that it claims that there are defining properties that determine category membership in an unambiguous manner. That is, this view claims that categories have defining properties and that people's mental representations that are used in classifying consist of just these defining properties. Recently the classical view has been roundly criticized by cognitive psychologists as well as by philosophers, on the grounds that many of the categories we use are fuzzy, do not have clear definitions, and admit differences in goodness of examples based on number of typical properties (see Smith & Medin, 1981, for a detailed review). For instance, most people judge robins and sparrows to be better examples than penguins and ostriches of the category *bird*. Graded categories are consistent with the prototype view. The idea of mental averaging is consistent with differences in goodness of examples because the summary representation consists of characteristic or typical properties in addition to defining properties (if indeed any defining properties exist). The third view, often called the exemplar view, shares with the first view the idea that categories may be fuzzy, but it argues that examples of a category may be stored individually and used in "case-based reasoning." The extreme of this position is that there is no abstraction, and only information concerning particular examples enters into mental representations of categories.

A fair share of laboratory research on categorization over the last decade or so has been directed at contrasting the mental averaging and exemplar storage theories of categorization. In the area of problem solving, corresponding theories arise, although to our knowledge no one has argued that problem solving consists solely of storing examples. Still, there is the question of the degree to which mental representations associated with problem solving are abstract and removed from particular examples. We now briefly review categorization and problem-solving research bearing on "the priority of the specific."

Categorization: Prototypes Versus Examples

Natural Categories

Our discussion of mental representations in terms of features or properties presupposes some knowledge of just what those features are. Of course, there is no technology available today that would allow us to peer into people's minds and read off the details of mental representations. An alternative approach is to ask people to list properties of concepts. This strategy was followed by Rosch and Mervis (1975) in their pioneering studies that seriously undermined the classical view of conceptual structure. They asked people to list features of natural object categories such as *bird, fruit,* and *tool* as well as to provide properties of particular examples of these categories. In addition, people were asked to rate the goodness of example or typicality of examples. Rosch and Mervis found that people tended to list properties that were characteristic but not defining (e.g., for birds, *have feathers, fly, build nests in trees, sing*) and that people found rating typicality or goodness of example a natural and meaningful

d that people's mental repre-
ust these defining properties.
.cized by cognitive psycholo-
ıat many of the categories we
mit differences in goodness of
e Smith & Medin, 1981, for a
e robins and sparrows to be
the category *bird*. Graded
The idea of mental averaging
ımples because the summary
ıroperties in addition to defin-
exist). The third view, often
v the idea that categories may
y may be stored individually
of this position is that there is
articular examples enters into

tion over the last decade or so
raging and exemplar storage
lving, corresponding theories
that problem solving consists
stion of the degree to which
ving are abstract and removed
categorization and problem-
pecific."

rms of features or properties
atures are. Of course, there is
s to peer into people's minds
An alternative approach is to
ategy was followed by Rosch
ıat seriously undermined the
ed people to list features of
d *tool* as well as to provide
ies. In addition, people were
lity of examples. Rosch and
ıat were characteristic but not
nests in trees, sing) and that
ıple a natural and meaningful

task. Most important, however, they found that typicality rating could be accurately predicted from property listings—typical examples had many of the properties that were characteristic of the category, whereas atypical members had only a few of them. These findings are consistent with the idea that category representations consist of a set of characteristic properties and that categorization is based on a comparison of properties of examples with properties in the summary representation. This summary representation is referred to as a *prototype* and need not correspond to any specific example but rather corresponds more to an ideal or even a stereotype.

The work of Rosch and Mervis, as well as others, most notably Smith, Shoben, and Rips (1974), who proposed a specific model for typicality effects, inspired a considerable body of research on categorization involving natural object concepts. Not only do people rate robins to be better examples of *bird* than they rate ostriches, but people are also faster in verifying category membership ("true or false, a robin is a bird") of typical rather than of atypical category members (see Medin & Smith, 1984; Mervis & Rosch, 1981; Oden, 1987; for relevant reviews).

Artificial Categories

Although asking people to list properties of concepts provides suggestive evidence concerning fuzzy categories, one would not want to maintain that people can directly and accurately introspect on their mental representations. One way of addressing this issue is to artificially construct new categories, ask people to learn to correctly classify examples of them, and then present test probes designed to reveal something about resulting category representations. The examples presented during initial learning can vary in the number of typical properties they possess; the transfer tests usually involve both old and new examples, including the category prototype.

Initial work with artificial categories provided evidence converging on the idea that experience with examples leads to the abstraction of a prototype representing the central tendency of a category. This converging evidence consists of three principal findings:

Typicality Effects. The first robust result is that the more typical properties an example has, the faster and more accurately it is classified (e.g., Barresi, Robbins & Shain, 1975; Goldman & Homa, 1977; Homa & Vosburgh, 1976; Posner & Keele, 1968). This is true for both the original learning of examples and for transfer tests involving new examples.

Performance on Prototypes. On transfer tests given at the end of learning, people classify prototypic patterns as accurately or even more accurately than they classify the old examples on which they were trained (e.g., Homa &

Chambliss, 1975; Homa & Vosburgh, 1976; Posner & Keele, 1968). In addition, transfer performance on new examples is well predicted by the similarity of a new example to the category prototype.

Differential Retention. When delays on the order of several days are inserted between learning and transfer tests, significantly greater forgetting is observed for the old training stimuli than for the prototype and other new patterns (Goldman & Homa, 1977; Homa & Vosburgh, 1976; Posner & Keele, 1970; Strange, Keeney, Kessel, & Jenkins, 1970). These results are consistent with the ideas that judgments are based on a mixture of specific item- and category-level information and that the specific item information is forgotten more rapidly than is the abstract information. As the retention interval increases, judgments are increasingly likely to be based on the prototype.

These results appear to provide strong evidence for prototype abstraction. In this view, experience with examples of a category activates a mental averaging process that more or less automatically abstracts out the prototype. This process operates on examples and, in a sense, discards them. Some theorists have assumed that on transfer tests an old example may be classified on the basis of direct retrieval of information concerning that example, but new examples are assumed to be classified by reference to the prototype.

Challenges to the Prototype View

The idea that concepts are organized around prototypes has not gone unchallenged. The most radical alternative to the prototype view is that no abstract representation is created and that classification decisions are based solely on analogy to specific stored examples.

Brooks (1978, 1987) has been a strong advocate of the exemplar view of categorization. In defense of this view, he puts forth two general arguments. The first argument is that, outside of formal learning contexts, human learning is more analogical than analytical. That is, learning does not consist of explicit attempts to develop and evaluate rules and hypotheses but rather consists of attempts to apply previously stored knowledge to current contexts.[1] According to Brooks, this analogical learning is best served by storing examples rather than by forming abstractions. The second argument is that nonanalytic learning serves organisms very well. Brooks supports this general argument with many clever examples that reveal advantages of preserving information associated with examples. The general claim is that unless one knows exactly how information will

[1]Reber (Reber 1967, 1969; Reber & Lewis, 1977) has argued for nonexplicit or unconscious rule learning, but this claim has been confronted with strong counterarguments (e.g., Dulany, Carlson & Dewey, 1984; Vokey & Brooks, in press; and see also Reber, Allen, & Regan, 1985; and Dulany, Carlson, & Dewey, 1985). In any event, we do not discuss this issue further since it is not central to our claim.

r & Keele, 1968). In addition, edicted by the similarity of a

r of several days are inserted greater forgetting is observed ype and other new patterns 976; Posner & Keele, 1970; results are consistent with the cific item- and category-level s forgotten more rapidly than val increases, judgments are

for prototype abstraction. In activates a mental averaging it the prototype. This process them. Some theorists have be classified on the basis of mple, but new examples are otype.

prototypes has not gone un- otype view is that no abstract ecisions are based solely on

ate of the exemplar view of 1 two general arguments. The contexts, human learning is does not consist of explicit theses but rather consists of irrent contexts.[1] According to oring examples rather than by t nonanalytic learning serves argument with many clever ormation associated with ex- exactly how information will

d for nonexplicit or unconscious rule arguments (e.g., Dulany, Carlson & Allen, & Regan, 1985; and Dulany, issue further since it is not central to

be used later on, it makes little sense to discard information based on very local information. Indeed, local contexts may be part of the information that it is important to preserve. For instance, an important aspect of one's knowledge about mushrooms is that although mushrooms purchased in a grocery store are nonpoisonous, mushrooms encountered in a woods may be poisonous.

Exemplar Accounts of Categorization. It may not be obvious, but exemplar-based models of categorization can account for the three main results from research using artificial categories that had been taken as providing support for prototype theories of abstraction. Medin and Schaffer (1978) proposed an exemplar-based model of categorization as an alternative to prototype models. In their *context model,* category judgments are assumed to be based on the retrieval of stored examples. The model assumes that similarity between a test probe and a stored example determines the likelihood that a stored example will be accessed. This means that the judgments will be influenced most heavily by the most similar stored examples. Let us see how the model accounts for findings that seem to call for prototype abstraction.

First of all, the context model predicts typicality effects because a new example having typical features is more likely to be highly similar to numerous learning examples than a new example having atypical features. Second, the prototype is almost always the pattern having the greatest number of highly similar category members, and it also is unlikely to be similar to members of contrasting categories. Medin and Schaffer (1978) showed that a mathematical model based on their theory could indeed predict better performance on the prototype than on specific training examples. Finally, predictions concerning differential retention depend on the particular assumptions made about forgetting. Medin and Schaffer assumed that retention loss corresponds to a loss of distinctiveness of the values of components of stored examples. Under this assumption, performance on old patterns will suffer more over a retention interval than performance on prototypes or new patterns, because of the decreasing likelihood that an old example as a probe will successfully access its own representation. Hintzman and Ludlam (1980) ran computer simulations of forgetting in an exemplar-based categorization model and showed that the predicted differential retention function for the model corresponds closely with observed retention functions. In brief, exemplar-based models can account for all of the main results that had been claimed to support prototype models.

