Chapter Seven

Discovering the Present by Predicting the Future

David Klahr

> Remember that the future is neither ours nor wholly not ours, so that we may neither count on it as sure to come nor abandon hope of it as certain not to be
>
> (Epicurus, 300 B C)

> If a man carefully examine his thoughts he will be surprised to find how much he lives in the future His well-being is always ahead Such a creature is probably immortal
>
> (Ralph Waldo Emerson, 1827)

Marshall Haith solicited contributions to this volume with a claim and a challenge. His claim was that contemporary psychologists have had a "dearth of concern" with future-oriented phenomena. His challenge was for us to view our own work in such terms.

My initial response to the claim was skepticism: "Certainly people

This work was supported in part by grants from the National Institute of Child Health and Human Development (RO1-HD25211) and the Andrew W Mellon Foundation The chapter is based in part on work done in collaboration with Mitchell Robinson, Kevin Dunbar, and Anne L Fay I am indebted to the volume editors as well as to Sharon Carver, Anne L. Fay, Ken Kotovksy, David Penner, Christian Schunn, and especially Shari Ellis for their comments and suggestions on earlier drafts

have grappled with the psychology of the future," I thought. But as the opening quotations suggest, most of what has been said on the topic has come from philosophers rather than psychologists (although one might anachronistically interpret Epicurus's statement as a comment on decision making under uncertainty and Emerson's as relevant to metacognition). And though there is an extensive literature on the psychology of time (Friedman, 1990, estimates thousands of articles in the last 100 years), little of it has addressed the psychology of thinking about the future, and much of it tends to be inconsistent and inconclusive (see Benson's summary, this volume).

Before writing this chapter, I had not thought about my own research areas—problem solving and scientific discovery—in terms of future-oriented processes. However, Haith's challenge motivated me to reflect on the extent to which the behaviors I was studying involved future-oriented thinking. In so doing, I attempted to lay out a framework for classifying future-oriented processes within which I could locate my own work (and ultimately work in other domains). Therefore my present goal is to review some of my research on the development of problem solving and scientific discovery skills and to recast that work in terms of what it suggests about how adults and children think about the future.

In the first section, I propose a framework for considering future-oriented processes. In the second, I lay the groundwork for discussing future orientation in terms of problem solving. In the next two sections I summarize two lines of research that are based on this problem-solving orientation, and I interpret the results of those studies in terms of future-oriented processing: the third section summarizes some investigations of preschool children's ability to think ahead in the context of simple, well-defined puzzles, and the fourth summarizes research on the development of scientific reasoning skills. Here the problem-solving formulation is extended to the domains of hypothesis formation and experimental design. In the final section, I return to the framework and attempt to summarize the important future-oriented attributes of the processes of problem solving and scientific discovery.

Before closing this introductory section, I must explain the somewhat cryptic title I have chosen. It derives from considering the future-oriented implications of the work on scientific discovery to be described in the fourth section. Attempts to predict the future behaviors of a complex system are successful only insofar as its underlying principles—operating in the present—have been discovered. To the extent that our predictions turn out to be incorrect, we must revise our characterization of the current situation. That is, we discover the present by predicting the future.

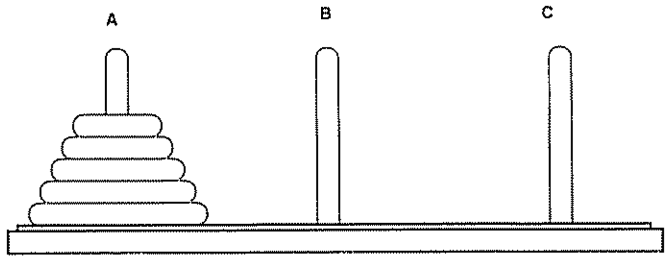

Figure 7.1 Five-disk Tower of Hanoi puzzle The goal is to move all the disks from peg A to peg C subject to the following constraints: (a) only one disk can be moved at a time; (b) a larger disk can never be placed above a smaller disk. The minimum path requires 31 moves. (In general, an n-disk problem requires $2^n - 1$ moves.)

