Exploring Science

The Cognition and Development of Discovery Processes

David Klahr (2000)

In collaboration with Kevin Dunbar, Anne L. Fay, David Penner, and Christian D. Schunn

Foreword by Herbert A. Simon

A Bradford Book

The MIT Press
Cambridge, Massachusetts
London, England

Contents

Chapter 2

Scientific Discovery as Problem Solving

It is understandable, if ironic, that "normal" science fits pretty well the description of expert problem solving, while "revolutionary" science fits the description of problem solving by novices. It is understandable because scientific activity, particularly at the revolutionary end of the continuum, is concerned with the discovery of new truths, not with the application of truths that are already well known. While it may incorporate some expert techniques in the manipulation of instruments and laboratory procedures, it is basically a journey into unmapped terrain. Consequently, it is mainly characterized, as is novice problem solving, by trial-and-error search. The search may be highly selective—the selectivity depending on how much is already known about the domain—but it reaches its goal only after many halts, turnings, and backtrackings.
(Simon, Langley, and Bradshaw, 1981, p. 5)

What does it mean to claim, as do Simon, Langley, and Bradshaw, that scientific discovery is a type of problem solving? The claim is neither controversial nor informative unless we go beyond a generic interpretation of "problem solving" as a synonym for "thinking," for it is unquestionable that scientific discovery involves some sort of thinking. Therefore, this chapter begins with a brief overview of some widely accepted (by cognitive psychologists) definitions and terminology associated with problem solving. This is followed by a description of the theoretical unification of the areas of problem solving and rule induction that was first suggested by Simon and Glen Lea (1974). Finally, I describe an extension of the Simon and Lea theory to the domain of scientific reasoning in which scientific discovery is characterized as a problem-solving process that involves search in two spaces: a space of hypotheses and a space of experiments.

Problem Solving, Search, and Weak Methods

Newell and Simon (1972) define a problem as consisting of an initial state, a goal state, and a set of permissible transformations from one state to another (called "operators") that, when executed in a correct sequence, result in a solution path from the initial state to the goal state, via a series of intermediate states and subgoals. Operators have constraints that must be satisfied before they can be applied. The set of states, operators, goals, and constraints is called a "problem space," and the problem-solving process can be conceptualized as a search for a path that links the initial state to the goal state.

This formal characterization can be applied to an extremely broad range of situations from laboratory puzzles to everyday situations and contexts. We begin with an example of the former type of problem, and then later in this section we provide an example of a naturally occurring problem-solving situation. Consider the well-known Tower of Hanoi puzzle. It consists of a stack of n disks of different size and three pegs (A, B, and C) on which they can be stacked. In the initial state, all the disks are stacked—in order of decreasing size—on peg A. That is, the largest disk is on the bottom of the stack, followed by the second largest, and so on. The goal state is to get them all stacked—in the same size order—on peg C. The only legal move (that is, the "move operator") is to take the top-most disk from the peg it is on and to place it on one of the other two pegs. The constraints on the move operator are (1) that a disk can only be moved if there is nothing on top of it and (2) that a larger disk can never be placed above a smaller disk. Although simply stated, this puzzle can be quite difficult for novices: the length of the most efficient solution is $2^n - 1$ moves. Thus, a three-disk problem can be solved in no fewer than seven moves, and a four-disk problem in no fewer than fifteen. (See Simon, 1975, for an elegant analysis of different strategies for solving this puzzle; see also Klahr and Robinson, 1981.)

Each of the basic components in a problem—initial state, goal state, operators, and constraints—can vary along a continuum from well- to ill-defined. For example, one could have a well-defined initial state and an ill-defined goal state and set of operators (e.g., make something with these materials), or one could have an ill-defined initial state and a well-defined final state (e.g., prove a particular mathematical conjecture). Scientific problems are obviously much less well defined than the Tower of Hanoi. Nevertheless, they can be characterized in the terms used here. Interestingly, the degree of well-definedness depends on the knowledge that is available to the problem-solver. For that reason, much of the training of scientists is aimed at increasing the degree of well-definedness of problems with respect to the scientists working on them.

In all but the most trivial problems, the problem solver is faced with a large set of alternative states and operators, so the search process can be quite demanding. For example, if we represent the problem space as a branching tree of m moves with b branches at each move, then there are b^m moves to consider in the full problem space. As soon as m and b get beyond small values, exhaustive search for alternative states and operators is beyond human capacity.[1] Thus, effective problem solving depends in large part on processes that judiciously constrain search to the exploration of a few branches. These search constraint processes can be placed into two broad categories: *weak methods* and *strong methods*. Weak methods are

> so called because they require relatively little knowledge of the problem structure for their application, but are correspondingly unselective in the way in which they search through the problem space. One mark of expertise is the possession of strong methods specifically adapted to the problem domain, which may allow solutions to be found with little or no search. For example, if someone who knows the calculus is asked to find the maximum of a function, he or she applies a known algorithm (taking the derivative and setting it equal to zero), and finds the answer without search. (Simon, 1986, p. 162)

Scientific practice involves a plethora of strong methods, such as statistical tools, standard experimental paradigms, routine instrumentation processes, and even highly constrained publication formats. In this book we do not address those aspects of scientific discovery other than to acknowledge that they can be viewed, as in the maximization example earlier, as the direct application of a method with little or no search. Of more interest to us are the ways in which weak methods are used in scientific discovery. They are of interest because they are, by definition, heuristic: they may work, or they may not, but they are applicable in a wide variety of contexts. Unlike the strong methods, which are only acquired after extensive formal training, the weak methods are easily acquired (indeed, some may be innate), and they are domain-general. We will describe five major methods: generate and test, hill climbing, means-ends-analysis, planning, and analogy.

Generate and Test

This method is commonly called "trial and error." It consists of simply applying some operator to the current state and then testing to determine if the goal state has been reached. If it has been, the problem is solved. If it has not been, then some other operator is applied. In the most primitive types of generate-and-test methods, the evaluation function is binary: either the goal has been reached or it has not, and the next "move" does not depend on any properties of the discrepancy between the current state and the goal state, or the operator that was just unsuccessfully applied. An example of a "dumb" generating process would be one in which you are searching in a box of keys for a key to fit a lock, and you sample with replacement: tossing failed keys back into the box without noting anything about the degree of fit, the type of key that seems to fit partially, and so on. A slightly "smarter" generator would, at the least, sample from the key box without replacement.

Hill Climbing

This method gets its name from the analogy to someone attempting to reach the top of a hill whose peak cannot be directly perceived (imagine a foggy day with severely limited visibility). One makes a tentative step in each of several directions, and then heads off in the direction that has the steepest gradient. More generally, the method computes an evaluation function whose maximum value corresponds to the goal state. Potential moves are generated and the evaluation function is applied to each potential state. The state that maximizes the increment to the evaluation function is chosen, that move is made, and then the process iterates from the new state. Hill climbing utilizes more information than generate-and-test about the discrepancy between the current state and the goal state. Instead of a simple all-or-none evaluation, it computes a measure of goodness of fit between the two, and uses that information to constrain search in the problem space. For the Tower of Hanoi, the hill-climbing evaluation function might be a simple count of the number of disks currently on the goal peg. The problem with this method is that it would be reluctant to remove any disks from the goal peg, even though that is a necessary part of the solution, as disks are temporarily placed there so that others can be moved.

Means-Ends Analysis

Of all the weak methods, perhaps the best known is means-ends analysis (Dunker, 1945; Newell and Simon, 1972). Means-ends analysis compares the current state and the goal state, and it produces a description of what differences exist. Then it searches for an operator that can reduce those differences. This search is effective only to the extent that the system already has indexed operators in terms of the types of differences they can reduce. It selects an operator that is designed to reduce the most important differences, and it attempts to apply that operator to the current state. However, it may be the case that the operator cannot be immediately applied, because the conditions for its applicability are not met. Means-ends analysis then formulates a subproblem in which the goal is to reduce the difference between the current state and a state in which the desired operator can be applied. Then it recursively attempts to solve the subproblem.

For the Tower of Hanoi, this method might define the most important difference between the initial state and the goal state as getting the largest disk (at the bottom of the initial stack of *n* disks) to the goal peg. But this move cannot be made unless the largest disk has nothing above it, and the goal peg is empty. The only state that fits this situation is one in which a stack of $n - 1$ disks is on the "other" peg, and the biggest disk is on the start peg. Thus, means-ends analysis would establish this configuration as a subgoal and then attempt to achieve that subgoal.

Means-ends analysis is often used in everyday situations. Consider the problem faced recently by an unnamed psychologist in getting from an office at Carnegie Mellon University to a conference room at a Colorado resort in order to present a talk about scientific reasoning. The "difference" was one of distance, and among the set of distance-reduction operators were flying, walking, biking, and so forth. Flying was the operator that would most rapidly reduce distance, but it could not be applied to the initial condition: that is, one could not fly directly from the office to the conference site. This presented the subproblem of creating conditions for flying (i.e., getting to an airport). Getting to the airport could best be done via taxi, but there was no taxi at Carnegie Mellon. The sub-subproblem involved making a phone call to the cab company. But all the university phones were out of order for the day during a transition to a new system: only the pay phones worked. An even deeper subproblem: make a call on a pay phone. But a lack of coins made it impossible to apply that operator (no pun intended). However, a Coke machine was handy, and it accepted dollar bills and gave change. So the problem solver bought and discarded a Coke in order to get on the solution path to transport himself to Colorado.