Exemplar Versus Prototype Models. Although the just described results show that exemplar models can account for results that had been taken to implicate prototype models uniquely, they do not provide direct contrasts of the models. The context model implies that typicality effects are a by-product of similarity to examples, not similarity to a prototype. The model predicts that with distance of an example from the prototype held constant, performance will vary (improve) with the number of stored exemplars similar to the probe in question. In a series of four experiments, Medin and Schaffer (1978) found clear support for this density prediction, and the context model provided a better quantitative account of transfer test data than did prototype models.

We will consider further findings comparing prototype models with exemplar models in the section on conservative induction. So far we wish only to claim that results that appear to require theorists to posit abstract representations can be handled by models that assume more specific information is the basis of performance. We now extend our argument by turning to research that has looked specifically at remindings. This work comes from the area of problem solving where the conventional wisdom has been that experience with examples leads to content- and context-independent abstract representations.

Remindings in Problem Solving

As in the work on categorization, much of the focus of research in problem solving has been on the use of abstract knowledge. At first, the pioneering work of Newell and Simon (1972) investigated general heuristic search methods that might underlie intelligence. Their view of problem solving was that it consisted largely of search through a problem space, so that any general search methods people had for exploring such a problem space would be useful across all types of problems. The work on expert systems (cf. Barr & Feigenbaum, 1981) and the study of experts and novices in various domains (e.g., Chase & Simon, 1973), however, suggested that much of what underlies improved performance for these experienced individuals is not better general search methods but rather more domain-specific knowledge. Thus, much of the current work on problem solving is concerned with trying to characterize this domain-specific knowledge, how it is used, and how it is acquired. A common view is that experience leads to the establishment of domain-specific schemata that alert the expert to a known pattern and also contain the associated procedures (e.g., Chi, Feltovitch, & Glaser, 1981). Although this perspective represents a considerable departure from domain-independent search methods, it still leads one to focus on the idea that some abstract procedures are being applied to the current problem.

Although the view of problem solving as schema instantiation may be appropriate for experienced individuals solving familiar types of problems, it still leaves open the question as to how less familiar types of problems are solved. Two broad classes of possibilities may be distinguished. First, people may rely on general methods, often called "weak methods" because they only make use of domain-independent knowledge. This approach involves starting from scratch, in that no domain knowledge is employed. Second, people may rely on domain knowledge that is not perfectly relevant to the current problem, but will allow them to at least get a good start on the solution. This latter view comes from the belief that much of what we learn is learned from making use of extant knowl-

·obe in question. In a series
ɔund clear support for this
. better quantitative account

ɪtype models with exemplar
ı far we wish only to claim
stract representations can be
ıation is the basis of perfor-
:o research that has looked
he area of problem solving
ence with examples leads to
ations.

cus of research in problem
At first, the pioneering work
euristic search methods that
solving was that it consisted
any general search methods
l be useful across all types of
Feigenbaum, 1981) and the
g., Chase & Simon, 1973),
ɪroved performance for these
:h methods but rather more
ent work on problem solving
ı-specific knowledge, how it
that experience leads to the
lert the expert to a known
s (e.g., Chi, Feltovitch, &
ıts a considerable departure
:ads one to focus on the idea
ı the current problem.
ı instantiation may be appro-
r types of problems, it still
'pes of problems are solved.
.shed. First, people may rely
ecause they only make use of
volves starting from scratch,
, people may rely on domain
rent problem, but will allow
is latter view comes from the
making use of extant knowl-

edge structures or procedures (e.g., Remelhart & Norman, 1981; such a view also derives from the Piagetean notion of accommodation). A principal type of knowledge that people have about a domain in which they do not have expertise is about examples that they have seen earlier. That is, much experience, especially for formal domains in usual instructional settings, consists of the presentation of examples, both worked-out ones and ones for the student to solve. The claim of this section is that these examples are a major source for solving later problems in the domain.

Reliance on Earlier Examples

This reliance on earlier examples is widespread. When students are given problems to solve for homework, a common strategy is to look through the textbook for worked-out problems that "look like" the current problem. In less formal settings, problems may also often remind us of earlier problems from which we may take some or all of the solution. When we have no well-developed schemata for the current problem and are not reminded of any particular earlier problem, a common strategy, as in the just mentioned textbook searchers, is to try to think of earlier examples that might be relevant. Indeed, books on heuristic problem solving explicitly suggest such a method (e.g., Polya, 1945; Wickelgren, 1974).

Much of the work on how earlier examples are used by analogy has focused on understanding how to transfer aspects of the earlier problem that are appropriate and not transfer aspects that are inappropriate or irrelevant (e.g., Anderson & Thompson, in press; Burstein, 1986; Gentner, 1983; Holyoak, 1985; Winston, 1980). We will refer to the problem aspects that affect the solution as structural and those that do not as superficial. A reasonable view of transfer, based on the idea that abstraction is natural, is that people use the structural aspects to determine what knowledge to transfer from the earlier example and that the superficial information is discarded. We shall see, however, that transfer is more complicated than this view implies.

Ross (1984, Experiment 2) provides one demonstration of how people rely on earlier examples. In that study, novices learned six principles of probability theory (e.g., permutations, conditional probability) along with a word problem example to illustrate each principle. Each of these word problems made use of a different superficial story line content, such as golfers at a tournament, weather forecasting predictions, or booths at a coin fair. After learning these principles, the learners were asked to solve new word problems. These word problems varied in their superficial similarity to the problems that illustrated the principles. Two problems had story line contents that were similar to the story lines used to illustrate the same principle. Thus, if permutations had been illustrated by the golf story line, the test problem for permutations would have a story line about golf as well. To the extent that these learners were using the earlier examples, these superficial similarities should facilitate performance because they made it easier to access the appropriate earlier problem (thus, this condition is the *appropriate* condition). Two problems had story lines unrelated to any earlier problems. This *unrelated* condition serves as a baseline from which the other effects may be judged. Finally, two test problems had story line contents used in illustrating word problems, but they had illustrated a different principle than the one being tested. For instance, if a boat race word problem had been used to illustrate conditional probability, a test problem for the probability of either or both of two independent events might be about a boat race. To the extent that subjects were making use of the earlier examples on the basis of superficial similarity, this use would be *inappropriate* for these problems and would lead to lowered performance relative to the unrelated condition. The results showed clear evidence that these superficial similarities were being used. The proportions of correct answers for the appropriate, unrelated, and inappropriate conditions were .77, .43, and .22, respectively. Thus, the same story line for the same principle led to higher performance than the baseline, whereas the same story line for different principles led to lower performance.

Before examining in detail the effect of earlier examples, it is important to distinguish two principal components of how this effect may occur: access and use. First, some earlier example has to be accessed, noticed, or selected. This selection may occur from memory or from a written record such as a textbook. If no earlier example is accessed, clearly one cannot make use of it. The Ross (1984) experiment just described used superficial similarity to bias access. Second, once some example is selected, it must be applied to the current problem. Before bringing out the implications of these observations, we turn to some more direct evidence.

Access. Before an example can be used by analogy, it must be selected or accessed. How are earlier examples chosen to be used in an analogy? Two main means have been examined. First, a teacher might supply the earlier example to use (e.g., Gentner & Gentner, 1983; Rumelhart & Norman, 1981). Although analogies may be provided during some formal instruction, they are rarely provided during much of the intensive problem-solving periods nor are they provided if no teacher is available (i.e., during self-instructed learning). Second, the learner might be reminded of an earlier example. How might this reminding work?

One possibility for how remindings work is that the learner's analysis of the structural aspects of a problem reminds him or her of another example with the same or similar structural analysis (e.g., Carbonell, 1983). This view claims that structural aspects of a problem are the *only* aspects used in remindings. A second possibility is that both superficial and structural aspects affect remindings. Although it might seem useful to have remindings based on purely structural aspects, the evidence is quite convincing that both superficial and structural

mance because they made it
(thus, this condition is the
ınes unrelated to any earlier
seline from which the other
ıd story line contents used in
a different principle than the
l problem had been used to
r the probability of either or
ɔoat race. To the extent that
on the basis of superficial
problems and would lead to
ıdition. The results showed
were being used. The pro-
ınrelated, and inappropriate
ıs, the same story line for the
baseline, whereas the same
formance.
examples, it is important to
effect may occur: access and
d, noticed, or selected. This
record such as a textbook. If
t make use of it. The Ross
l similarity to bias access.
t be applied to the current
ıese observations, we turn to

logy, it must be selected or
ed in an analogy? Two main
supply the earlier example to
& Norman, 1981). Although
instruction, they are rarely
olving periods nor are they
instructed learning). Second,
e. How might this reminding

the learner's analysis of the
of another example with the
1983). This view claims that
ısed in remindings. A second
aspects affect remindings.
s based on purely structural
th superficial and structural

aspects affect what earlier example is thought of (e.g., Gentner & Landers, 1985; Holyoak & Koh, 1987; Ross, 1984, 1987, 1989). Superficial similarity manipulations can have very large effects on access.

These results indicate that the access of earlier examples is affected by multiple sources of knowledge. The idea that various sources might be combined during retrieval is a common one in memory theories (e.g., Anderson, 1983; McClelland & Rumelhart, 1985; Raiijmakers & Shiffrin, 1981), and a number of researchers have suggested that the usual memory retrieval processes are used in these remindings (e.g., Anderson & Thompson, in press; Holyoak & Koh, 1987; Ross, 1984, in press; Schank, 1982). Further evidence indicates that these remindings are sensitive to factors that usually affect memory retrieval. For instance, the retrieval depends upon the relative similarity of the current problem to the earlier example (Ross, 1987, Experiment 2) and can be shown to exhibit standard interference effects (Ross, 1984, Experiment 3).