A FRAMEWORK FOR CONSIDERING FUTURE-ORIENTED PROCESSES

Future-oriented processes appear to differ in the following critical attributes: uncertainty, control, contingency, abstraction, social dependency, and grain of abstraction. I can best explain these attributes by example. Suppose you are thinking about making the first move in a puzzle like the well-known Tower of Hanoi (TOH), shown in Figure 7.1. One could ask the following questions about your thought processes:

Uncertainty. How certain can you be about the possible outcomes of the move? In principle, deterministic puzzles with perfect information—such as the TOH—have no uncertainty. You simply consider every possible first move you could make, all possible second moves, and so on, until you have reached the goal. Then you make the first move leading to a minimum path solution. However, even in simple puzzles the combinatorics of this approach may require considering an enormous number of moves, and such computations would overwhelm the limited human cognitive capacity. Thus, even in a formally deterministic puzzle, the degree of uncertainty depends on both the demands of the external environment and the limitations of the human system. In contrast, some situations are inherently uncertain, such as predicting the roll of dice, or the deal of a hand of cards.

Control. To what extent are you in control of the current and future situation? In two-person games like chess, control is only partial, because your opponent's possible moves must form a part of your view of the future. In contrast, in the TOH there is no adversary. Once you make a legal move, you can be sure of reaching the state that move is supposed to produce. As noted, even in the TOH, planning a full solution path involves some uncertainty, but there is no agent beyond you that will be making any of the moves.

Type of contingency. What kind of environmental response will be evoked by your move? In zero-sum games like chess and checkers, it is clear that the response will be *adversarial.* The opponent will attempt to undermine your attempt to determine the future course of events. In the TOH there is no such contingent reaction. Nothing in the context responds to whatever state you may reach. In collaborative or joint problem-solving situations, the contingency is *supportive.* There are other agents in the context who will support and guide your outcomes.

Abstraction. Is the "grain"—the level of detail—of your thinking at the same level as the grain of the relevant events? One function of a plan is to suppress detail so that a rough sketch of the solution path can be formulated. In chess one can suppress detail and make a general plan (center control, queen side attack, etc.), but each of these plans is ultimately unpacked into finer-grain units that eventuate in a legal move.

Social. Do you have to consider the future orientation of other people in your own deliberations? This is related to the contingency issue listed earlier, but it is focused entirely on the impact that other human problem solvers, and your representation of *their* decision-making and problem-solving situation is required in your own. This would involve some estimate of their own goals and priorities. (The classic Prisoner's Dilemma paradigm for studying cooperation and conflict exemplifies the issues here.)

Temporal extent. What is the order of magnitude of the temporal interval involved? Thinking about the future can be limited to very brief intervals or extend to planning one's life. It may be that the cognitive processes involved for vastly different temporal extents are quite distinct. For example, the range of future-oriented processes discussed in this chapter varies from a couple of seconds to an hour or so, and real scientific reasoning can go on for months and years.

At this point in the development of the taxonomy, it is likely that these attributes are neither mutually exclusive nor exhaustive nor independent. Nevertheless, I offer them as a useful starting point in the endeavor. I will attempt to answer the questions listed above in the context of research on the development of problem-solving and scientific reasoning skills. The projects to be described are based on the view that a wide range of higher-order cognitive processes can be viewed as different types of problem solving. Therefore I will preface the description of those projects with a few general comments on what I mean by a problem.

PROBLEM SOLVING

Newell and Simon (1972) define a problem as comprising an initial state, a goal state, and a set of operators that allow the problem solver to transform the initial state into the goal state via a series of intermediate states. Operators have constraints that must be satisfied before they can be applied. The set of states, operators, and constraints is called a "problem space," and the problem-solving process can be characterized as a search for a path that links the initial state to the goal state. (But the search need not be constrained to start with the initial state. Indeed, working backward—from goal state to initial state—is an appropriate procedure in some situations.)