The hierarchy of goals and subgoals generated by this method is called the "goal tree." At any point in the goal tree, if you asked *how* a goal was going to be achieved, the answer could be found in the subgoals (or, if possible, the immediately applied operators). If you asked *why* a goal was being attempted, the answer would lie in the parent goal for that subgoal. Thus: "Why do I need change?" "To make a phone call." (Or, more obscurely, because it skips a few levels in the goal tree: "To get to the airport.") "How will I get change?" "By using a vending machine."

One would not usually think of a vending machine as an means for distance reduction. However, as illustrated by this example, means-ends analysis can, in the process of proliferating subgoals, take the problem solver along unexpected paths into situations that might, at first blush, seem unlikely to bear any relevance to the initial goal. The nature of these paths can change so much that, in some cases they are more appropriately characterized as distinct problem spaces. Indeed, the determination of the number of distinct problem spaces involved in solving complex problems remains an unresolved issue, to which we will return at the end of this book. Later in this chapter we will characterize scientific discovery in terms of just two spaces, but that will be expanded later.

Planning

Newell and Simon (1972) define planning[2] as another problem-solving method consisting of (1) forming an abstract version of the problem space by omitting certain details of the original set of states and operators, (2) forming the corresponding problem in the abstract problem space, (3) solving the abstracted problem by applying any of the methods listed here (including planning), (4) using the solution of the abstract problem to provide a plan for solving the original problem, (5) translating the plan back into the original problem space and executing it.

For example, if we apply the planning method to the Tower of Hanoi puzzle we might get the following three-step plan: (1) Move the smallest disk to the intervening peg, (2) move a stack of $n - 1$ disks to the goal peg, (3) move the smallest disk to the goal peg. Or for the Colorado problem, we might get (1) taxi to airport; (2) fly to Denver; (3) drive to Breckenridge. Because planning suppresses some of the detail in the original problem space, it is not always possible to implement the plan, for some of the simplifications result in planned solution paths that cannot be effected. For example, moving a stack of $n - 1$ disks is not a legal move, so more problem solving is required to execute that part of the plan. Or, for the Colorado plan, there might be no rental cars at the Denver airport.

Analogy

Analogy involves a mapping between a new target domain (i.e., the current problem) and a previously encountered base domain (i.e., some similar, previously solved, problem). The mappings may vary in their complexity from surface mappings—involving only the simple recognition that the current problem can be solved by a known procedure—to relational mappings, to complex structural mappings (Forbus and Gentner, 1991; Gentner and Jeziorski, 1989; Halford, 1992). Although the potential power of solving a novel problem via analogy to a previously solved problem is obvious, the mapping process itself may fail. Thus, like the other weak methods, analogy is not guaranteed to produce a solution.

Although analogy is not included in Newell and Simon's list of problem-solving methods,[3] in the past twenty-five years, analogy has assumed a central role in theories of problem solving and scientific discovery and its underlying mechanisms have been studied in great detail. Keith Holyoak and Paul Thagard (1995) provide several examples of analogical problem solving in major scientific discoveries, ranging from a first-century analogy between sound and water waves to Turing's mind/computer analogy. Holyoak and Thagard also emphasize the role of analogical thinking in cognitive development. (See Goswami, 1996, for an extensive review.) Analogical reasoning also plays a central role in recent analysis of the thinking processes of contemporary scientists working in their labs (Dunbar, 1997; Thagard, 1997; Ueda, 1997), and, as we will argue in later chapters, it is often the method of choice for formulating initial hypotheses and experiments in a variety of discovery contexts.

Analogical mappings thus provide the principal bridge between weak and strong methods when the source of the analogy is a well-defined procedure. Used in conjunction with domain-specific knowledge, analogy may enable the search process to be greatly abridged when patterns are noticed in the current problem state. Prestored knowledge can be evoked and used to plan the next steps toward solution of the problem, provide macros to replace whole segments of step-by-step search, or even suggest an immediate problem solution. The recognition mechanism (with its associated store of knowledge) is a key weapon in the arsenal of experts and a principal factor in distinguishing their performance in the domain of expertise from that of novices.

Scientific Reasoning: Problem Solving or Concept Formation?

The problem-solving view of scientific discovery has its roots in the Gestalt tradition. For example, Wertheimer (1945) implicates search processes in his historical anecdotes about Einstein and Gauss, and Bartlett (1958) is quite explicit in structuring his discussion of the "thinking of the experimental scientist" in terms of search through a set of knowledge states.[4] Simon (1977) elaborates this position in characterizing scientific reasoning as a search process, similar in structure to any problem-solving task, albeit within a complex search space. Simon's contribution to the discovery-as-problem-solving view is to demonstrate how one could go beyond the search metaphor by explaining the discovery process in terms of an explicit theory of human problem solving (Newell, Shaw, and Simon, 1958). This basic idea has since been extended substantially by Simon and his colleagues in the computational models described in chapter 1.

There is another characterization of the process of scientific reasoning that was considered in the initial period when cognitive psychologists first began to study the topic. This view, exemplified by the Bruner and Wason tasks described in chapter 1, we call the *concept-formation view.* The argument here is that much of scientific reasoning consists of forming new concepts—via induction—on the basis of experimental evidence. This view tended to dominate the early laboratory investigations of the scientific discovery process.

Although the concept-formation and problem-solving views appear to be radically different characterizations of the scientific reasoning process, both traditions can be organized into a coherent theory of scientific reasoning. The key to this integration comes from Simon and Lea's (1974) insight that both concept learning and problem solving are information-gathering tasks and that both employ guided search processes. Simon and Lea have shown how a single information-processing system—called the Generalized Rule Inducer—can account for performance in problem-solving tasks and a range of rule-induction tasks, including concept attainment, sequence extrapolation, and grammar induction. The Generalized Rule Inducer uses the same general methods for both problem-solving tasks and rule-induction tasks. The main difference between problem solving and rule induction lies in the problem spaces that are used in the two tasks.

Consider a rule induction task such as the 2-4-6 task (Wason, 1960). In this task, the experimenter tells subjects that their goal is to figure out what rule the experimenter has in mind, and he tells them that 2-4-6 is an instance of that rule.[5] Subjects then propose a series of triads—such as 8-10-12 or 3-4-5—one at a time and, for each triad, the experimenter says whether or not it is an instance of the to-be-discovered rule. Whenever subjects are ready, they also can state what they think the rule is. (In the original form of this experiment, the rule was simply "increasing numbers," so that both of the triads proposed above would be positive instances of the rule.) Subjects typically start out with much more limited hypotheses about the rule, such as "increasing evens" or "the third is the sum of the first two."

This kind of task requires search in two problem spaces: a space of rules and a space of instances. That is, subjects have to search for hypotheses about what the rule underlying all the triples is, and they also have to search for a good instance—a particular triple—to test their hypotheses. Problem-solving search, however, takes place in a single space: a space of rules.

Thus, the distinctive feature of rule-induction tasks is that proposed rules (hypotheses) are never tested directly, but only by applying them to instances and then testing whether the application gives the correct result. In rule-induction tasks the participant selects (or is shown) an instance and checks to see whether the instance confirms or disconfirms the rule. Instance selection requires search of the instance space, and changing from one rule to another requires search of the rule space. Because rule induction requires two spaces, the tests operate in a different space from the hypothesis (rule) generator. Simon and Lea's analysis illustrates how information from each space may guide search in the other space. For example, information about previously generated rules may influence the generation of instances, and information about the classification of instances may determine the modification of rules.

The Generalized Rule Inducer view makes it possible to characterize some further differences between the previous research on concept formation and problem solving. Because the concept-learning research is concerned with rules derived from well-defined instances, the rule space is usually very simple: it consists of all possible combinations of the values and attributes in the instance space. Even when participants have some control over instance selection, as in the Bruner, Goodnow, and Austin (1956) work, the full set of permissible instances is predetermined. In problem-solving experiments, the structure of the problem space is usually much more complicated. Rather than being merely the concatenation of a new set of given features, it consists of a series of knowledge states that the participant can generate by following a wide variety of strategies.

The Goal Structure of the Scientific Discovery Process

The Generalized Rule Inducer was designed to account for the results of traditional laboratory studies of problem solving and rule induction. Recall that, in the Wason 2-4-6 task, participants have to search both a space of instances and a space of rules that might account for how the experimenter is classifying number triples. Analogously, in scientific discovery, one has to search both a space of observations or experiments, and a space of hypotheses that can account for them—"Nature's rules," in effect. In considering the complexity of the mental processes that support this kind of thinking, we applied one of the standard tools of cognitive psychology—task analysis—to the domain of scientific discovery. We formulated a framework that expands the top-level discovery goal into a hierarchically organized goal structure that elucidates the nature of this dual search and shows the relations among the components and subcomponents. In this chapter, we introduce that goal structure and in subsequent chapters we use it to account for people's behavior in a variety of scientific discovery tasks.

We start by summarizing the key features of our model of scientific discovery as dual search (SDDS). There are two parts to the presentation. First, we give a very broad overview of the model's main phases and then we give an elaborated description of its goal structure.

Dual Search

Searching the Hypothesis Space The process of generating new hypotheses is a type of problem solving in which the initial state consists of some knowledge about a domain, and the goal state is a hypothesis that can account for some or all of that knowledge in a more concise or universal form. Once generated, hypotheses are evaluated for their initial plausibility. Expertise plays a role here, as participants' familiarity with a domain tends to give them strong biases about what is plausible in the domain. Plausibility, in turn, affects the order in which hypotheses are evaluated: highly likely hypotheses tend to be tested before unlikely hypotheses (Klayman and Ha, 1987; Wason, 1968). Furthermore, participants may adopt different experimental strategies for evaluating plausible and implausible hypotheses.