We believe this effect of superficial similarity on access of earlier examples is another instance of the reliance on specific information. Although in some cases it is possible that this effect occurs because people cannot distinguish superficial and structural aspects (Holyoak, 1985; Holyoak & Koh, 1987), this explanation does not cover many of the cases. For instance, Hinsley, Hayes, and Simon (1977) showed that proficient algebra word problem solvers make use of superficial characteristics in problem solving. These experts have noted correlations between superficial and structural characteristics that they may use in helping to predict the appropriate schema from the readily available superficial information. The subjects would sometimes be able to make a good guess as to how the problem would be solved simply by hearing the first couple of words (e.g., the riverboat). As Lewis and Anderson (1985) noted, experts may have learned when it is all right to rely on superficial features. Such use of superficial information may be helpful in efficient artificial intelligence programs as well. For instance, Shavlik, DeJong, and Ross (1987) argued that relying on the most abstract characterization of a principle that can be used greatly hampers the speed with which a system may solve a problem. Programs that keep various special cases of this abstract characterization can be more quickly accessed and applied to problems than programs that rely on abstract characterizations alone.

Why does superficial information have such large effects on access? Although we may only speculate, two main reasons can be suggested. First, as just mentioned, these superficial aspects may be predictive of structural aspects and be more readily available. In many domains, entities that are superficially similar tend to share deeper similarities (Medin & Ortony, in press). Even when there is no necessary reason for such correlations, as in the domain of formal problem solving, correlations between surface and structural properties tend to be present nonetheless (e.g., Lewis & Anderson, 1985; Mayer, 1982). Second, it appears that the representation of earlier problems may include much of the superficial information. Viewing these remindings as part of the usual memory retrieval process, one might expect the inclusion of superficial information in the memory for an earlier example to evoke this example under various circumstances. For novices, who do not have a good understanding of many of the structural aspects of a problem (e.g., Chi et al., 1981; Silver, 1981), there may be little alternative but to make use of superficial aspects. Even for experts, however, who presumably do encode the structural information in earlier examples, remindings may occur if the relative similarity between the current and earlier example is high enough. For instance, if an expert problem solver solved a problem about aliens on the planet Quarck, a test problem about the Planet Quarck may well remind the expert of the earlier example (though, of course, the expert may or may not make use of this reminding). Usually, however, problems are not so distinctive and because experts have experienced so many examples, the current test problem may not bring any particular one to mind (at least with enough clarity for the expert to realize he or she is thinking of an earlier example). From this view, at least part of the drop-off in remindings that is assumed to come with expertise may be due to interference among the many possible remindings (Thorndyke & Hayes-Roth, 1979, presented a related interference view of schema acquisition).

Use. Perhaps superficial similarities have effects on how people access potentially relevant knowledge but not on how they use it. A number of current theories allow for the superficial similiarity effects on access but claim little if any effect of superficial similarity on use (e.g., Anderson & Thompson, in press; Holyoak, 1985). These theories assume that once an earlier example is selected for analogical use, the structural aspects of the earlier example are used to constrain the analogy that is made. Thus, what actually gets kept from a potential analogy, according to these theories, is some structural-related measure such as functionality (Anderson & Thompson, in press; Kedar-Cabelli, 1985), relevance (Holyoak, 1985), or systematicity (Gentner, 1983).

The idea that only structural information is involved in the use of analogies is consistent with strong forms of abstraction. Some recent evidence, however, suggests that this view is not an accurate one. Gentner and Toupin (1986) showed that the transparency of the object correspondences across stories (i.e., whether similar characters played similar or different roles across the stories) affected mapping performance. If the structural correspondences matched the superficial correspondences, performance was excellent; if not, performance was poor. Ross (1987, 1989) showed that in solving probability theory problems, object correspondences have large effects on performance, even when the appropriate formula is provided. That is, subjects tended to assign objects in the test problem to the variable roles that similar objects occupied in the study example (Ross, 1987). This use of superficial similarity occurred even when the two problems were superficially very different in other respects (Ross, 1989). Theo-

l information in the memory
various circumstances. For
any of the structural aspects
here may be little alternative
erts, however, who presum-
examples, remindings may
and earlier example is high
lved a problem about aliens
et Quarck may well remind
, the expert may or may not
blems are not so distinctive
examples, the current test
(at least with enough clarity
earlier example). From this
at is assumed to come with
many possible remindings
elated interference view of

n how people access poten-
se it. A number of current
on access but claim little if
erson & Thompson, in press;
n earlier example is selected
earlier example are used to
lly gets kept from a potential
ural-related measure such as
ar-Cabelli, 1985), relevance

ved in the use of analogies is
recent evidence, however,
Gentner and Toupin (1986)
ndences across stories (i.e.,
ent roles across the stories)
rrespondences matched the
lent; if not, performance was
robability theory problems,
nance, even when the appro-
to assign objects in the test
cupied in the study example
curred even when the two
espects (Ross, 1989). Theo-

ries that try to account for how analogies are used without any reference to superficial similarities are unable to account for such effects.

Again, this effect of superficial similiarity is not limited to novices. Hinsley et al. (1977) showed that their proficient algebra problem solvers used very different solution procedures for word problems with typical and atypical contents, even though these problems were structurally identical.

This work on remindings in problem solving indicates a strong reliance on specific content-rich examples during learning. We are not claiming that experts have no problem schemata and rely totally on specific examples. However, the heavy use of earlier examples by novices and the finding that experts still can make use of superficial similarities that have proven useful in the past suggest that any understanding of the acquisition and representation of these schemata is going to need to incorporate this use of and learning from specific examples. In the next section, we begin to examine how abstractions may arise.

NONAUTONOMOUS ABSTRACTION

Introduction

Up to this point, we have tried to convince the reader that people are making use of knowledge other than high-level abstractions. In this section, we will start by assuming that some abstracted knowledge (i.e., more abstract than previous instances) exists and is used in categorization and problem solving. Evidence for this assumption will be given shortly. If abstractions are used, two questions require answers. First, how do these abstractions come about? Second, what is the nature of these abstractions? Although these questions are not strictly independent, this section will primarily deal with the first question, whereas the next section will deal primarily with the second question.

We may distinguish two basic views of how abstractions come about. First, a common view is that abstractions come about automatically from commonalities across instances. By this account, one can imagine an autonomous abstraction engine scouring knowledge, devouring instances, and spewing forth abstractions. From a perspective that abstract knowledge is the goal of a system, such a view has much to recommend it, since abstractions are automatically taken care of by a mechanism whose sole purpose is to abstract. A number of rule induction systems in artificial intelligence and cognitive science have this view of abstraction (e.g., Anderson, Kline, & Beasley, 1979; Carbonell, 1983; Michalski, 1983). A common scheme of this type is to claim that maximal common generalizations (essentially abstractions that include features common to the instances) are made over all instances. In addition, although not explicit, such a view appears to underlie both associationist and prototype views of category learning. In associationist views, instances are assumed to combine in some way such that features common to members of the category are strengthened (i.e., the association between these features and the category are strengthened) and features distinctive to some members are weakened. The prototype view has not generally been concerned with specifying how prototypes are acquired, but a similar process is assumed to occur. After a number of instances have been experienced, the person is assumed to have a representation of the category that consists of some central tendency for each feature, such as those values that occur most often in the instances.

A second, rather different, view of how abstractions come about will be referred to as *nonautonomous abstraction*. By this account, abstractions come about selectively, from the use of one instance (or abstraction) being applied to another instance. Although the specific schemes vary, the basic idea is that commonalities are necessary, but not sufficient, to lead to abstractions. It is also necessary that the commonalities be used before an abstraction is made. This view eschews making all possible abstractions for the policy of making only those that are useful at least once. To better understand the implications of this nonautonomous abstraction, it is important to note that one now needs to specify what are the conditions under which one instance might be used in processing another. These conditions are exactly what the first part of the chapter has been concerned with, reliance on the specific. The use of an earlier instance, which we have tried to argue is a basic processing principle, not only is effective for processing a current instance but also has later effects through the abstraction that is made as a by-product of this use. We illustrate this idea with two sets of experimental results from our individual research programs.

One research project, conducted by Ross, Susan Perkins, and Patricia Tenpenny (Ross, Perkins, & Tenpenny, 1988) has investigated abstractions evolving from exemplar comparisons. The first study showed that if people use similarity between two items to classify the new item, later performance is also affected. Subjects learned arbitrary facts about four people. For example "Shirley likes ice cream and bought nails at the store" and "Julia likes canaries and bought a swimsuit at the store." After learning these facts, subjects were told that two of the people were from Group 1 (say, Shirley and Julia) and two were from Group 2. They were then given descriptions of new people and asked to say to which group they belonged, One of these new people was a person who "bought wood and a towel." Half of the subjects also heard that the person liked *sherbet*, whereas the other half heard that the person liked *parakeets*. We reasoned that these facts would tend to lead subjects to use similarities with either Shirley or Julia to categorize the new people. (Of course, in either case the new example would go into Group 1. The interesting question is what happens next.) If subjects connected the new example with Shirley (because of the sherbet/ice cream similarity), then they might also notice the relation between nail and wood. If they linked the example with Julia (because of parakeet/canary), then they might also notice the relation between swimsuit and towel. Finally, a list of items was presented, and subjects ranked how likely a person from each group was to buy the items. The list included an item related to nail and wood (chisel) and an item related to swimsuit and towel (sunglasses).