Weak Methods

In all but the most trivial problems, the problem solver is faced with a very large set of alternative states and operators, so the search process can be 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 very small values, exhaustive search for alternative states and operators is beyond human capacity,[1] so effective problem solving depends in large part on how well the search is constrained. Newell and Simon (1972) divided different approaches to search constraint into two broad categories: *strong methods* and *weak methods*.

Strong methods are algorithmic procedures, such as those for long division or for computing means and standard deviations. The most important aspect of strong methods is that—by definition—they guarantee a solution to the problem they are designed to solve. However, strong methods have several disadvantages for human problem solvers. First, they may require extensive computational resources. For example, a strong method for minimizing cost (or maximizing protein) of a list of grocery items subject to other dietary and budget constraints is to apply a standard linear-programming algorithm (Hadley, 1962). Of course doing this in one's head while pushing a shopping cart is hardly feasible. Second, strong methods may be difficult to learn because they may require many detailed steps (for example, the procedure for inverting a matrix, or computing a correlation coefficient by hand). Finally, strong methods, by their very nature, tend to be domain specific and thus have little generality.

1. For example, Newell and Simon (1972, p. 669) 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."

Weak methods are heuristic: they may work or they may not, but they are highly general. The trade-off for the lack of certainty associated with weak methods is that they make substantially lower computational demands, are more easily acquired (indeed, some may be innate), and are domain general. Newell and Simon (1972) describe several kinds of weak methods, used by both human problem solvers and artificial intelligence systems. I will describe only four such weak methods and comment on the differences in their "future orientedness."

Generate and Test

The generate and test 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, then some other operator is applied. In the most primitive 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 is searching in a box of keys for a key to fit a lock, and sampling 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 forth. A slightly "smarter" generator would, at the least, sample from the key box without replacement.

It is difficult to attribute much of a future orientation to this method, beyond noting that the goal state is something that has not yet happened and that, if it does occur, it will occur in the future. Neither progress nor history is represented in the most primitive generate and test methods.

Hill Climbing

The hill climbing method gets its name from the analogy of 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, 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 about the discrepancy between the current state and the goal state than does generate and test.

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. However, its representation of the future—the evaluation function—is primitive in two regards. First, it is very local because it anticipates only the next step in the solution path. Second, it is at an aggregate level that does not include any details about the structure of future states.

Means-Ends Analysis

Of all the weak methods, perhaps the best known is means-ends analysis (Dunker, 1945; Newell & Simon, 1972). Means-ends analysis compares the current state with the goal state and describes any differences. Then it searches for operators that can reduce those differences. It selects an operator designed to reduce the most important differences and attempts to apply it to the current state. However, it may be that the operator cannot be immediately applied because the conditions for doing so 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.

As a homely example, consider the problem I faced in getting from my office at Carnegie Mellon University to the conference room in Breckenridge, Colorado, where I first presented the talk this chapter is based on. 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 of choice, but I could not fly directly from my office to Breckenridge. This presented the subproblem of creating conditions for flying (getting to an airport). Getting to the airport could best be done by 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 I could not apply that operator (no pun intended) because I had no change. A Coke machine was handy, however, and it accepted dollar bills and gave change. So I bought a Coke in order to get on the solution path to transport myself to Colorado.

Means-ends analysis constructs a hierarchy of goals and subgoals, with the parent node for a state corresponding to why something is being done and the descendants of a node corresponding to how it will be done. Because means-ends analysis computes a qualitative evaluation of the difference between the current state and the goal state, and because it generates a goal tree, it requires a highly articulated representation of both the present and the future. Of the methods described thus far, means-ends analysis has the greatest future orientation.

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.

If we apply the planning method to the problem of getting to Breckenridge, we might produce a three-step plan: (1) take taxi to airport; (2) fly to Denver; (3) drive rental car to Breckenridge. This plan contains none of the details about phone calls, change, and Coke machines. 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, there might be no rental cars at the Denver airport.