Searching the Experiment Space One of the most important constraints on this search is the need to produce experiments that will yield interpretable outcomes. This requires *domain-general* knowledge about one's own information-processing limitations, as well as *domain-specific* knowledge about the pragmatic constraints of the particular discovery context. As we will see, there are important developmental differences in people's ability to constrain search in the experiment space.

Evaluating Evidence Search in the two spaces is mediated by the evidence evaluation process. This process both assesses the fit between theory and evidence and guides further search in both the hypothesis space and the experiment space. In contrast to the binary feedback provided to participants in the typical psychology experiment, real-world evidence evaluation is not very straightforward. Relevant features must first be extracted, potential noise must be suppressed or corrected (Gorman, 1992; Penner and Klahr, 1996), and the resulting internal representation must be compared with earlier predictions. When people are reasoning about real world contexts, their prior knowledge imposes strong theoretical biases (Brewer and Chin, 1994). These biases influence not only the initial strength with which hypotheses are held—and hence the amount of disconfirming evidence necessary to refute them—but also the features in the evidence that will be attended to and encoded.

The Goal Structure of SDDS

This model is intended to depict the goal structure that is generated in a broad range of scientific reasoning contexts. The fundamental assumption is that scientific discovery involves search in problem spaces. In the case of scientific discovery, there are *two* primary spaces to be searched: a space of hypotheses and a space of experiments. Search in the hypothesis space is guided both by prior knowledge and by experimental results. Search in the experiment space may be guided by the current hypothesis, and it may be used to generate information to formulate hypotheses.

Before presenting the model, we need to briefly describe the format we use to depict the mental representation of hypotheses. We use Marvin Minsky's "frame" notation: data structures that have a set of elements with slots whose values can vary according to context, and whose set of elements bear fixed relations to one another. Initial hypotheses are constructed by a series of operations that result in the instantiation of a frame with default values. Subsequent hypotheses within that frame are generated by changes in values of particular slots and new frames are generated either by a search of memory or by generalizing from experimental outcomes.

SDDS consists of a set of basic *components* that guide search within and between these two problem spaces. In the following description of SDDS we sometimes refer to the components as goals and subgoals, and at other times as processes and subprocesses, depending on the context. Because we are proposing SDDS as a general framework within which to interpret human behavior in any scientific reasoning task, we introduce it at a very general level without reference to any particular discovery domain.

The full model is complex, so perhaps an advance organizer will help. The target we are aiming for in the next several pages is the fully-elaborated structure depicted in figure 2.9 (see page 37). Our approach is to describe various subcomponents of the full model in a piecemeal fashion—from left to right, and from top to bottom. Then, at the end of the chapter, we put it all together. So let us begin.

Three main components control the entire process from the initial formulation of hypotheses, through their experimental evaluation, to the decision that there is sufficient evidence to accept an hypothesis. The three components, shown in figure 2.1, are *Search Hypothesis Space, Test Hypothesis*, and *Evaluate Evidence*. The output from *Search Hypothesis Space* is a fully specified hypothesis, which provides the input to *Test Hypothesis*. The output of *Test Hypothesis* is a description of evidence for or against the current hypothesis, based on the match between the prediction derived from the current hypothesis and the actual experimental

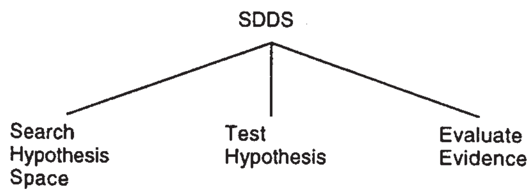

Figure 2.1
The three top-level components of SDDS.

result. Next, *Evaluate Evidence* decides whether the cumulative evidence—as well as other considerations—warrants acceptance, rejection, or continued consideration of the current hypothesis. Each of these three top-level components will be described more fully in the following paragraphs.

Search Hypothesis Space
This process, depicted in figure 2.2, has two components. One component generates the broad scope of a new hypothesis, and the second component refines it and further specifies it. Because we represent hypotheses as frames, we show this as first *Generate Frame* and then *Assign Slot Values*. Where do these initial frames and their associated slot values come from? We propose two different types of sources for new hypotheses. One source comes from prior knowledge stored in memory, and the other source comes from experimental (or observational) data. The two different sources are evoked in both *Generate Frame* and in *Assign Slot Values*.

Generate Frame Depicted in figure 2.3, *Generate Frame* has two components corresponding to the two ways that a frame may be generated. The

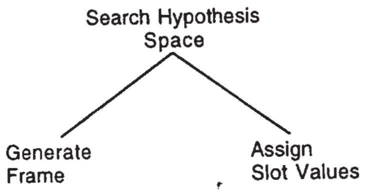

Figure 2.2
Top-level structure of Search Hypothesis Space.

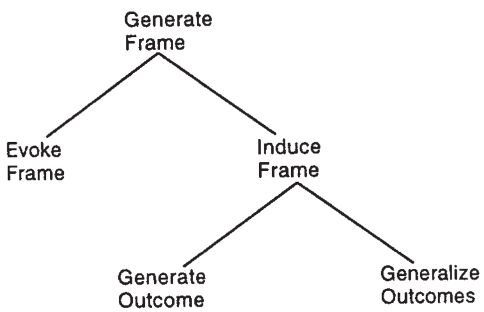

Figure 2.3
The structure of Generate Frame.

first component to be activated, *Evoke Frame*, is a search of memory for information that could be used to construct a frame. Prior knowledge plays an important role here. In cognitive psychology, several mechanisms have been proposed to account for the way in which initial hypotheses are generated, including analogical mapping (Gentner, 1983; Gick and Holyoak, 1983), heuristic search (Kaplan and Simon, 1990; Klahr and Dunbar, 1988), priming (Dunbar and Schunn, 1990; Schunn and Dunbar, 1996), remindings (Ross, 1984), and conceptual combination (Shrager, 1985, 1987). Each of these mechanisms emphasizes a different aspect of the way in which search in the hypothesis space is initiated.

But it is not always possible to evoke a new hypothesis from memory. In such cases, it is necessary to induce a frame from data. Thus, the second component of *Generate Frame* is *Induce Frame*. It generates a new frame by induction from a series of outcomes. The first subprocess in *Induce Frame* generates an outcome, and the second subprocess generalizes over the results of that (and other) outcomes to produce a frame. *Generate Outcome* will be described later.

The result from *Generate Outcome* is a data pattern that is input to *Generalize Outcomes*, which then attempts to generalize over the outcomes in order to produce a frame. The difference between *Evoke Frame* and *Induce Frame* is important. It corresponds to the difference between situations in which participants are able to recall similar situations and use them as the basis for constructing initial frames, and situations in which participants must observe some results before they can venture even an initial hypothesis.

Assign Slot Values The second major component of *Search Hypothesis Space* is shown in figure 2.4. Its purpose is to take a partially instantiated

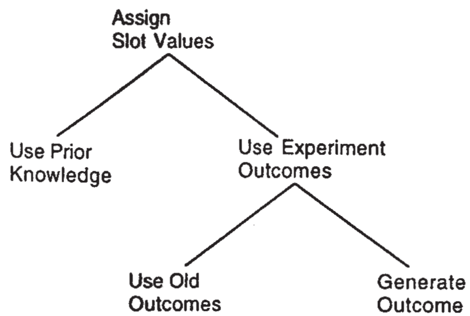

Figure 2.4
The structure of Assign Slot Values.

frame and assign specific values to the slots so that a fully specified hypothesis can be generated. Like *Generate Frame*, it also has two subcomponents corresponding to memory search or induction over data. Slot values may be assigned by using prior knowledge (*Use Prior Knowledge*) or by using specific experimental outcomes (*Use Experimental Outcomes*). As noted, the distinction here is parallel to the earlier distinction between evoking and inducing a frame, but now it is with respect to filling in slot values in an existing frame rather than generating a new frame. If there are already some experimental outcomes, then they can be examined to determine specific slot values (*Use Old Outcomes*). Alternatively, the system can use *Generate Outcome* to produce some empirical results solely for the purpose of determining slot values, that is, for refining a partially specified hypothesis.

In the early phases of the discovery process, *Use Prior Knowledge* plays the major role in assigning values, whereas later in the course of experimentation, *Use Experimental Outcomes* is more likely to generate specific slot values. If the system is unable to assign slot values to the current frame (because they have all been tried and rejected), then the frame is abandoned and the system returns to *Generate Frame* (see figure 2.3).

In figure 2.5, we have assembled all of these components of *Search Hypothesis Space* to show its full structure. The end result of *Search Hypothesis Space* is a fully specified hypothesis which is then input to *Test Hypothesis* (to be described in the next few paragraphs). Note that *Generate Outcome* occurs in two different subcontexts in the service of *Search Hypothesis Space*. This requires the running of "experiments" even though neither of these contexts involve the evaluation of an hypothesis, for it is still being formed. We will return to this point.