The logic of this test is perhaps easiest to see in the more abstract description provided in Figure 6.1. The new category member either shared two features with Shirley and one with Julia or shared one feature with Shirley and two with Julia. The critical test contrasts the category relevance of f2 with f4. Note that both of these features have appeared twice in the category. Therefore, neither the models that assume that feature frequencies are abstracted nor the pure exemplar storage models have any basis for predicting that either feature will be more important than the other. A more dynamic model of reminding does, however, predict a difference. If being reminded of Shirley leads to abstractions related to f1 and f2, and being reminded of Julia leads to abstractions related to f3 and f4, then for subjects reminded of Shirley, f2 should be more important than f4, and for subjects reminded of Julie, the reverse should hold. This is exactly the result obtained by Ross, et al. (1988), suggesting that our hypothesized exemplar comparison process is a key aspect of abstraction.

One alternative possibility, however, is that the reminding does not lead to an abstraction but simply makes the individual (and her associated properties) more memorable. By this account, if one is reminded of Shirley, then the features associated with Shirley, ice cream and nails (f1 and f2), are well remembered. At the final ranking, chisel is chosen over sunglasses because nail is better remembered than swimsuit, not because of any abstraction concerning woodworking. Ross et al. (1988) tested this possibility by having the first test item still lead to a reminding (by sherbet or parakeet), but the wood and towel items were replaced with two other unrelated shopping items. If the reminding is having its effect by

Study			
	Shirley		*Julia*
	f1		f3
	f2		f4
	Group 1		Group 1
First test (Categorization)			
	f1		f3
	f2	*OR*	f2
	f4		f4
Final test (Ranking)			
		f2	
		f4	
		filler	
		filler	

FIG. 6.1. Partial design of Ross, Perkins, and Tenpenny Study.

making the earlier individual more memorable, the use of these unrelated shopping items should not change the effect. If, however, the reminding is having its effect by leading to an abstraction from the earlier features and test features, then the use of unrelated shopping items should lead to no effect. No effect was obtained, undermining the idea that the earlier results were due to exemplar strengthening. Instead, the results support the abstraction-from-comparison account.

A second line of work by Medin and Edelson (1988) showed exemplar comparison effects, albeit more indirectly. Medin and Edelson were concerned with the use of base rate or frequency information in simulated medical diagnosis. When people are given ambiguous information that is consistent with either a common or a rare disease, they ought to think it is more likely that the common disease is present. Hence the aphorism in medicine, "When you hear hoofbeats, think of horses, not zebras." Medin and Edelson taught people to classify common and rare diseases on the basis of symptom information. A symptom was either unique to a disease or appeared with both a common and a rare disease. For example, people might learn that a common disease, burlosis, was associated with the symptoms of *headaches* and *dizziness* and that a rare disease, midosis, was associated with *headaches* and *blurred vision* (see Fig. 6.2). After people learned to classify examples of these (and other) diseases from symptom information, they were then presented with new patterns of symptoms and asked to give their best judgment (again see Fig. 6.2). One test involved the symptom of *headaches* alone, and people were expected to predict that the more common disease, burlosis, was more likely to be present. They did. On an equally ambiguous test involving *dizziness* and *blurred vision*, however, people showed a strong tendency to predict the rare disease, midosis. Finally, when all three symptoms were present people once again tended to predict that the more common disease was present.

These results are quite puzzling, both from the point of exemplar storage theories and models of abstraction. Exemplar storage models would predict that all three tests would lead to judgments that the more common disease was present. Many models of abstraction imply that people would ignore the symptom *headaches* because it is not a reliable predictor of the disease and that people would learn that *dizziness* goes with burlosis whereas *blurred vision* goes with midosis. These abstraction models would imply no preferences on the three ambiguous tests.

The results do make sense, however, when abstraction is tied more closely to exemplar comparison processes. The main idea is that the more frequent or common disease is usually learned first (because it appears more often) and that symptoms of midosis, the rare disease, may remind the learner of the common disease, burlosis (because they share the symptom, *headaches*). If people were reminded of the common disease and responded by analogy (that is, indicated that burlosis was present), their classification would be incorrect. Medin and Edelson assumed that when people are reminded of an earlier example, respond

the use of these unrelated
however, the reminding is
the earlier features and test
should lead to no effect. No
e earlier results were due to
pport the abstraction-from-

n (1988) showed exemplar
and Edelson were concerned
on in simulated medical di-
ation that is consistent with
ink it is more likely that the
medicine, "When you hear
d Edelson taught people to
of symptom information. A
d with both a common and a
a common disease, burlosis,
nd *dizziness* and that a rare
and *blurred vision* (see Fig.
ese (and other) diseases from
th new patterns of symptoms
. 6.2). One test involved the
ected to predict that the more
present. They did. On an
red vision, however, people
e, midosis. Finally, when all
ded to predict that the more

e point of exemplar storage
ge models would predict that
more common disease was
ople would ignore the symp-
of the disease and that people
eas *blurred vision* goes with
no preferences on the three

action is tied more closely to
s that the more frequent or
appears more often) and that
d the learner of the common
headaches). If people were
y analogy (that is, indicated
ld be incorrect. Medin and
an earlier example, respond

	Symptoms: earaches, dizziness *1.* burlosis *2.* namitis *3.* terrigitis *4.* coralgia *5.* althrax *6.* buragamo
Choice? 3 Correct diagnosis is coralgia	
	Symptoms: skin rash, sore muscles *1.* terrigitis *2.* burlosis *3.* althrax *4.* namitis *5.* buragamo *6.* coralgia
Choice? 2 That's correct!	
	Symptoms: back pain, earaches *1.* coralgia *2.* althrax *3.* buragamo *4.* terrigitis *5.* namitis *6.* burlosis
Choice? 1 Correct diagnosis is terrigitis	

FIG. 6.2. Illustrative sequence of trials from the Medin and Edelson (1988) Experiments.

by analogy, and are incorrect, they then pay special attention to what is novel or different about the current example. That is, they would attend to *blurred vision* more than *headaches*. If people followed this strategy, they would, in effect, learn that burlosis is characterized by *headaches* and *dizziness* and that midosis is characterized by *blurred vision*. On the test consisting of *blurred vision* plus *dizziness*, people would predict midosis is more likely because the symptom *blurred vision* had received more attention and because the symptom of *headaches* associated most closely with burlosis was missing. When the symptom *headaches* is added, preferences switch to the more common disease, as expected. In short, the idea of reminding and exemplar comparison can account for all three main results. Although Medin and Edelson do not offer the sort of direct evidence for exemplar comparison that was adduced by Ross, et al. (1988), neither pure abstraction models nor simple exemplar storage models can account for these results.

The theme of both of these research projects, then, is that the use of instances during categorization has effects not only on the categorization of the current instance, but on one's representation of the category. In particular, these exemplar-based categorizations influence the abstractions that are made by the learner. We believe that this evolution of abstractions as a by-product of exemplar comparison and use underlies much of our knowledge. In the next section we present more detail and evidence concerning the form of these abstractions.

IV. CONSERVATIVE INDUCTION

A. Introduction

So far we have argued in favor of two principal claims. The first is that categorization and problem solving rely more on specific examples than many accounts of learning have implied. The second claim is that abstraction is not the product of some autonomous process but rather that abstraction arises as a by-product of the use of examples. If people frequently use the specific in reasoning, and abstraction comes from use, then the form of these abstract representations may be constrained by use in an interesting manner. In particular, induction should be *conservative*. By conservative we mean that, in general, abstractions will, on the one hand, preserve more than the minimal information necessary to perform a task and, on the other hand, not add information that is not supported by data. To give a simple-minded example, suppose that people were required to correctly classify examples of the categories *elephant* and *zebra*. Single dimensions like size and color might be used to achieve perfect classification, and it is conceivable that the resulting mental representations would be as impoverished as "elephants are big" or "zebras are striped." Certainly the fact that elephants have four legs would have little diagnotic value because zebras also have four legs. If mental representations are the minimum necessary to perform a task, then one would not expect them to include information about number of legs. If, on the other hand, the mental representations were just representations of examples, then information about legs would be retained. Our claim is that when abstraction takes place the resulting mental representation is more like that of examples than like the minimum needed to perform a task (that is, in the case of elephants versus zebras, information about number of legs might well be part of the abstract representation). The other aspect of conservative induction is that inductions need to be tied to data. Technically, a striped zebra supports both the hypothesis "zebras are striped" and the hypothesis "zebras are striped or bright orange." Conservative induction would disallow the latter hypothesis unless some specific bright orange zebra had been observed.

Some Evidence

Prototype Versus Exemplar Models of Categorization

Earlier we mentioned that exemplar models of categorization could account for the main findings that had been taken as support for prototype models and that there was at least a bit of evidence favoring exemplar models. Actually there have been numerous contrasts of prototype versus exemplar models, and these tests have tended to favor exemplar models uniformly. Recall that a prototype representation only preserves information about mean or modal values and all other information is ignored. This is not conservative induction! We believe that the lack of conservative induction in prototype models is the main reason exemplar models fare better than prototype models.

Comparisons of exemplar and prototype models have generally taken the following form:

1. Prototype models assume that some form of potential category-relevant information is not preserved, whereas the contrasting exemplar model implies that the information is not discarded.
2. A test is made that finds that people are, in fact, sensitive to this information, contrary to the prediction of prototype models.

For example, representing a category in terms of central tendency implies that information concerning category size and the frequency with which individual examples appeared is thrown away. Categorization is, however, influenced by both category size and exemplar frequency (e.g., Estes, 1986b; Nosofsky, in press) In addition, people are influenced by correlated properties within categories. For example, large birds are less likely to sing than small birds, and as a consequence, a large nonsinging bird may be more typical than a large, singing bird of the category *bird*. Again, prototype models imply insensitivity to such correlations contrary to data obtained with both artificial categories (Medin, Altom, Edelson, & Freko, 1982) and natural categories (Malt & Smith, 1983; Medin & Shoben, 1988).[2]

Our reading of the contrasts between exemplar versus prototype models is that exemplar models fare better because they are more conservative with respect to induction. The claim that abstraction is carried out over individual, independently processed components appears to be too strong. These results do not unequivocally suppott exemplar models per se.[3] Rather they suggest that cate-

[2]Yet another problem for prototype models is that they imply sensitivity to certain category properties that do not seem to matter to people. For example, it makes no difference whether or not categories can be separated perfectly on the basis of a weighted, additive function of typical properties (Medin & Schwanenflugel, 1981; Wattenmaker, Murphy, Dewey, & Medin, 1986). This is required in order for a prototype process to work successfully. This sensitivity is a consequence of the fact that other category-relevant information is not available, according to prototype theory.