Of the four methods described here, planning has the strongest future orientation in its broad sweep from the current state to the goal state. It actually produces a sketch of the solution before the solution itself is executed and in that sense anticipates the future.

Weak Methods and the Attributes of Future Orientation

This set of weak methods can be applied in a wide variety of contexts involving different levels of uncertainty, control, contingency, and so on. That is, the weak methods can be crossed with the future-oriented attributes listed in the previous section to form a large space of future-oriented situations. In the following sections I describe a small subset

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, & Pribram, 1960, p. 16) to "the predetermination of a course of action aimed at achieving a goal" (Hayes-Roth & Hayes-Roth, 1979, p. 275). An elaboration and discussion of the many definitions can be found in Scholnick and Friedman (1987). I use the Newell and Simon (1972) version here because it is much better defined than the others and fits nicely in the set of weak methods.

of this space, based on several investigations of problem solving and scientific reasoning in children and adults.

PROBLEM-SOLVING METHODS USED BY PRESCHOOL CHILDREN

The broad generality of these four methods (as well as several others) raises the question of their developmental course. To the extent that these methods vary in their degree of future orientation, understanding their developmental trajectory might give us some insight into children's ability to anticipate and represent the future.

My interest in children's problem solving was stimulated by what I perceived to be a discrepancy between Piaget's (1976) claims about the limited problem-solving capacities of preschoolers and my observations of my own young children. Piaget (1976) used two-, three-, and four-disk TOH problems (see Figure 7.1) with children from about 5½ to 12 years old. He reported that most 5- and 6-year-old children "cannot move the three-disk tower even after trial and error. They do succeed in moving the two-disk tower, but only after all sorts of attempts to get around the instructions and without being conscious of the logical links" (p. 288). From this performance Piaget concluded that "none of these subjects make a plan or even understand how they are going to move the tower" (p. 290); and later, "There is . . . a systematic primacy of the trial-and-error procedure over any attempt at deduction, and no cognizance of any correct solution arrived at by chance" (p. 291). In contrast, I could see behavior in my children that strongly suggested a capacity to plan modestly complex action sequences, such as using one kind of object to facilitate getting another that could then be used to accomplish a goal (Klahr, 1978). In this section I will describe two studies designed to investigate the kinds of methods preschool children use when faced with novel puzzles that require them to "think ahead."

Preschoolers' Problem Solving on the Tower of Hanoi

The first task I will describe is a modification of the Tower of Hanoi shown earlier. This puzzle has been used extensively to study adults' problem solving (Simon, 1975; Anzai & Simon, 1979). It conforms to the definition of a well-defined problem given earlier in that it contains unambiguous descriptions of an initial state, a final state, and legal moves (operators). The difficulty lies in discovering the sequences of legal moves that transform the initial configuration into the desired one. To use this task with young children, we modified it in several ways that changed its superficial appearance while maintaining its basic structure (Klahr & Robinson, 1981).