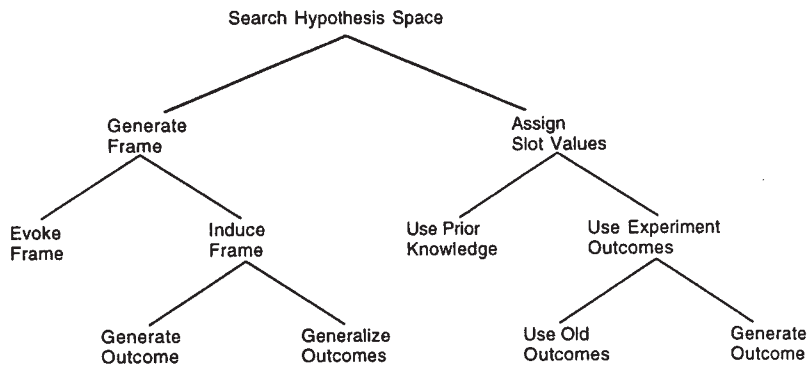

Figure 2.5
The full structure of *Search Hypothesis Space*.

Test Hypothesis
Now we describe the second of the three main components shown in figure 2.1: *Test Hypothesis* (figure 2.6). It generates an experiment (*Search E-Space*) appropriate to the current hypothesis, makes a prediction, runs the experiment, and matches the outcome to the prediction. *Test Hypothesis* uses three subcomponents. The first, *Search E-Space*, produces an experiment. It will be described later, as it is used in several places in the model. The second, *Make Prediction*, takes the current hypothesis and the current experiment and predicts specific results, centered on the current focal values. The third component, *Run*, executes the experiment, and then *Match* produces a description of a discrepancy between the prediction and the actual outcome. As depicted here, the expected outcome is generated prior to the running of the experiment (during *Predict*). However, SDDS allows the computation of what "should have happened" to occur *following* the running of the experiment, during the *Match* process. *Match* requires descriptions of both the expectation and the observation as input. When completed, *Test Hypothesis* outputs a representation of evidence for or against the current hypothesis; this representation is then used as input by *Evaluate Evidence*.

Evaluate Evidence
Evaluate Evidence (figure 2.7) determines whether or not the cumulative evidence about the experiments run under the current hypothesis is

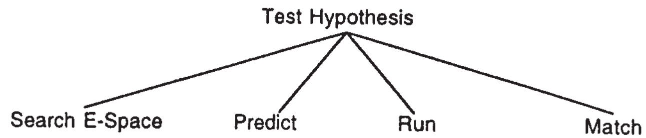

Figure 2.6
Test Hypothesis.

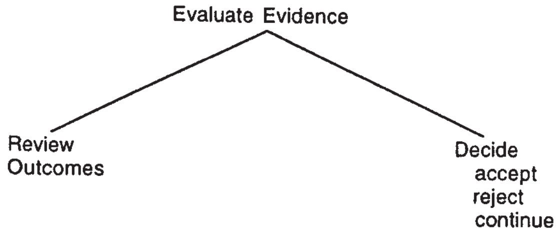

Figure 2.7
Evaluate Evidence.

sufficient to reject or accept it. It is possible that the evidence is inconclusive and neither situation obtains, in which case *Evaluate Evidence* loops back to *Test Hypothesis.* Note that the input to the review process consists of an accumulation of output from earlier *Test Hypothesis* cycles. The scope of this accumulation could range from the most recent result, to the most salient ones, to a full record of all the results thus far. The content of this record could be one of either consistency or inconsistency. Additional factors may play a role in *Evaluate Evidence*, for example, plausibility, functionality, or parsimony. Although these factors appear to influence evidence evaluation, we do not yet have a general understanding of how they work, and we will return to them in subsequent chapters describing the results of specific studies.

Generate Outcome

This process, depicted in figure 2.8, starts with a *Search E-Space,* which produces an experiment. Once produced, the experiment is *Run* and the result is *Observed. Generate Outcome* occurs twice in the subgoal hierarchy of *Search Hypothesis Space:* (1) when outcomes are being generated in order to induce frames and (2) when they are being run to refine hypotheses by assigning slot values (see figure 2.5). As noted, each time that *Generate Outcome* is invoked, so is *Search E-Space,* which we describe next.

Search E-Space Experiments are designed by *Search E-Space.* The most important step is to *Focus* on some aspect of the current situation that the experiment is intended to illuminate. "Current situation" is not just a circumlocution for "current hypothesis," because there may be situations in which there is no current hypothesis, but in which *Search E-Space* must function nevertheless. (This is an important feature of the model, and it will be elaborated in the next section.) If there is an hypothesis, then *Focus* determines that some aspect of it is the primary reason for the experiment. If there is a frame with open slot values, then *Focus* will select any one of those slots as the most important thing to be resolved. If there is neither a frame nor an hypothesis—that is, if *Search E-Space* is being called by *Induce Frame,* then *Focus* makes an arbitrary decision about what aspect of the current situation to focus on. Once a focal aspect has been determined, *Choose* sets a value in the experiment space that will provide information relevant to it, and *Set* determines the values of the remaining, but less important, experimental features necessary to produce a complete experiment.

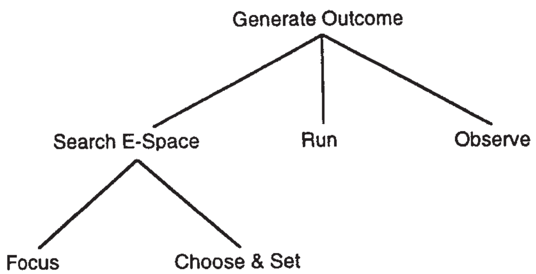

Figure 2.8
Generate Outcome.

The Multiple Roles of Experimentation in SDDS

Examination of the relationship among all these processes and subprocesses, depicted in figure 2.9, reveals both the conventional and unconventional characteristics of the model. At the top level, the discovery process is characterized as a simple repeating cycle of generating hypotheses, testing hypotheses, and evaluating the accumulated evidence. Below that level, however, we can begin to see the complex interaction among the subprocesses. Of particular importance is the way in which *Search E-Space* occurs in three places in the hierarchy: (1) as a subprocess deep within *Generate Frame,* where the goal is to generate a data pattern over which a frame can be induced, (2) as a subprocess of *Assign Slot Values* where the purpose of the "experiment" is simply to resolve the unassigned slots in the current frame, and (3) as a component of *Test Hypothesis,* where the experiment is designed to play its conventional role of generating an instance (usually positive) of the current

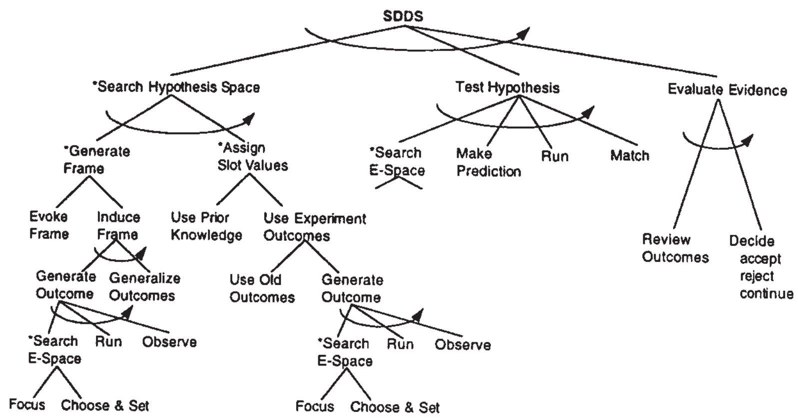

Figure 2.9
Complete SDDS Goal Structure. Components connected by an arrow are usually executed in the order shown. Component names preceded by an asterisk include conditional tests that subcomponents execute.

hypothesis. Note that the implication of the first two uses of *Search E-Space* is that, in the absence of hypotheses, experiments are generated atheoretically, by moving around the experiment space.

SDDS also elaborates the details of what can happen during the *Evaluate Evidence* process. Recall that three general outcomes are possible: The current hypothesis can be accepted, it can be rejected, or it can be considered further. In the first case, the discovery process simply stops, and asserts that the current hypothesis is the true state of nature. In the second case—rejection—the system returns to *Search Hypothesis Space*, where two things can happen. If the entire *frame* has been rejected by *Evaluate Evidence*, then the model must attempt to generate a new frame. If *Evoke Frame* is unable to generate an alternative frame, then the system will wind up in *Induce Frame* and will ultimately start to run experiments (in *Generate Outcome*) in order to find some empirical elements over which to do the induction. Having induced a new frame, or having returned from *Evaluate Evidence* with a frame needing new slot values (i.e., a rejection of the hypothesis but not the frame), SDDS executes *Assign Slot Values*. Here, too, if prior knowledge is inadequate to make slot assignments, the system may wind up making moves in the experiment space in order to make the assignments (i.e., *Generate Outcome* under *Use Experimental Outcomes*). In both these cases, the behavior we would observe would be the running of "experiments" without fully specified hypotheses. This is precisely what we see in several of the studies to be described later in this book. It is particularly common in children's experimentation. In the third case, SDDS returns to *Test Hypothesis* in order to further consider the current hypothesis. The experiments run in this context correspond to the conventional view of the role of experimentation. During *Move In E-Space*, *Focus* selects particular aspects of the current hypothesis and designs an experiment to generate information about them.

Conclusion

In this chapter we have elucidated our general claim that scientific discovery is a type of problem solving involving search in problem spaces. We then summarized the important weak methods used to control that search. Next, we argued that scientific discovery requires search in (at least) two problem spaces: a space of hypotheses and a space of experiments. Finally, we described a detailed goal structure consisting of three main components that coordinate the search in two spaces.