[3]We should note that exemplar storage does not necessarily imply conservative induction, at least if induction is defined over combinations of representation and processing assumptions, not over representations alone. Preserving information by storing examples does not provide conservative induction if the processes that operate on the exemplar representations can only extract a tiny subset of the potential information. One exemplar-based model of classification, known as an average distance model, assumes that classifications are based on computing the average similarity (or its converse, distance) to the examples of each category and classifying the new instance to the category with the greatest average similarity. For many categorization tasks the average distance model makes predictions identical to those of a prototype model (see Estes, 1986a, for a formal proof).

gory representations are more like and preserve more information associated with specific examples than prototype models imply (Medin, 1986). To be successful, models of abstraction would need to be sensitive to these additional types of information.

The above observations are consistent with reminding-based abstraction, because reminding-based abstraction will be conservative. If categorization and problem solving often involve comparing examples within a category, then attention will be directed at properties that the examples share. For comparisons arising from an example reminding the learner of an earlier example, one would expect that only the most similar instances would be selected. For example, in learning the category *elephant* it might be that bull elephants would tend to remind the learner primarily of other bull elephants. Thus, if there were two distinct types of elephant (bulls and cows), one might expect this to be reflected in the abstract representations that are formed (for example, the representation of elephant might consist of distinct bull elephant and cow elephant representations). Such representations would be consistent with the evidence of sensitivity to within-category property correlations that we mentioned earlier. In brief, within-category property correlations that we mentioned earlier. In brief, within-category comparisons should tend to promote conservative induction.

Cue Versus Category Validity

One can also make the case that conservative induction is very adaptive. The idea that mental representations ought to be the minimum necessary to produce successful classifications treats categorization as an end in itself. One of the main functions of categorization is to allow one to make predictions and inferences about the future. In this view, categorization provides access to critical information, namely, the set of properties that may be inferred from access to a representation. If the representation of elephants is that "they are big," then little useful information will be accessed. Alternatively, one's accessing the information that bull elephants may suddenly charge at jeeps could be vitally important. More formally, we are suggesting that category validity (the probability of some feature given a category) may be at least as important as cue validity (the probability of a category given some feature). Category validity may influence performance even in tasks that, nominally, involve only cue validity.

There is suggestive evidence that category validity influences categorization. Wolf, Gruppen, and Billi (1985) studied the performance of medical personnel in a simple symptom-testing task. The medical personnel were presented with cards labeled with two diseases (A and B) and given information about the prevalence of one of the symptoms in one of the disease categories. Participants were allowed to select one of the other three sources of information (prevalence of that same symptom in the other disease, prevalence of the alternative symptom in that or the other disease). To determine cue validity or diagnostic value, one would

more information associated with
(Medin, 1986). To be successful,
itive to these additional types of

ith reminding-based abstraction,
onservative. If categorization and
:amples within a category, then
examples share. For comparisons
of an earlier example, one would
ould be selected. For example, in
hat bull elephants would tend to
ephants. Thus, if there were two
e might expect this to be reflected
for example, the representation of
unt and cow elephant representa-
nt with the evidence of sensitivity
we mentioned earlier. In brief,
nentioned earlier. In brief, within-
: conservative induction.

ve induction is very adaptive. The
e minimum necessary to produce
is an end in itself. One of the main
make predictions and inferences
rovides access to critical informa-
y be inferred from access to a
ts is that "they are big," then little
vely, one's accessing the informa-
t jeeps could be vitally important.
y validity (the probability of some
is important as cue validity (the
. Category validity may influence
nvolve only cue validity.
validity influences categorization.
rformance of medical personnel in
ersonnel were presented with cards
information about the prevalence
ase categories. Participants were
of information (prevalence of that
of the alternative symptom in that
ty or diagnostic value, one would

need to test for the prevalence of the given symptom in the alternative disease category. Only a minority of the personnel (24%) consistently selected this optimal information, and the large majority of nonoptimal choices were testing for the alternative symptom in the initial disease category. In general, if physicians organize their medical knowledge in terms of diseases and the likelihood that different symptoms are associated with them, then category validity may play an important role in induction and diagnostic reasoning. In fact, there is evidence in clinical medicine (Eddy, 1982) that physicians often act as if category validity (probability of a symptom given a disease) were the same as cue validity (probability of a disease given a symptom).

Category Validity and Redundant Rules

A simple illustration of the influence of category validity is associated with the trains shown in Fig. 6.3. The task is to come up with a rule that can be used to determine whether a train is eastbound or westbound. The reader may wish to look at the trains in Fig. 6.3 and come up with such a rule.

There are many possible rules that would successfully classify the trains. The

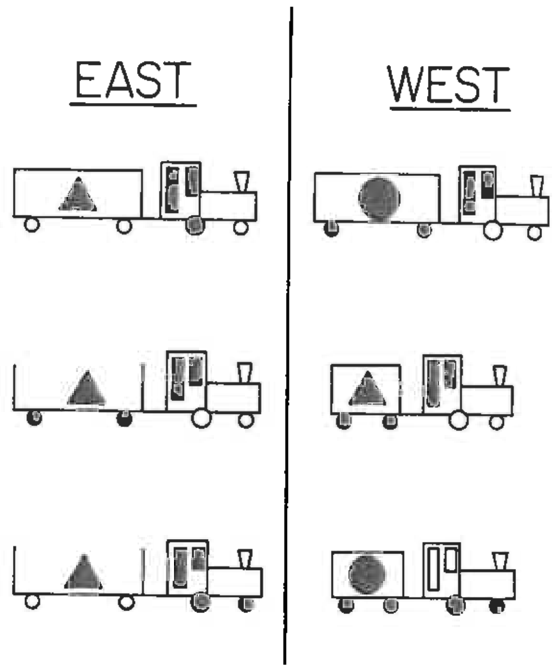

FIG. 6.3. Set of trains from the Medin et al. (1987) rule induction studies. The Participant's task is to come up with a rule that could be used to classify trains as Eastbound versus Westbound.

question is, Which rules are people most likely to develop? If cue validity were paramount, then one might focus on properties that are unique to one category and combine them into an overall rule. For example, one might notice that only westbound trains have circular loads or short cars, and develop the rule, "Westbound trains have short cars or a circular load."

On the other hand, if category validity is important, then one might expect rules to be composed of properties that are true of all of the members of the category. For example, one might notice that all eastbound cars have a triangular load and all have a long car, and develop the rule, "Eastbound trains have long cars and a triangular load."

These alternative possibilities were evaluated in a series of rule induction experiments by Medin, Wattenmaker, and Michalski (1987). For the trains shown in Fig. 6.3, almost all of the rules developed by participants were conjunctions of properties with perfect category validity.

Even more surprising is the observation that the rules given were not necessarily the most simple possible. For example, a considerable number of people gave the rule, "Eastbound trains have a circular load, a long car, and dark engine windows." Of course, the rule would be equally accurate without the information concerning engine windows. These rules with redundant components provide strong support for conservative induction in that the task seems to represent ideal circumstances (few examples, no memory load) for discovering and using the most simple (minimal) rules. Yet the most simple rules were not always given. In addition, the fact that category validity effects can be obtained in contexts where only cue validity is needed is also completely consistent with conservative induction.

Context Specificity

The standard account of abstraction treats concepts as relatively static and context-independent. The concept *bird*, for example, is thought to be instantiated with the prototype of the best example of bird, independent of the prevailing context. Roth and Shoben (1983), however, demonstrated that typicality judgments vary substantially with context. For example, tea is a more typical beverage than milk in the context of librarians taking a break, but this ordering reverses in the context of truck drivers taking a break. When people read the sentence "The birds waded into the pond," they think of ducks, not robins. If one goal of categorization is to make predictions, it makes a great deal of sense for representations to be context-specific (see Barsalou & Medin, 1986, for more extensive arguments). Again, we have conservative induction.

Interproperty Relations

Independent feature theories, such as prototype abstraction models, do not embody conservative induction in that they assume that abstractions are per-formed over individual dimensions independent of other dimensions. One might expect, however, that interproperty relationships are important if representations are knowledge-rich and context-dependent. For example, *wings, hollow bones,* and *feathers* are related (by the concept of flying) in a way that *red, large,* and *triangles* are not. Medin, Wattenmaker, and Hampson (1987) reported evidence from a task where people are given examples and asked to construct categories that interproperty relationships rather than independent features serve to organize categories. They observed sorting by the principle of "family resemblance" only when knowledge was provided that linked the properties together conceptually. These observations are consistent with induction being more conservative than independent-feature models imply.

Content-Specific Problem Schemata

The common view of problem schemata is that they contain abstract knowledge about how to identify a problem of that type and abstract knowledge about how to solve such identified problems. Although little work has addressed the exact nature of these schemata, it appears that under some circumstances they may contain a great deal of specific and formally unnecessary information. The Hinsley et al. (1977) work, which has been often mentioned in this chapter, provides the best illustration. As a reminder, they found that if word problems used contents that were often used for a particular problem type, the identification of the problem type and the solution of the problem were different from when the content was atypical for the problem type. From this research it appears that if irrelevant aspects are often used in a similar way for a particular problem type, a problem schema will be constructed that takes advantage of these correlations for identification and solution. Such an idea follows clearly from the view presented here but is problematic for most views of problem schemata.