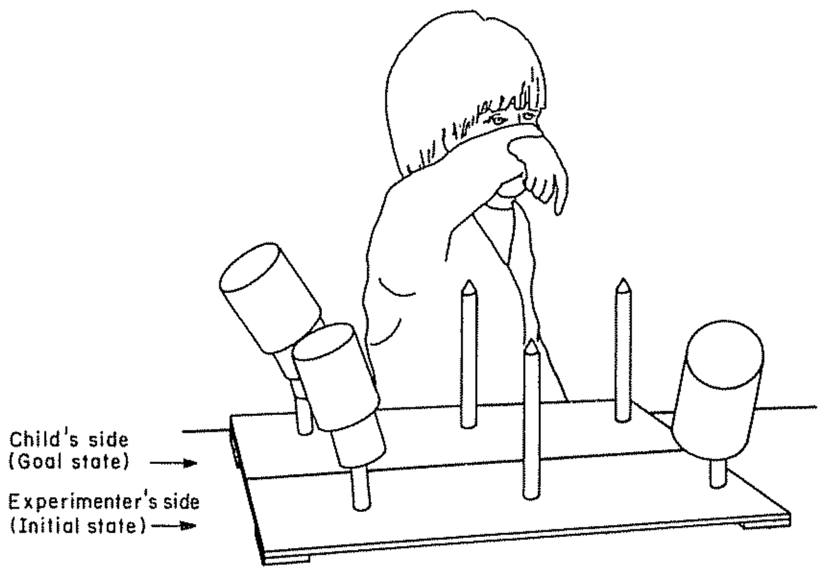

Figure 7.2 Child seated in front of "monkey cans" working on a one-move problem (state 2 to state 1: see Figure 7.3).

Materials. We reversed the size constraint and used a set of nested inverted cans that fit loosely on the pegs. When they were stacked up it was impossible to put a smaller can on top of a larger can (see Figure 7.2). Even if the child forgot the relative size constraint, the materials provided an obvious physical consequence of attempted violations: smaller cans fell off bigger cans.

Externalization of final goal. In addition to the initial configuration, the goal configuration was always physically present. We arranged the child's cans in a goal configuration and the experimenter's cans in the initial configuration. Then the child was asked to tell the experimenter what she (the experimenter) should do in order to get her (experimenter's) cans to look just like the child's. This procedure was used to elicit the child's reasoning about several future states: children were asked to describe the complete *sequence* of moves necessary to solve the problem.

Cover story. Problems were presented in the context of a story in which the cans were monkeys (large daddy, medium-size mommy, and small baby), who jump from tree to tree (peg to peg). The child's monkeys were in some good configuration, the experimenter's monkeys were "copycat" monkeys who wanted to look just like the child's monkeys. The cans were redundantly classified by size, color, and family membership for easy reference. Children found the cover story easy to comprehend and remember, and they readily agreed to consider the cans as