Each of these three components is a potential source of developmental change and, as indicated by the taxonomy presented in table 1.1, most psychologists have studied them in isolation. But such decomposition begs the very question of interest: the coordination of search in two spaces. We wanted to try a different approach in which we could study discovery behavior in situations that required coordinated search in *both* the experiment space and the hypothesis space. In order to do this, we set up laboratory situations that were designed to place participants in various parts of this framework and then look at how they managed the dual search process. In the next chapter, we describe our basic paradigm for such investigations.

Notes

1. For example, Newell and Simon (1972) note that in chess "there are something like 10^{120} continuations to be explored, with much less than 10^{20} nanoseconds available in a century to explore them" (p. 669).
2. Planning has had a very wide variety of definitions, ranging from "little computer programs that program the mind to perform certain cognitive tasks, such as long division, brushing your teeth, or generating a sentence" (Wickelgren, 1974, p. 357) to "any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed" (Miller, Galanter, and Pribram, 1960, p. 16) to "the predetermination of a course of action aimed at achieving a goal" (Hayes-Roth and Hayes-Roth, 1979, p. 275). An elaboration and discussion of the many definitions can be found in Scholnick and Friedman (1987). We use the Newell and Simon definition here because (1) it is much better defined than the others, and (2) it fits nicely in the set of weak methods.
3. Instead, it is characterized by them as one of several sources of information that could be used to create a representation for the problem space. "If the given task bears a similarity to another task with which the problem solver is already familiar, then he may attempt to use the problem space and programs from the familiar one. To do so requires that the knowledge and structures be adapted to the new task. The result will be useful only if the initial recognition of similarity reflects fundamental isomorphisms between the structures of the tasks and if the adaption is successful. (Newell and Simon, 1972, p. 851)
4. In a remarkably prescient inquiry in 1932, Erika Fromm, then a young graduate student of Max Wertheimer, asked a sample of famous scientists and philosophers to reflect on their own "productive thinking processes" and to "write a few lines about the beginning and the course of a cognitive process from your concrete scientific research" (Fromm, 1998, p. 1194). Among the replies she received was one from Albert Einstein, who, in reflecting on the thinking that led to his special theory of relativity, said, "The psychological situation is comparable with the attitude of somebody who wants to solve a puzzle or a chess problem, who is convinced that the solution exists, because the creator of the problem possess the solution" (Einstein, 1932, quoted in Fromm, 1998, p. 1198).
5. There is an extensive literature on this task. See Klayman and Ha (1987) for an excellent review and theoretical reformulation.

own strengths and limitations, and a complete account of the discovery process will include findings from all of them. Fortunately, the strengths of one approach tend to offset the weaknesses of another, and some important convergent findings are emerging from the full set (Klahr and Simon, 1999). In this volume I have tried to demonstrate the strengths and contributions of the discovery microworld approach.

This book will have achieved its aims if it convinces the reader that, rather than being an impenetrable mystery, the psychology of scientific discovery can be studied by using the same methods that apply to the study of any other natural or behavioral phenomena. It will have been even more successful if the particular findings reported here are taken as acceptable explanations for certain behaviors associated with the discovery process and if the methods used here are extended to explore further the many remaining questions. For example, one aspect of scientific practice that I have not dealt with at all in this book is collaboration. Obviously, this is an important factor in scientific discovery, as well as many other forms of problem solving, and it has generated an extensive literature in cognitive and developmental psychology (e.g., Azmitia, 1988; Castellan, 1993; Resnick, Levine, and Teasley, 1991). The microworld studies described in this volume have proven to be easily extended to investigations of collaborative discovery (e.g., Teasley, 1995; Okada and Simon, 1997). Given the important role of collaboration in science, I hope that others interested in the problems addressed here will join my colleagues and me in an extended collaboration as we continue to search for ways to advance our knowledge of the discovery process.

Notes

1. Some computational models focus mainly on search in the hypothesis space: for example, the BACON models and variants (Langley et al., 1987), IDS (Nordhausen and Langley, 1993), PHINEAS (Falkenhainer, 1990), COPER (Kokar, 1986), MECHEM (Valdés-Pérez, 1994), HYPGENE (Karp, 1990, 1993), AbE (O'Rorke, Morris, and Schulenburg, 1990), OCCAM (Pizzani, 1990), and ECHO (Thagard, 1988). Other models focus mainly on the process of experiment generation and evaluation: for example, DEED (Rajamoney, 1993) and DIDO (Scott and Markovitch, 1993). A few deal with both processes: for example, KEKADA (Kulkarni and Simon, 1988), STERN (Cheng, 1990), HDD (Reimann, 1990), IE (Shrager, 1985), and LIVE (Shen, 1993).
2. Indeed, some studies (Sloutsky, Rader, and Morris, 1998) demonstrate that children prefer to seek evidence even when it is unnecessary, such as when they are asked to evaluate the truth or falsity of tautologies or contradictions.

References

Amsel, E., and S. Brock. (1996). The development of evidence evaluation skills. *Cognitive Development* 11: 523–550.

Azmitia, M. (1988). Peer interaction and problem solving: when are two heads better than one? *Child Development* 59: 87–96.

Baker, L. M., and K. Dunbar. (1996). Problem spaces in real-world science: what are they and how do scientists search them? In G. W. Cottrell, ed., *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, pp. 21–22. Mahwah, N.J.: Erlbaum.

Bartlett, Frederic C. (1958). *Thinking: An Experimental and Social Study*. New York: Basic Books.

Bazerman, C. (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison: University of Wisconsin Press.

Bijker, W. E., T. P. Hughes, and T. Pinch. (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge, Mass.: MIT Press.

Bloor, D. (1981). *Knowledge and Social Imagery*. London: Routledge and Kegan Paul.

Boden, M. A. (1990). *The Creative Mind: Myths and Mechanisms*. London: Basic Books.

Boden, M. A. (1994). Précis of the creative mind: myths and mechanisms. *Behavioral and Brain Sciences* 17: 519–570.

Bower, G. H., and T. R. Trabasso. (1964). Concept identification. In R. C. Atkinson, ed., *Studies in Mathematical Psychology*, pp. 32–94. Stanford, Calif.: Stanford University Press.

Brewer, W. F., and C. A. Chinn. (1994). The theory-ladeness of data: an experimental demonstration. In A. Ram and K. Eiselt, eds., *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, pp. 61–65. Hillsdale, N.J.: Erlbaum.

Brewer, W. F., and A. Samarapungavan. (1991). Child theories versus scientific theories: differences in reasoning or differences in knowledge? In R. R. Hoffman and D. S. Palermo, eds., *Cognition and the Symbolic Processes: Applied and Ecological Perspectives*, pp. 209–232. Hillsdale, N.J.: Erlbaum.

Brown, A. L., J. D. Bransford, R. A. Ferrara, and J. C. Campione. (1983). Learning, remembering, and understanding. In P. H. Mussen, ed., *Handbook of Child Psychology: Cognitive Development* (4th ed.) Vol. 3, pp. 77–166. New York: Wiley.

Bruner, J. S., J. J. Goodnow, and G. A. Austin. (1956). *A Study of Thinking*. New York: NY Science Editions.

Bullock, M., A. Ziegler, and S. Martin. (1993). Scientific thinking. In F. E. Weinert and W. Schneider, eds., *LOGIC Report 9: Assessment Procedures and Results of Wave 6*, pp. 66–110. Munich: Max Plank Institute for Psychological Research.

Bullock, M., and A. Ziegler. (1999). Scientific reasoning: developmental and individual differences. In F. E. Weinert and W. Schneider, eds., *Individual Development from 3 to 12:*

Findings from the Munich Longitudinal Study, pp. 38–54. New York: Cambridge University Press.

Burns, B. D., and R. Vollmeyer. (1997). A three-space theory of problem solving. In M. G. Shafto and P. Langley, eds., *Proceedings of the Nineteenth Annual Meeting of the Cognitive Science Society*, p. 879. Mahwah, N.J.: Erlbaum.

Callahan, J. D., and S. W. Sorensen. (1992). Using TETRAD II as an automated exploratory tool. *Social Science Computer Review* 10(3): 329–336.

Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge, Mass.: Bradford Books/MIT Press.

Case, R. (1974). Structures and strictures: some functional limitations on the course of cognitive growth. *Cognitive Psychology* 6: 544–573.

Castellan, J. J., ed. (1993). *Current Issues in Individual and Group Decision Making*. Hillsdale, N.J.: Erlbaum.

Chase, W. G., and H. A. Simon. (1973). The mind's eye in chess. In W. G. Chase, ed., *Visual Information Processing*, pp. 215–281. New York: Academic Press.

Chen, Z., and D. Klahr. (1999). All other things being equal: children's acquisition of the control of variables strategy. *Child Development* 70(5): 1098–1120.

Cheng, P. C.-H. (1990). *Modeling Scientific Discovery*. Ph.D. dissertation, The Open University.

Cheng, P. C.-H., and H. A. Simon. (1992). The right representation for discovery: finding the conservation of momentum. In D. Sleeman and P. Edwards, eds., *Machine Learning: Proceedings of the Ninth International Conference* (ML92), pp. 62–77. San Mateo, CA: Morgan Kaufmann.

Chinn, C. A., and W. F. Brewer. (1992). Psychological responses to anomalous data. In J. K. Kruschke, ed., *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, pp. 165–170. Hillsdale, N.J.: Erlbaum.