RELATED WORK

We have been arguing that specific or exemplar-based information plays a significant role in intellectual functioning and have tried to make the case for this perspective by reviewing psychological research on categorization and problem solving. One can also trace a trend toward a greater recognition of the "priority of the specific," nonautonomous abstraction, and conservative induction in related areas of research. We provide a brief description of a few of these projects so that the reader may see alternative instantiations of this general approach.

Artificial Intelligence

As we mentioned earlier, most of the initial work in AI was concerned with abstract, domain-independent reasoning heuristics. This early work was followed by a trend toward using domain-specific schemata, frames, or scripts in reasoning. These organized packets of information contained background knowledge that would be important for an intelligent system to know but, at the same time, were quite rigid. For example, one might have a single script for what goes on in a restaurant. But then how does one decide when a variation is a "normal" variation versus one that requires subscripts? A script that would need to cover both a posh French restaurant and a fast-food hamburger haunt would be quite strained. One response to the lack of flexibility of scripts has been to build systems that can exploit single precedents that may be tailored to fit current situations more closely. We now briefly review a few systems that employ case-based reasoning.

Dynamic Memory and Case-Based Reasoning

Schank (1982) presented a model of memory in which much of the processing of new events is affected by remindings of earlier events (see also Kolodner 1983a, 1983b, 1984). When one is reminded of an earlier event, a generalization of the commonalities of the two events is made, and the events are stored under this generalization with their separate distinctive features. Thus, generalizations of two instances occur, but only through remindings. (Generalizations may also be made if an already formed generalization is used to process a new event, in the same manner as remindings lead to generalizations.) This model leads to generalizations that have proven useful in previous cases, and these generalizations will tend to be more low-level or conservative than those associated with scripts.

In associated work, Kolodner (1983a, 1983b, 1984) has addressed the issue of how memory organization is used in problem solving, focusing on case-based reasoning. Even if a person has the well-organized knowledge, there will be times when abstract knowledge will be insufficient. As in Schank's model, in Kolodner's approach abstract knowledge is built up from comparing episodes and noting their commonalities. The episodes are then stored beneath this abstract knowledge, with indices as to their distinctive features. If a person is trying to solve a problem with which they have a great deal of practice, they are likely to have relevant abstract knowledge that can be used to solve the problem. However, if this problem does not fit the abstract schema very well, the system looks for particular earlier experiences that may fit better.

Kolodner's work on case-based reasoning examined instances in which the person uses a previous experience to suggest means of solving the new problem. She pointed out that the earlier experience may be used in a variety of ways. First, it may help in understanding the focus and intricacies of the new problem. Second, it may help in generating a plan for solving the new problem. Third, if the current problem solving results in a failure, the earlier experience may help in explaining or remedying the failure.

These ideas may be illustrated using one of the domains Kolodner has used,

:ic schemata, frames, or scripts in
ation contained background knowl-
it system to know but, at the same
it have a single script for what goes
ide when a variation is a "normal"
A script that would need to cover
l hamburger haunt would be quite
ility of scripts has been to build
hat may be tailored to fit current
view a few systems that employ

easoning

ry in which much of the processing
earlier events (see also Kolodner
of an earlier event, a generalization
le, and the events are stored under
ive features. Thus, generalizations
indings. (Generalizations may also
used to process a new event, in the
ations.) This model leads to gener-
is cases, and these generalizations
than those associated with scripts.
3b, 1984) has addressed the issue of
n solving, focusing on case-based
ganized knowledge, there will be
fficient. As in Schank's model, in
built up from comparing episodes
des are then stored beneath this
distinctive features. If a person is
ve a great deal of practice, they are
it can be used to solve the problem.
tract schema very well, the system
may fit better.
; examined instances in which the
means of solving the new problem.
nay be used in a variety of ways.
and intricacies of the new problem.
solving the new problem. Third, if
:, the earlier experience may help in

of the domains Kolodner has used,

psychiatric diagnoses (e.g., Kolodner & Kolodner, 1982). If a psychiatrist is trying to diagnose a patient, it is not uncommon to be reminded of another patient. This reminding may alert the psychiatrist to examine in more detail some aspects of the patient's complaints or medical history. It may also suggest a treatment plan that worked well with the patient of whom the psychiatrist is reminded. Finally, should a treatment plan not work, that failure may remind the psychiatrist of another patient for whom the same treatment failed and suggest possible routes for alternative treatment.

Kolodner's system shares with Schank's the properties of nonautonomous abstraction and conservative induction. Tendencies to overgeneralize or miss nuances at the schema-level are counteracted by the saving of precedents that are exceptions that can enter into lower-level, more precise generalizations.

Abstractionless Induction

A very different use of earlier examples was proposed by Stanfill and Waltz (1986). These researchers work on the Connection Machine, a computer that makes use of a parallel architecture with a large number of simple processing units operating simultaneously. They hypothesized that reasoning from memory of specific instances is the foundation of an intelligent system, and they, therefore, explicitly eschewed the abstraction view of intelligence. Their research attempts to demonstrate how an abstractionless system can make excellent decisions.

The method they used can be simply explained. For any application, they require a fairly rich data base, but the data base does not have to be edited to include only useful information. Again using an example that they have used in their research, the data base may consist of patient records, with each having personal history, symptoms, disease diagnosed, and treatment. When the system is given information about a new patient (e.g., personal history and symptoms), it uses this data base to arrive at a diagnosis in the following way. First, it gets counts of various feature set frequencies, such as the number of times that a symptom and disease occur together. Second, it produces some metric, such as the conditional probability of a disease given a particular symptom. Third, it computes some similarity measure between the current case and each of the earlier cases, using the metric from step two. Finally, it retrieves those cases that best match the current case and combines their information to arrive at a diagnosis (e.g., it could retrieve the best matching case and just use that diagnosis).

This approach differs radically from many other approaches in artificial intelligence. Unlike the case-based reasoning approach, no strong domain model is required. In addition, unlike the rule induction work, the reasoning relies solely on the data base and does not make use of rules derived from comparing instances. Similar to the exemplar approach to categorization, the hypothesis is that the exemplars are likely to contain the important structural information implicitly.

This perspective has a number of advantages over abstraction-oriented views, two of which are particularly relevant here. First, because no rules are formed, the system has a much easier time correcting inappropriate inferences. That is, while rule induction systems need to deal with ways of avoiding inappropriate rules that were formed earlier (e.g., see Anderson et al., 1979; Michalski, 1983), this model does not. As more cases are added to the data base, any inappropriate inferences that occurred because of coincidental correlations of features will be overcome. Second, this view leads in a simple way to a context-sensitivity about what is important. Rule-based views need to incorporate all qualifications into their conditions to avoid making inappropriate inferences as a function of context. The Stanfill and Waltz model, because it always refers to the specific exemplars, naturally incorporates context-sensitivity. For instance, what to do about high blood pressure depends upon, among other things, whether the patient is pregnant or not.

The Stanfill and Waltz model probably does not have much validity as a psychological model because it requires very considerable computation and because it assumes perfect access to relevant knowledge. Nonetheless, at a more abstract level of analysis, their model illustrates all three of our major themes. Indeed, it represents the most radical form of them. Of course, one does not have to subscribe to the strong claim that no abstraction occurs to recognize the costs of nonconservative induction.

Reasoning and Judgment

Given that one of the most important items on the agenda of intelligent organisms is predicting what will happen next, it is not surprising that models of reasoning and judgment have been geared toward describing how our prior experiences are summarized to generate expectations about the future. The past is a norm for the future. A standard view of judgment is that prior experience with examples leads to norms or standards that are used to generate expectations and evaluate current experiences. Recently this view has been challenged. Kahneman and Miller (1986) took the distinct position that norms are not prestored but rather are frequently created on the spot. The idea is that a current experience retrieves or "recruits" related earlier experiences that form an ad hoc norm. As applied to categorization, norm theory implies that categories are really not stored at all. Rather, the category is determined post hoc at the time it is needed and is heavily influenced by the current example.

In more detail, Kahneman and Miller (1986) assumed that when category norms are needed (e.g., to categorize an item or determine how typical it is), the current instance is used to retrieve some similar examples, these examples are combined to generate some norm, and the current instance is then compared to these generated norms (see also, Barsalou, 1983; Barsalou & Medin, 1986, for related ideas). For example, suppose that John almost always drives home from work on a particular route but decides one day to take a more scenic route. Unfortunately, on that day he becomes involved in a car accident. According to Kahneman and Miller (1986), John ought to retrieve episodes involving typical trips home, and since taking the scenic route represents a departure from a norm, tend to attribute at least some of the blame to his taking the alternative route even though there may be no basis at all for doing so (that is, one route may not be any more dangerous than the other).

Norm theory differs from both the common prototype view and the complete exemplar view espoused by Stanfill and Waltz. First, unlike the prototype view, no abstract knowledge of the category is used to form the norms that are used. Second, unlike the Stanfill and Waltz proposal, not all earlier exemplars are used. In fact, norm theory relies upon the fact that the exemplars that are used are ones retrieved by their similarity to the current instance to account for a variety of effects. This norm theory also differs from some other exemplar models (e.g., Medin & Schaffer, 1978) at least in focus, by highlighting the idea that exemplars are combined to form an ad hoc norm. Of course, norm theories share with exemplar theories the general idea of case-based reasoning.

Learning

The same trends we have been discussing are evident in the area of learning, broadly construed.

ACT

Anderson, working within his ACT framework, has been investigating learning for the last decade, using a variety of specific models of abstraction for procedural knowledge. Anderson represents procedural knowledge in terms of If-Then rules called productions. These productions are initially formed from particular experiences, so abstraction requires mechanisms for generalizing the conditions (i.e., aspects in the If part) and the actions (i.e., aspects in the Then part). In his early work, he proposed that there exist automatic generalization and discrimination mechanisms that operate on these productions (e.g., Anderson et al., 1979).