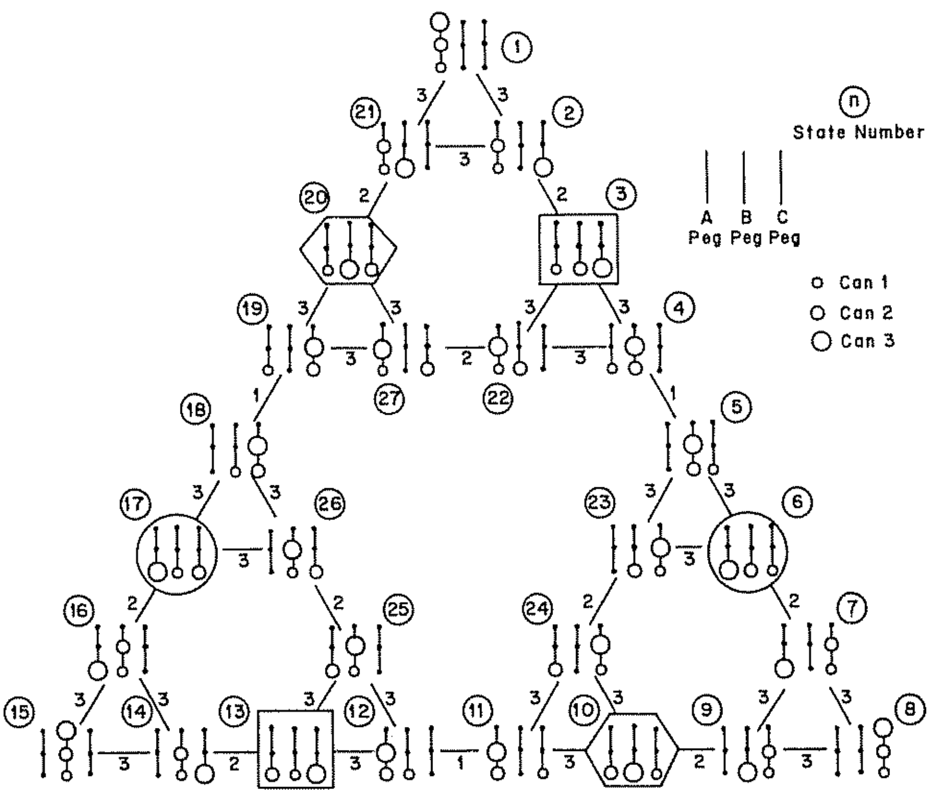

Figure 7.3 State space of all legal configurations and moves for three-can problem. States are arbitrarily numbered from 1 to 27. The can that moves between one state and the next is indicated on the line connecting the states. The minimum path solutions for all tower-to-tower problems (see text) are seven moves. Flat-to-flat problems enclosed by matching geometric shapes (e.g., state 17 to state 6) also have seven-move minimum path solutions.

monkeys. The remaining variations are best described after considering some of the formal properties of this task.

Figure 7.3 shows the state space for this problem: all possible legal states and moves. Each state is one move distant from its neighbors, and the can that is moved is indicated by the number on the line connecting adjacent states. The solution to a problem can be represented as a path through the state space. For example, the minimum solution path for the problem that starts with all three cans on peg A (state 1) and ends with them on peg C (state 8) is shown along the right-hand side of the large triangle in Figure 7.3. The first move involves shifting the largest can (can 3) from peg A to peg C, producing state 2. The next move places can 2 on peg B (state 3), followed by a move of can 3 to peg B (state 4), and so on. The "standard" TOH problems always end with all the cans stacked up on one peg. We call these "tower-ending" (T-end) problems. In the six states indicated by the large squares, circles, and hexagons in Figure 7.3, all pegs are occupied. We call any problem that ends in one of these states "flat-ending" (F-end). We used F-end as well as the more commonly used T-end problems. As we shall see, for the children in this study, F-ends were much more difficult.

Subjects

Fifty-one children attending the Carnegie Mellon University Children's School participated in the study. There were 19 children each in the 4-year and 5-year groups and 13 in the 6-year group. The children came predominantly, but not exclusively, from white, middle-class backgrounds. There were approximately equal numbers of boys and girls at each age level.

Procedure

Children were familiarized with the materials shown in Figure 7.2, in the context of the following cover story.

> Once upon a time there was a blue river (experimenter points to space between rows of pegs). On your side of the river there were three brown trees. On my side there were also . . . , etc. On your side there lived three monkeys: a big yellow daddy (present yellow can and place on peg), a medium size blue mommy (present and place), and a little red baby. The monkeys like to jump from tree to tree [according to the rules]; they live on your side of the river. (Establish legal and illegal jumps.) On my side there are also three: a daddy, [etc.] (introduce experimenter's cans). Mine are copycat monkeys. They want to be just like yours, right across the river from yours. Yours are all stacked up like so [state 1] mine are like so [state 2 or state 21]. Mine are very unhappy because they want to look like yours, but right now they are a little mixed up. Can you tell me what to do in order to get mine to look like yours? How can I get my daddy across from your daddy [etc.]?

During the initial part of the familiarization phase, the child was allowed to handle the cans but was gradually dissuaded from doing so and was instead encouraged to tell the experimenter what she should do in order to get her cans to look like the child's.

The final procedural variation we used was designed to satisfy two opposing constraints. On the one hand, in order to give a precise diagnostic of children's problem-solving strategies, we wanted to use a rule assessment procedure (Siegler, 1981), but that required a relatively long series of problems (up to 40). On the other hand, we wanted to minimize test effects that might result from children's learning about different path segments in the state space. Therefore we proceeded as follows:

for each problem the child told the experimenter the full sequence of proposed moves, and the experimenter gave supportive acknowledgment *but did not move the cans*. Then the next problem was presented.

We used a set of 40 problems: four problems having minimum path lengths of one, two, three, and four, and eight problems each with path lengths five, six, and seven. For each path length, half the problems were T-end and half were F-end. Problems were presented in two blocks with only T-end problems in one block, only F-end problems in the other. Children were randomly assigned to one of the two block orders (F-T or T-F).