Chinn, C. A., and W. F. Brewer. (1998). An empirical test of a taxonomy of responses to anomalous data in science. *Journal of Research in Science Teaching* 35(6): 623–654.

Crick, F. (1988). *What Mad Pursuit: A Personal View of Scientific Discovery*. New York: Basic Books.

Darden, L. (1992). Strategies for anomaly resolution. In R. N. Giere, ed., *Cognitive Models of Science*, pp. 251–273. Minneapolis: University of Minnesota Press.

Darden, L. (1997). Recent work in computational scientific discovery. In M. G. Shafto and P. Langley, eds., *Proceedings of the Nineteenth Annual Meeting of the Cognitive Science Society*, pp. 161–166. Mahwah, N.J.: Erlbaum.

DeBoer, G. E. (1991). *A History of Ideas in Science Education: Implications for Practice*. New York: Teachers College Press.

DeLoache, J. S., K. F. Miller, and S. L. Pierroutsakos. (1998). Reasoning and problem solving. In D. Kuhn and R. S. Siegler, eds., *Handbook of Child Psychology* (5th ed.), Vol. 2, *Cognition, Perception, and Language*, pp. 801–850. New York: Wiley.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science* 17: 397–434.

Dunbar, K. (1995). How scientists really reason: scientific reasoning in real-world laboratories. In R. J. Sternberg and J. Davidson, eds., *The Nature of Insight*, pp. 365–395. Cambridge, Mass.: MIT Press.

Dunbar, K. (1997). How scientists think: on-line creativity and conceptual change in science. In T. Ward, S. Smith, and S. Vaid, eds., *Conceptual Structures and Processes: Emergence, Discovery, and Change*, pp. 461–492. Washington, D.C.: APA Press.

Dunbar, K., and L. A. Baker. (1994). Goals, analogy, and the social constraints of scientific discovery. *Behavioral and Brain Sciences* 17: 538–539.

Dunbar, K., and D. Klahr. (1989). Developmental differences in scientific discovery strategies. In D. Klahr and K. Kotovsky, eds., *Complete Information Processing: The Impact of Herbert A. Simon*. Hillsdale, N.J.: Erlbaum.

Dunbar, K., and C. D. Schunn. (1990). The temporal nature of scientific discovery: the roles of priming and analogy. In M. Piattelli-Palmarini, ed., *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, pp. 93–100. Hillsdale, N.J.: Erlbaum.

Duncan, S. C., and R. D. Tweney. (1997). The problem-behavior map as cognitive-historical analysis: the example of Michael Faraday. In M. G. Shafto and P. Langley, eds., *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, p. 901. Hillsdale, N.J.: Erlbaum.

Dunker, K. (1945). On problem solving. *Psychological Monographs* 58(5): 270.

Einhorn, H. J., and R. M. Hogarth. (1986). Judging probable cause. *Psychological Bulletin* 99: 3–19.

Einstein, A. (1936). Physics and reality. Reprinted in A. Einstein (1950), *Out of My Later Years*. New York: Philosophical Library.

Elder, J. (1990). Hands on. *Education Life*, supplement to the *New York Times*, January 7.

Erdos, P., S. Fajtlowicz, and W. Staton. (1991). Degree sequences in the triangle-free graphs. *Discrete Mathematics* 92(1): 85–88.

Ericsson, K. A., and H. A. Simon. (1993). *Protocol Analysis: Verbal Reports as Data*. Cambridge, Mass.: MIT Press.

Falkenhainer, B. C. (1990). A unified approach to explanation and theory formation. In J. Shrager and P. Langley, eds., *Computational Models of Scientific Discovery and Theory Formation*, pp. 157–196. San Mateo, Calif.: Morgan Kaufmann.

Fay, A. L., and D. Klahr. (1996). Knowing about guessing and guessing about knowing: preschoolers' understanding of indeterminacy. *Child Development* 67: 689–716.

Feist, G. J. (1991). *The Psychology of Science: Personality, Cognitive, Motivational and Working Styles of Eminent and Less Eminent Scientists*. Ph.D. dissertation, University of California at Berkeley.

Feist, G. J. (1994). Personality and working style predictors of integrative complexity: a study of scientists' thinking about research and teaching. *Journal of Personality and Social Psychology* 67: 474–484.

Feist, G. J., and M. E. Gorman. (1998). The psychology of science: review and integration of a nascent discipline. *Review of General Psychology* 2(1): 3–47.

Flavell, J. H. (1977). *Cognitive Development*. Englewood Cliffs, N.J.: Prentice-Hall.

Forbus, K. D., and D. Gentner. (1991). MAC/FAC: a model of similarity-based access and mapping. In K. J. Hammond and D. Gentner, eds., *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, pp. 504–509. Hillsdale, N.J.: Erlbaum.

Fromm, E. (1998). Lost and found half a century later: letters by Freud and Einstein. *American Psychologist* 53(11): 1195–1198.

Galison, P. (1987). *How Experiments End*. Chicago: University of Chicago Press.

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

Gentner, D., and M. Jeziorski. (1989). Historical shifts in the use of analogy in science. In B. Gholson, W. Shadish, R. Beimeyer, and A. Houts, eds., *The Psychology of Science: Contributions to Metascience*, pp. 296–325. Cambridge: Cambridge University Press.

Gholson, B., M. Levine, and S. Phillips. (1972). Hypotheses, strategies, and stereotypes in discrimination learning. *Journal of Experimental Child Psychology* 13: 423–446.

Gick, M. L., and K. J. Holyoak. (1983). Schema induction and analogic transfer. *Cognitive Psychology* 15: 138.

Giere, R. N. (1988). *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press.

Giere, R. N. (1993). Précis of cognitive models of science. *Psycoloquy* 4(56): <www.princeton.edu/~harnad/psyc.html>.

Gooding, D. (1990). *Experiment and the Making of Meaning*. Dordrecht: Nijhoff/Kluwer.

Gopnik, A. (1996). The scientist as child. *Philosophy of Science* 63(4):485–514.

Gorman, M. E. (1989). Error, falsification and scientific inference: An experimental investigation. *Quarterly Journal of Experimental Psychology* 41A(2): 385–412.

Gorman, M. E. (1992). *Simulating Science: Heuristics, Mental Models, and Technoscientific Thinking*. Bloomington: Indiana University Press.

Goswami, U. (1996). Analogical reasoning and cognitive development. *Advances in Child Development and Behavior* 26: 91–139.

Greeno, J. G. (1976). Cognitive objectives of instruction: theory of knowledge for solving problems and answering questions. In D. Klahr, ed., *Cognition and Instruction*, pp. 123–159. Hillsdale, N.J.: Erlbaum.

Groen, G. J. (1978). The theoretical ideas of Piaget and educational practice. In P. Suppes, ed., *Impact of Research in Education: Some Case Studies*. Washington, D.C.: National Academy of Education.

Gross, P. R., and N. Levitt. (1994). *Higher Superstition: The Academic Left and Its Quarrels with Science*. Baltimore: Johns Hopkins University Press.

Gruber, H. E. (1974). *Darwin on Man: A Psychological Study of Scientific Creativity*. New York: E. P. Dutton.

Hadamard, J. (1945). *The Psychology of Invention in the Mathematical Field*. New York: Dover.

Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Human Development* 35: 193–217.

Hayes-Roth, B., and F. Hayes-Roth. (1979). A cognitive model of planning. *Cognitive Science* 3: 275–310.

Hendrickson, J. B., and T. Sander. (1995). COGNOS: A Beilstein-type system for organizing organic reactions. *Journal of Chemical Information and Computer Sciences* 35(2): 251.

Holmes, F. L. (1985). *Lavoisier and the Chemistry of Life: An Exploration of Scientific Creativity*. Madison: University of Wisconsin Press.

Holyoak, K. J., and P. Thagard. (1995). *Mental Leaps: Analogy in Creative Thought*. Cambridge, Mass.: MIT Press.

Hovland, C., and E. Hunt. (1960). The computer simulation of concept attainment. *Behavioral Science* 5: 265–267.

Inhelder, B., and J. Piaget. (1958). *The Growth of Logical Thinking from Childhood to Adolescence*. New York: Basic Books.

Ippolito, M. F., and R. D. Tweney. (1995). The inception of insight. In R. J. Sternberg and J. E. Davidson, eds., *The Nature of Insight*, pp. 433–462. Cambridge, Mass.: MIT Press.

John-Steiner, V. (1997). *Notebooks of the Mind*. New York: Oxford University Press.

Judson, H. F. (1979). *The Eighth Day of Creation*. New York: Simon and Schuster.

Kaplan, C. A., and H. A. Simon. (1990). In search of insight. *Cognitive Psychology* 22: 374–419.

Karmiloff-Smith, A. (1988). A child is a theoretician, not an inductivist. *Mind and Language* 3: 183–195.

Karp, P. (1990). Hypothesis formation as design. In J. Shrager and P. Langley, eds., *Computational Models of Scientific Discovery and Theory Formation*, pp. 275–317. San Mateo, Calif.: Morgan Kaufmann.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review* 88: 197–227.

Kern, L. H., H. L. Mirels, and V. G. Hinshaw. (1983). Scientists' understanding of propositional logic: an experimental investigation. *Social Studies of Science* 13: 131–146.

King, M. L., Jr. (1963). *Strength to Love*. New York: Harper & Row.