In addition to these inductive learning mechanisms, Anderson proposed knowledge compilation processes that operate on the traces of the applications of productions to create new productions. These knowledge compilation processes (proceduralization and composition) essentially act to collapse operations that have hitherto been conducted separately (see Anderson, 1983, 1986). Thus, proceduralization collapses information retrieval and production matching so that frequently used information can be retrieved and applied in a single step.

Composition collapses productions that often occur in a sequence into a single production. Anderson's recent work (1986) makes the claim that these knowledge compilation processes are sufficient to produce the inductive learning for which he had previously used separate generalization and discrimination processes. For the present purposes, the important point for how this is accomplished is that these knowledge compilation processes are applied to the traces of an analogy from one instance to another. That is, he proposes that the inductive learning occurs as a result not of separate automatic mechanisms, but from the conscious application of knowledge compilation to an analogy. Again, learning is closely tied to how earlier examples are used.

General constraints

Inductive learning in its general form is impossible, because there are simply too many possibilities to be considered (e.g., Gold, 1967; Valiant, 1984). Necessarily, then, people are not completely generalized information processors but rather come prepared to learn some things and not others. Thus, an important part of a general learning theory would be identifying any domain-independent constraints that people use. Berwick (1986) has suggested that one particular type of conservative induction, the Subset Principle, may be one of these domain-independent constraints. Although we shall not describe the Subset Principle in detail, essentially it leads to making the smallest possible generalization.[4] Berwick pointed out that conservative inductions protect the learner from inappropriate (over)generalizations that may be difficult to recover from. The general claim is that it is very risky to discard information, especially in the early stages of learning. We take the observations favoring exemplar models of categorization not as ironclad proof that people do not risk abstractions but rather as suggestive evidence that abstractions are conservative and rich rather than parsimonious and poor.

Parallel Distributed Processing Models

There has been much recent interest in parallel distributed processing, or PDP, models (e.g., Rumelhart, McClelland, & PDP Group 1986). Because these models differ greatly from the symbolic models discussed so far, it is interesting to consider how they fit in this distinction. In these models, knowledge is represented as a pattern of activation over a large number of individual units, with the same units participating in the representation of a great deal of knowledge. Learning occurs by adjustments in the weight of the connections between the individual units. Although these models are very different from the ones discussed so far, they share the property that the generalizations are not formed by explicitly comparing all possible instances but rather occur as a by-product of

[4]Berwick (1986) or the work by Angluin (1978) cited there should be consulted for details.

ur in a sequence into a single
:s the claim that these knowl-
uce the inductive learning for
ation and discrimination pro-
point for how this is accom-
ises are applied to the traces of
he proposes that the inductive
tic mechanisms, but from the
o an analogy. Again, learning

ible, because there are simply
Gold, 1967; Valiant, 1984).
ralized information processors
not others. Thus, an important
ying any domain-independent
suggested that one particular
iciple, may be one of these
hall not describe the Subset
aking the smallest possible
vative inductions protect the
iat may be difficult to recover
iscard information, especially
servations favoring exemplar
people do not risk abstractions
ns are conservative and rich

lel distributed processing, or
P Group 1986). Because these
scussed so far, it is interesting
these models, knowledge is
e number of individual units,
tion of a great deal of knowl-
ht of the connections between
very different from the ones
generalizations are not formed
ather occur as a by-product of

re should be consulted for details.

the memory storage. As new information is encoded into the units, the adjustment of weights will tend to lead to abstractions over the stored knowledge. An important property of PDP models is that they are highly context- and content-sensitive. For example, if there are distinct subtypes within a category, then multiple prototypes corresponding to these subtypes will evolve (Knapp & Anderson, 1984).

Rumelhart, Smolensky, McClelland, and Hinton (1986) applied these ideas to the notion of schemata. Contrary to the earlier espoused view (e.g., Rumelhart, 1975) that schemata are knowledge structures for interpreting events and storing knowledge, these investigators suggest that schemata correspond to stable patterns of activated units. This view allows a high degree of context- and content-sensitivity, as the inputs constrain the effects of each other and the interpretation of the overall configuration. They provide an illustration of how starting with the idea of a room (via the concept "ceiling") and a particular object in a room (e.g., an oven) will activate units corresponding to schema-related (kitchen) objects, with the extent of activation a function of the schema-relatedness. Thus, refrigerator and sink might be strongly activated, but even within these objects the particular features (e.g., type of sink) would be affected by the oven unit activation. Thus, PDP models allow detailed specificity from their abstractions.

The above examples are consistent with our general thesis. The traditional view of abstraction, which we have been arguing against, does not have the flexibility that apparently is needed to capture intellectual functioning.

SUMMARY

Our paper has been organized around three major claims:

1. Reasoning is often based on specific examples rather than abstract principles.
2. Induction is not necessarily autonomous but derives from how examples are used.
3. Induction is conservative.

Although these claims go against the typical view of intelligence as abstract thought, in a sense we are taking a middle-of-the-road position in light of the surprising evidence on the effectiveness of case-based reasoning without any abstraction.

Throughout this chapter, we have provided arguments for the advantages of this specificity in case-based reasoning, nonautonomous abstraction, and conservative induction. These three influences can be traced through the major components of problem solving and categorization. First, one accesses relevant knowledge. Second, one applies that knowledge to the current situation. Third, the knowledge may be affected, either through some major revision or through some minor updating. Specificity is important in each of these aspects.

First, relevant knowledge must be accessed. This component may be further subdivided into which cues provide access to knowledge and the form that this knowledge takes. Implicit in most abstraction views is that it is best to strip away all but the relevant information and to store information of the greatest generality, so that it will have the widest applicability. We have argued that both of these assumptions are probably incorrect. These assumptions might work well if people could always easily analyze any instance or problem to obtain this abstraction and if the irrelevant aspects were independent of the relevant aspects, but neither of these conditions are commonly met. The abstraction view omits formally irrelevant but empirically correlated cues that both novices and experts use (e.g., Ross, 1984, 1987; Hinsley et al., 1977; Medin & Ortony, in press). We see this natural incorporation of such cues as an important advantage for specificity views. These same cues allow this access to have the very important characteristic of context specificity. For any given category or schema, the knowledge that is accessed does not have to always be the same, but can change with various changes in the situation. Finally, access in case-based reasoning allows the learner to take advantage of knowledge from the start, before any abstraction might be formed, using the same processes that will be used when additional relevant knowledge is acquired.

After knowledge is accessed, it must be applied. Again, a common abstraction view is that if the knowledge has been stored at a high level of generality, then it can be instantiated for a large number of different situations. Even granting this assumption, this view still does not take into account any costs of instantiation of high-level variables. We would argue that such costs can be great for three reasons. First, the variables will often be highly interdependent, as well as context-sensitive. The schema must contain information about all these interdependencies and be able to access and apply it appropriately. Second, if certain irrelevant aspects are correlated with the relevant aspects, this view would needlessly force the person to instantiate the variables from scratch each time. If more specific abstractions could also be stored, all of these interdependencies would not have to be constantly kept track of but could be divided such that only ones in the current situation would be used. Finally, it is frequently the case that the same knowledge can be used in a variety of different ways. Premature abstraction would tend to freeze knowledge into a less flexible form.

A third important aspect of performance is updating of knowledge in light of current experiences. Many abstraction models would claim that instances encountered early in learning are automatically used to form an abstraction. Once this abstraction is formed, as long as no major problems are encountered, the later instances only have minor effects in updating this abstraction. Our alternative proposal is that abstractions are not formed automatically but only when

:o the current situation. Third,
ıme major revision or through
ı each of these aspects.
his component may be further
ıwledge and the form that this
's is that it is best to strip away
nation of the greatest general-
We have argued that both of
ssumptions might work well if
:e or problem to obtain this
endent of the relevant aspects,
t. The abstraction view omits
that both novices and experts
'; Medin & Ortony, in press).
s an important advantage for
ess to have the very important
ven category or schema, the
s be the same, but can change
:cess in case-based reasoning
ge from the start, before any
:esses that will be used when

d. Again, a common abstrac-
. at a high level of generality,
of different situations. Even
:ake into account any costs of
ue that such costs can be great
highly interdependent, as well
formation about all these in-
' it appropriately. Second, if
: relevant aspects, this view
e variables from scratch each
be stored, all of these in-
/ kept track of but could be
ı would be used. Finally, it is
: used in a variety of different
knowledge into a less flexible

ating of knowledge in light of
ould claim that instances en-
to form an abstraction. Once
roblems are encountered, the
this abstraction. Our alterna-
automatically but only when

aspects of one instance have proven useful to accomplishing some goal with a later instance. Thus, only generalizations that have proven useful will be formed. In addition, the current view does not necessarily discard the instances once the generalization is formed. This maintenance of instances is particularly important if the information abstracted originally would exclude important information, a phenomenon that one might expect is not uncommon during learning. For example, further experience may lead the learner to realize that additional distinctions are required, triggering a reevaluation of earlier instances. The maintenance of later instances is also important in updating, because if subtle changes occur in a concept or schema over time, these instances allow the learner to be sensitive to that change. Finally, any inappropriate inferences that are discovered can be corrected by the use of the stored instances, a safeguard that is not available to views that discard these instances.

All told, we find these advantages to be strong motivation for the specificity view in categorization, problem solving, and in general, induction. Such a view allows flexibility and a sensitivity in intellectual functioning, properties that are often lacking in models of abstraction.

Perhaps the most surprising result of our survey is the ability of exemplar-based models to parallel the predictions of models based on more abstract representations. This mimicry may hold some clues to the puzzle of why cognitive psychologists find it so natural to equate intelligence with abstract thought. Models of abstraction are a good first-order approximation to what people do. We have seen, however, that the many models of abstraction discard too much information and are rigid and brittle compared with the context-sensitivity and flexibility of human cognition. In that sense, models of abstraction may fail to get at the true character of human intelligence.

ACKNOWLEDGMENTS

This chapter was supported by National Science Foundation grant 84-19576 and National Library of Medicine grant LM 04375 to Douglas Medin, and by National Science Foundation grant IST 83-08670 to Brian Ross. Marie Banich, Gordon Logan, and Larry Barsalou provided helpful comments on earlier drafts of this paper.

REFERENCES

Angluin, D. (1978). Inductive inference of formal languages from positive data. *Information and Control, 45,* 112–135.

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

Anderson, J. R. (1986). Knowledge compilation: The general learning mechanism. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach: Vol. 2* (pp. 289–310). Los Altos, CA: Morgan Kaufman.

Anderson, J. R., Kline, P. J., & Beasley, C. M. (1979). A general learning theory and its application to schema abstraction. In G. H. Bower (Ed.), *The psychology of learning and motivation: Vol. 13* (pp. 227–318). New York: Academic Press.

Anderson, J. R., & Thompson, R. (in press). Use of analogy in a production system architecture. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Barr, A., & Feigenbaum, E. (1981). *The handbook of artificial intelligence, Vol. I, II, and III*. Los Altos, CA: William Kaufman, Inc.

Barresi, J., Robbins, D., & Shain, K. (1975). Role of distinctive features in the abstraction of related concepts. *Journal of Experimental Psychology: Human Learning and Memory, 104*, 360–368.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11*, 211–227.

Barsalou, L. W., & Medin, D. L. (1986). Concepts: Fixed definitions or dynamic context-dependent representations? *Cahiers de Psychologie Cognitive, 6*, 187–202 (invited paper).

Berwick, R. C. (1986). Learning from positive-only examples: The subset principle and three case studies. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach: Vol 2* (pp. 625–645). Los Altos, CA: Morgan Kaufman.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–215). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brooks, L. R. (1987). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 141–174). London/New York: Cambridge University Press.

Burstein, M. H. (1986). Concept formation by incremental analogical reasoning and debugging. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach: Vol. 1* (pp. 351–370). Palo Alto: Tioga.

Carbonell, J. G. (1983). Learning by analogy: Formulating and generalizing plans from past experience. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach: Vol. 1* (pp. 137–161). Palo Alto: Tioga.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–282). New York: Academic Press.

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

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General, 113*, 541–555.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249–267). New York: Cambridge University Press.

Estes, W. K. (1986a). Memory storage and retrieval processes in category learning. *Journal of Experimental Psychology: General, 115*, 155–175.

Estes, W. K. (1986b). Array models for category learning. *Cognitive Psychology, 18*, 500–549.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models*. (pp. 99–129). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gentner, D., & Landers, R. (1985). *Analogical reminding: A good match is hard to find*. Paper presented at the International Conference of Systems, Man & Cybernetics. Tucson, AZ.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277–300.

Gold, E. M. (1967). Language identification in the limit. *Information and Control, 10*, 447–474.

Goldman, D., & Homa, D. (1977). Integrative and metric properties of abstracted information as a function of category discriminability, instance variability, and experience. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 375–385.

Hajek, F. A. (1969). The primacy of the abstract. In A. Koestler & J. R. Smythies (Eds.), *Beyond reduction* (pp. 309–323). New York: Macmillan.

Hinsley, D. A., Hayes, J. R., & Simon, H. A. (1977). From words to equations: Meaning and representation in algebra word problems. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 89–105). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hintzman, D. L., & Ludlam, G. (1980). Differential forgetting of prototypes and old instances: Simulation by an exemplar-based classification model. *Memory & Cognition, 8*, 378–382.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), *The psychology of learning and motivation: Vol. 19* (pp. 59–87). New York: Academic Press.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15*, 332–340.

Homa, D., & Chambliss, D. (1975). The relative contributions of common and distinctive information on the abstraction from ill-defined categories. *Journal of Experimental Psychology: Human Learning and Memory, 104*, 351–359.

Homa, D., & Vosburgh, R. (1976). Category breadth and the abstraction of prototypical information. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 322–330.

Jacoby, L. L., & Brooks, L. R. (1984). Non-analytic cognition: Memory, perception, and concept learning. In G. Bower (Ed.), *The psychology of learning and motivation: Vol. 18* (pp. 1–47). New York: Academic Press.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternative. *Psychological Review, 93*, 136–153.

Kedar-Cabelli, S. (1985). Purpose-directed analogy. *Proceedings of the Seventh Annual Conference of the Cognitive Science Society* (pp. 150–159). Irvine, CA.

Knapp, A. G., & Anderson, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10*, 616–637.

Kolodner, J. L. (1983a). Reconstructive memory: A computer model. *Cognitive Science, 7*, 281–328.

Kolodner, J. L. (1983b). Maintaining organization in a dynamic long-term memory. *Cognitive Science, 7*, 243–280.

Kolodner, J. L. (1984). *Retrieval and organizational structures in conceptual memory: A computer model.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kolodner, J. L., & Kolodner, R. M. (1982). Towards a computer model of psychiatric reasoning. In *Proceedings of the Sixth Annual Conference on Computers in Medicine*, Washington, D.C.

Lewis, M. W., & Anderson, J. R. (1985). Discrimination of operator schemata in problem solving: Learning from examples. *Cognitive Psychology, 17*, 26–65.

Malt, B. C., & Smith, E. E. (1983). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior, 23*, 250–269.

Mayer, R. E. (1982). Memory for algebra story problems. *Journal of Educational Psychology, 74*, 199–216.

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. (1986). Commentary on "Memory Storage and Retrieval Processes in Category Learning." *Journal of Experimental Psychology: General, 115*, 373–381.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 37–50.

Medin, D. L., & Edelson, S. (1988). Problem structure and the use of base rate information from experience. *Journal of Experimental Psychology: General. 117,* 68–85.

Medin, D. L., & Ortony, A. (in press). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*. Cambridge, England: Cambridge University Press.

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

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

Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology, 20,* 158–190.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology, 35,* 113–138.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, concept cohesiveness, and category construction. *Cognitive Psychology, 19,* 242–279.

Medin, D. L., Wattenmaker, W. D., & Michalski, R. S. (1987). Constraints in inductive learning: An experimental study comparing human and machine performance. *Cognitive Science, 11,* 319–359.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology, 32,* 89–115.

Michalski, R. S. (1983). A theory and methodology of inductive learning. *Artificial Intelligence, 20,* 111–161.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nosofsky, R. M. (in press). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory and Cognition.*

Oden, G. C. (1987). Concept, knowledge, and thought. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology, 38,* 203–227.

Polya, G. (1945). *How to solve it.* Princeton, NJ: Princeton University Press.

Posner, M. K., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83,* 304–308.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6,* 855–863.

Reber, A. S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology, 81,* 115–119.

Reber, A. S., Allen, R., & Regan, S. (1985). Syntactical learning and judgment: Still unconscious and still abstract. *Journal of Experimental Psychology: General, 114,* 17–24.

Reber, A. S., & Lewis, S. (1977). Toward a theory of implicit learning: The analysis of the form and structure of a body of tacit knowledge. *Cognition, 5,* 333–361.

Rosch, E., & Mervis, C. G. (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.

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). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15.* Manuscript

Ross, B. H. (in press). Remindings in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogy workshop*. Cambridge, England: Cambridge University Press.

Ross, B. H., Perkins, S. J., & Tenpenny, P. L. (1988). Reminding-based category learning. Manuscript submitted for publication.

Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15,* 346–378.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding* (pp. 211–236). New York: Academic Press.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (1986). *Parallel distributed processing: Vol. 1* (Cambridge, MA: The MIT Press.

Rumelhart, D. E., & Norman, D. A. (1981). Analogical processes in learning. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 335–360). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rumelhart, D.E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Vol. 2* (pp. 7–57). Cambridge, MA: MIT Press.

Schank, R. C. (1982). *Dynamic memory: A theory of learning in people and computers*. Cambridge, England: Cambridge University Press.

Shavlik, J. W., DeJong, G. F., & Ross, B. H. (1987). Acquiring special case schemata in explanation-based learning. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*. Seattle, WA.

Silver, E. A. (1981). Recall of mathematical information: Solving related problems. *Journal for Research in Mathematics Education, 12,* 54–64.

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

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and processes in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214–241.

Stanfill, C., & Waltz, D. (1986). Toward memory-based reasoning. *Communications of the ACM, 29,* 1213–1228.

Strange, W., Keeney, T., Kessel, F. S., & Jenkins, J. J. (1970). Abstraction over time of prototypes from distortions of random dot patterns: A replication. *Journal of Experimental Psychology, 83,* 508–510.

Thorndyke, P. W., & Hayes-Roth, B. (1979). The use of schemata in the acquisition and transfer of knowledge. *Cognitive Psychology, 11,* 82–106.

Valiant, L. G. (1984). A theory of the learnable. *Communications of the ACM, 27,* 1134–1142.

Vokey, J. R., & Brooks, L. R. (in press). Taming the clever unconscious: Analogic and abstractive strategies in artificial grammar learning. *Cognitive Psychology*.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology, 18,* 158–194.

Wickelgren, W. A. (1974). *How to solve problems: Elements of a theory of problems and problem solving*. San Francisco: W. H. Freeman.

Winston, P. H. (1980). Learning and reasoning by analogy. *Communications of the ACM, 23,* 689–703.

Wolf, F. M., Gruppen, L. D., & Billi, J. E. (1985). Differential diagnosis and the competing-hypothesis heuristic. *Journal of the American Medical Association, 253,* 2858–2862.