Klahr, D. (1985). Solving problems with ambiguous subgoal ordering: preschoolers' performance. *Child Development* 56: 940–952.

Klahr, D. (1994). Searching for cognition in cognitive models of science. Psycoloquy 5(94): <www.princeton.edu/~harnad/psyc.html>.

Klahr, D., and S. M. Carver. (1988). Cognitive objectives in a LOGO debugging curriculum: instruction, learning, and transfer. *Cognitive Psychology* 20: 362–404.

Klahr, D., and S. M. Carver. (1995). Scientific thinking about scientific thinking. *Monographs of the Society for Research in Child Development* 245, Vol. 60, no. 3.

Klahr, D., and K. Dunbar. (1988). Dual space search during scientific reasoning. *Cognitive Science* 12: 1–55.

Klahr, D., K. Dunbar, and A. L. Fay. (1990). Designing good experiments to test bad hypotheses. In J. Shrager and P. Langley, eds., *Computational Models of Discovery and Theory Formation*, pp. 355–402. San Mateo, Calif.: Morgan Kaufmann.

Klahr, D., A. L. Fay, and K. Dunbar. (1993). Heuristics for scientific experimentation: a developmental study. *Cognitive Psychology* 24(1): 111–146.

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

Klahr, D., and H. A. Simon. (1999). Studies of scientific discovery: complementary approaches and convergent findings. *Psychological Bulletin* 125: 524–543.

Klayman, J., and Y. Ha. (1987). Confirmation, disconfirmation and information in hypothesis testing. *Psychological Review* 94: 211–228.

Klayman, J., and Y. Ha. (1989). Hypothesis testing in rule discovery: strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15(4): 596–604.

Kokar, M. M. (1986). Determining arguments of invariant functional descriptions. *Machine Learning* 1(4): 403–422.

Koslowski, B. (1996). *Theory and Evidence: The Development of Scientific Reasoning*. Cambridge, Mass.: MIT Press.

Koslowski, B., and L. Okagaki. (1986). Non-Humean indices of causation in problem-solving situations: causal mechanisms, analogous effects, and the status of rival alternative accounts. *Child Development* 57: 1100–1108.

Koslowski, B., L. Okagaki, C. Lorenz, and D. Umbach. (1989). When covariation is not enough: the role of causal mechanism, sampling method, and sample size in causal reasoning. *Child Development* 60: 1316–1327.

Krems, J., and T. R. Johnson. (1995). Integration of anomalous data in multicausal explanations. In J. D. Moore and J. F. Lehman, eds., *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*, pp. 277–282. Mahwah, N.J.: Erlbaum.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review* 96: 674–689.

Kuhn, D. (1992). Piaget's child as scientist. In H. Beilin and P. Pufall, eds., *Piaget's Theory: Prospects and Possibilities*, pp. 185–208. Hillsdale, N.J.: Erlbaum.

Kuhn, D., E. Amsel, and M. O'Loughlin. (1988). *The Development of Scientific Reasoning Skills*. Orlando, Fla.: Academic Press.

Kuhn, D., and J. Angelev. (1976). An experimental study of the development of formal operational thought. *Child Development* 47: 697–706.

Kuhn, D., M. Garcia-Mila, A. Zohar, and C. Andersen. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development* 60(4): pp. 1–128.

Kuhn, D., and E. Phelps. (1982). The development of problem-solving strategies. In H. Reese, ed., *Advances in Child Development and Behavior* Vol. 17, pp. 1–44. New York: Academic Press.

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kulkarni, D., and H. A. Simon. (1990). Experimentation in machine discovery. In J. Shrager and P. Langley, eds., *Computational Models of Scientific Discovery and Theory Formation*, pp. 255–273. San Mateo, Calif.: Morgan Kaufmann.

Lakatos, I., and A. Musgrave, eds. (1970). *Criticism and the Growth of Knowledge.* New York: Cambridge University Press.

Langley, P., H. A. Simon, G. L. Bradshaw, and J. M. Zytkow. (1987). *Scientific Discovery: Computational Explorations of the Creative Processes.* Cambridge, Mass.: MIT Press.

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

Latour, B., and S. Woolgar. (1986). *Laboratory Life: The Construction of Scientific Facts.* Princeton, N.J.: Princeton University Press.

Laudan, L. (1977). *Progress and Its Problems.* Berkeley: University of California Press.

Lovett, M. C., and J. R. Anderson. (1995). Making heads or tails out of selecting problem solving strategies. In J. D. Moore and J. F. Lehman, eds., *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*, pp. 265–270. Hillsdale, N.J.: Erlbaum.

Lovett, M. C., and J. R. Anderson. (1996). History of success and current context in problem solving: combined influences on operator selection. *Cognitive Psychology* 31: 168–217.

Mahoney, M. J. (1979). Psychology of the scientist: an evaluative review. *Social Studies of Science* 9: 349–375.

Mahoney, M. J., and B. G. DeMonbreun. (1978). Psychology of the scientist: an analysis of problem-solving bias. *Cognitive Therapy and Research* 1(3): 229–238.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner and A. L. Stevens, eds., *Mental Models*, pp. 299–324. Hillsdale, N.J.: Erlbaum.

Miller, G. A., E. Galanter, and K. Pribram. (1960). *Plans and the Structure of Behavior.* New York: Holt, Rinehart & Winston.

Minsky, M. (1975). *A Framework for Representing Knowledge: The Psychology of Computer Vision.* New York: McGraw-Hill.

Mitroff, L. L. (1974). *The Subjective Side of Science.* New York: Elsevier.

Mynatt, C. R., M. E. Doherty, and R. D. Tweney. (1977). Confirmation bias in a simulated research environment: an experimental study of scientific inference. *Quarterly Journal of Experimental Psychology* 29: 85–95.

Mynatt, C. R., M. E. Doherty, and R. D. Tweney. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology* 30: 395–406.

Nersessian, N. J. (1992). How do scientists think? Capturing the dynamics of conceptual change in science. In R. N. Giere, ed., *Cognitive Models of Science*, pp. 3–44. Minneapolis: University of Minnesota Press.

Newell, A. (1973). You can't play 20 questions with nature and win: projective comments on the papers of this symposium. In W. G. Chase, ed., *Visual Information Processing*, pp. 283–308. New York: Academic Press.

Newell, A. (1990). *Unified Theories of Cognition.* Cambridge, Mass.: Harvard University Press.

Newell, A., J. C. Shaw, and H. A. Simon. (1958). Elements of a theory of human problem solving. *Psychological Review* 65: 151–166.

Newell, A., and H. A. Simon. (1972). *Human Problem Solving.* Englewood Cliffs, N.J.: Prentice-Hall.

Nordhausen, B., and P. Langley. (1993). An integrated framework for empirical discovery. *Machine Learning* 12: 17–47.

O'Brien, D. P., G. Costa, and W. F. Overton. (1986). Evaluations of causal and conditional hypotheses. *Quarterly Journal of Experimental Psychology* 38: 493–512.

Okada, T. (1994). *Collaborative Scientific Discovery Processes.* Ph.D. dissertation, Carnegie Mellon University.

Okada, T., and H. A. Simon. (1997). Collaborative discovery in a scientific domain. *Cognitive Science* 21: 109–146.

O'Rorke, P., S. Morris, and D. Schulenburg. (1990). Theory formation by abduction: a case study based on the chemical revolution. In J. Shrager and P. Langley, eds., *Computational Models of Scientific Discovery and Theory Formation*, pp. 197–224. San Mateo, Calif.: Morgan Kaufmann.

Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas.* New York: Basic Books.

Penner, D., and D. Klahr. (1996*a*). The interaction of domain-specific knowledge and domain-general discovery strategies: a study with sinking objects. *Child Development* 67: 2709–2727.

Penner, D., and D. Klahr. (1996*b*). When to trust the evidence: further investigations of the effects of system error on the Wason 2-4-6 task. *Memory & Cognition* 24(5): 655–668.

Perkins, D. N. (1981). *The Mind's Best Work.* Cambridge, Mass.: Harvard University Press.

Piaget, J. (1952). *The Origins of Intelligence in Children.* New York: International University Press.

Pickering, A., ed. (1992). *Science as Practice and Culture.* Chicago: University of Chicago Press.

Pizzani, M. J. (1990). *Creating a Memory of Causal Relationships.* Hillsdale, N.J.: Erlbaum.

Poincaré, H. (1929). *The Foundation of Science.* New York: The Science Press.

Popper, K. R. (1959). *The Logic of Scientific Discovery.* London: Hutchinson.

Rajamoney, S. A. (1993). The design of discrimination experiments. *Machine Learning* 12: 185–203.

Rasmussen, J. (1981). Models of mental strategies in process plant diagnosis. In W. B. Rouse and J. Rasmussen, eds., *Human Detection and Diagnosis of System Failures*, pp. 241–258. New York: Plenum Press.

Reimann, P. (1990). *Problem Solving Models of Scientific Discovery Learning Processes.* Frankfurt am Main: Peter Lang.

Resnick, L. B., J. M. Levine, and S. D. Teasley. (1991). *Perspectives on Socially Shared Cognition.* Washington, D.C.: American Psychological Association.

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

Rowe, A. (1953). *The Making of a Scientist.* New York: Dodd, Mead.

Ruffman, T., J. Perner, D. R. Olson, and M. Doherty. (1993). Reflecting on scientific thinking: children's understanding of the hypothesis-evidence relation. *Child Development* 64: 1617–1636.

Sadish, W. R., and S. Fuller, eds. (1994). *Social Psychology of Science.* New York: Guilford Press.

Samarapungavan, A. (1992). Children's judgments in theory-choice tasks: scientific rationality in childhood. *Cognition* 45: 1–32.

Schauble, L. (1990). Belief revision in children: the role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology* 49: 31–57.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology* 32(1): 102–119.

Schauble, L., and R. Glaser. (1990). Scientific thinking in children and adults. In D. Kuhn, ed., *Developmental Perspectives on Teaching and Learning Thinking Skills*, pp. 9–26. Basel and New York: Karger.

Schauble, L., R. Glaser, K. Raghavan, and M. Reiner. (1991). Causal models and experimentation strategies in scientific reasoning. *Journal of the Learning Sciences* 1: 201–238.

Schauble, L., L. E. Klopfer, and K. Raghavan. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching* 28: 859–882.

Scholnick, E. K., and S. L. Friedman. (1987). The planning construct in the psychological literature. In S. L. Friedman, E. K. Scholnick, and R. R. Cocking, eds., *Blueprints for Thinking*, pp. 3–38. Cambridge: Cambridge University Press.

Schunn, C. D., and K. Dunbar. (1996). Priming, analogy, and awareness in complex reasoning. *Memory & Cognition* 24(3): 271–284.

Schunn, C. D., and D. Klahr. (1992). Complexity management in a discovery task. In J. K. Kruschke, ed., *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, pp. 177–182. Mahwah, N.J.: Erlbaum.

Schunn, C. D., and D. Klahr. (1993). Other vs. self-generated hypotheses in scientific discovery. In W. Kintsch, ed., *Proceedings of the Fifteenth Annual Meeting of the Cognitive Science Society*, p. 900. Mahwah, N.J.: Erlbaum.

Schunn, C. D., and D. Klahr. (1995). A 4-space model of scientific discovery. In J. D. Moore and J. F. Lehman, eds., *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*, pp. 106–111. Mahwah, N.J.: Erlbaum.

Schunn, C. D., and D. Klahr. (1996). Integrated yet different: logical, empirical, and implementational arguments for a 4-space model of inductive problem solving. In G. Cottrell, ed., *Proceedings of the Eighteenth Annual Meeting of the Cognitive Science Society*, pp. 25–26. Mahwah, N.J.: Erlbaum.

Schuster, M. L. (1940). *A Treasury of the World's Great Letters from Ancient Days to Our Own Time*. New York: Simon and Schuster.

Scott, P. D., and S. Markovitch. (1993). Experience selection and problem choice in an exploratory learning system. *Machine Learning* 12: 49–67.

Shaklee, H., and D. Paszek. (1985). Covariation judgment: systematic rule use in middle childhood. *Child Development* 56: 1229–1240.

Shen, W. M. (1993). Discovery as autonomous learning from the environment. *Machine Learning* 12: 143–165.

Shrager, J. (1985). *Instructionless Learning: Discovery of the Mental Model of a Complex Device*. Ph.D. dissertation, Carnegie Mellon University.

Shrager, J. (1987). Theory change via view application in instructionless learning. *Machine Learning* 2: 247–276.

Shrager, J., and P. Langley, eds. (1990). *Computational Models of Scientific Discovery and Theory Formation*. San Mateo, Calif.: Morgan Kaufmann.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology* 8: 481–520.

Siegler, R. S. (1987). The perils of averaging data over strategies: an example from children's addition. *Journal of Experimental Psychology* 116: 250–264.

Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking*. New York: Oxford University Press.

Siegler, R. S., and E. Jenkins. (1989). *How Children Discover New Strategies*. Hillsdale, N.J.: Erlbaum.

Siegler, R. S., and R. M. Liebert. (1975). Acquisition of formal scientific reasoning by 10- and 13-year-olds: designing a factorial experiment. *Developmental Psychology* 10: 401–402.

Simon, H. A. (1966). Scientific discovery and the psychology of problem solving. In R. Colony, ed., *Mind and Cosmos*, pp. 22–40. Pittsburgh: University of Pittsburgh Press.

Simon, H. A. (1975). The functional equivalence of problem-solving skills. *Cognitive Psychology* 7: 268–288.

Simon, H. A. (1977). *Models of Discovery*. Dordrecht: Reidel.

Simon, H. A. (1983). Understanding the processes of science: the psychology of scientific discovery. In T. Gamelius, ed., *Progress in Science and its Social Conditions*, pp. 159–170. Oxford: Pergamon.

Simon, H. A., and K. Kotovsky. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review* 70(6): 534–546.

Simon, H. A., P. Langley, and G. L. Bradshaw. (1981). Scientific discovery as problem solving. *Synthese* 47: 1–27.

Simon, H. A., and G. Lea. (1974). Problem solving and rule induction: a unified view. In L. Gregg, ed., *Knowledge and Cognition*, pp. 105–128. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Skinner, B. F. (1956). A case history in scientific method. *The American Psychologist* 11: 221–233.

Sloutsky, V. M., A. W. Rader, and B. J. Morris. (1998). Increasing informativeness and reducing ambiguities: adaptive strategies in human information processing. In M. A. Gernsbacher and S. J. Derry, eds., *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*, pp. 997–1002. Mahwah, N.J.: Erlbaum.

Sodian, B., D. Zaitchik, and S. Carey. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development* 62: 753–766.

Sokal, A. (1996). Transgressing the boundaries: towards a transformative hermeneutics of quantum gravity. *Social Text* 46/47: 217–252.

Teasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology* 31: 207–220.

Terman, L. M. (1954). Scientists and nonscientists in a group of 800 men. *Psychological Monographs* 68, no. 378.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences* 12: 435–502.

Thagard, P. (1992). *Conceptual Revolutions*. Princeton, N.J.: Princeton University Press.

Thagard, P. (1997). Medical analogies: why and how. In M. G. Shafto and P. Langley, eds., *Proceedings of the Nineteenth Annual Meeting of the Cognitive Science Society*, pp. 739–763. Mahwah, N.J.: Erlbaum.

Thagard, P. (1998). Ulcers and bacteria I: discovery and acceptance. *Studies in the History and Philosophy of Biology and Biomedical Sciences* 29: 107–136.

Tschirgi, J. E. (1980). Sensible reasoning: a hypothesis about hypotheses. *Child Development* 51: 1–10.

Tumblin, A., and B. Gholson. (1981). Hypothesis theory and the development of conceptual learning. *Psychological Bulletin* 90: 102–124.

Tweney, R. D. (1990). Five questions for computationalists. In J. Shrager and P. Langley, eds., *Computational Models of Discovery and Theory Formation*, pp. 471–484. San Mateo, Calif.: Morgan Kaufmann.

Tweney, R. D., M. E. Doherty, and C. R. Mynatt, eds. (1981). *On Scientific Thinking*. New York: Columbia University Press.

Udea, K. (1997). Actual use of analogies in remarkable scientific discoveries. In M. G. Shafto and P. Langley, eds., *Proceedings of the Nineteenth Annual Meeting of the Cognitive Science Society*, pp. 763–768. Mahwah, N.J.: Erlbaum.

Valdés-Pérez, R. E. (1994*a*). Conjecturing hidden entities via simplicity and conservation laws: machine discovery in chemistry. *Artificial Intelligence* 65(2): 247–280.

Valdés-Pérez, R. E. (1994*b*). Algebraic reasoning about reactions: discovery of conserved properties in particle physics. *Machine Learning* 17(1): 47–68.

Valdés-Pérez, R. E. (1994*c*). Human/computer interactive elucidation of reaction mechanisms: application to catalyzed hydrogenolysis of ethane. *Catalysis Letters* 28(1): 79–87.

Valdés-Pérez, R. E. (1995). Some recent human/computer discoveries in science and what accounts for them. *Artificial Intelligence* 16(3): 37–44.

Vosniadou, S., and W. F. Brewer. (1992). Mental models of the earth: a study of conceptual change in childhood. *Cognitive Psychology* 24: 535–585.

Vosniadou, S., and W. F. Brewer. (1994). Mental models of the day/night cycle. *Cognitive Science* 18: 123–183.

Ward, S. L., and W. F. Overton. (1990). Semantic familiarity, relevance, and the development of deductive reasoning. *Developmental Psychology* 26(3): 488–493.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12: 129–140.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology* 20: 273–281.

Weisberg, R. W. (1993). *Creativity: Beyond the Myth of Genius*. New York: W. H. Freeman.

Wellman, H. M. (1983). Metamemory revisited. In M. T. H. Chi, ed., *Trends in Memory Development Research*, pp. 31–51. New York: Karger.

Wellman, H. M., and S. Gelman. (1992). Cognitive development: foundational theories of core domains. *Annual Review of Psychology* 43: 337–375.

Wertheimer, M. (1945). *Productive Thinking*. (Reprint, 1959.) New York: Harper & Row.

Wickelgren, W. A. (1974). *How to Solve Problems: Elements of a Theory of Problems and Problem Solving*. San Francisco: W. H. Freeman.

Wiser, M. (1989). Does learning science involve theory change? Paper presented at the biannual meeting of the Society for Research in Child Development, Kansas City.

Wolf, D. F., and J. R. Beskin. (1996). Task domains in N-space models: giving explanation its due. In G. Cottrell, ed., *Proceedings of the Eighteenth Annual Meeting of the Cognitive Science Society*, pp. 27–28. Mahwah, N.J.: Erlbaum.