Within a given block increasingly difficult problems were presented in order until the child appeared to reach his upper limit. There were several indicators of this upper limit: (a) explicit statements of confusion or inability to continue; (b) abrupt violation of rules of the game (e.g., putting monkey in the river); (c) sudden loss of motivation; (d) consistent errors in planned moves. At this point the session was terminated.

Scoring

Videotape recordings of children's behavior were transcribed and scored as shown in the two examples in Table 7.1. The child's move sequences were encoded as shown on the right side of the table. Recall that no cans were actually moved during these protocols, so all the configurations shown in the "results" column (except the initial and final ones) are imagined rather than real. The two protocols shown in Table 7.1 were scored as perfect six-move solutions.

Results

The main question of interest is how far into the future a child could "see" in describing move sequences. To avoid overestimating this capacity on the basis of a few fortuitous solutions, we used a very strict criterion. A child was scored as able to solve *n*-move problems only after proposing the minimum path solution for *all four* of the problems of length *n*. For example, to be classified as having the capacity to see five moves into the future, a child would have to produce the minimum path solution for the four five-move problems.[3]

The proportion of subjects in each age group producing correct solutions for all problems of a given length is shown in Figure 7.4a for T-end problems and Figure 7.4b for F-end problems. It is important to reemphasize that the abscissa in Figure 7.4 is not overall proportion correct, but rather a much more severe measure: the proportion of

3. The four five-move problems included the following initial-final states shown in Figure 7.3: 23–1, 3–8, 22–8, and 6–1.

TABLE 7.1 Two Six-Move Protocols and Their Encodings

	3/21/– Initial (state 16)	321/–/– Goal (state 1)	Move	Result 3/21/–	(Initial)
What you do is you put the daddy (3) . . .					
What you do is you move the daddy (3) over this tree (points to C)					
and move and move the baby .					
and then you move the mommy (2)					
Wait; where could you move the mommy (2) to?					
Well first move the mommy (2) on this tree (points to C)			2BC	3/ 1/2	
then put the daddy (3) on that tree (points to C)			3AC	–/ 1/32	
and put the baby (1) over there (points to A)			1BA	1/ –/32	
Then how would the mother? . .					
And after you put the baby (1) over here (points to A)					
you could put the daddy (3) (points to B)			3CB	1/ 3/2	
then you could put the mommy (2) over the baby,			2CA	21/ 3/–	
and the daddy over the mommy			3BA	321/–/–	
	3/–/21 Initial (state 7)	321/–/– Goal (state 1)		3/–/21	(Initial)
Oh, that. OK. That's easy					
Just take the yellow one (3) and put it on there (B)					
Take the (pointing to 2[C]) . . and take . . . , take the ba					
No, take the blue one (2), put it on there (B),			2CB	3/ 2/ 1	
and the, then take the yellow (3)					
and put it on the blue (points toward C, then to B),			3AB	–/32/ 1	
and then take the red (1) one and put it on here (A)			1CA	1/32/–	
And then take the blue (2) one and					
no, and then . . . and then put the yellow (3) one here (C)			3BC	1/ 2/ 3	
and then put the blue one (2) on the red one,			2BA	21/–/ 3	
and then put the yellow one on the blue one.			3CA	321/–/–	

subjects with perfect solutions on all problems of a given length. For example, 69% of the 6-year-olds were correct on all four of the five-move problems, while only 16% of the 5-year-olds and 11% of the 4-year-olds produced four flawless five-move solutions.

The absolute level of performance was striking, given the results of previous studies with children on this task (e.g., Piaget, 1976). On the T-end problems over two-thirds of the 5-year-olds and nearly all of the 6-year-olds consistently gave perfect four-move solutions, and over half of the 6-year-olds gave perfect six-move solutions. Almost half of the 4-year-olds could do the three-move problems. Recall that these solutions required that the child manipulate mental representations of future states, because the cans were not moved during or after the child's

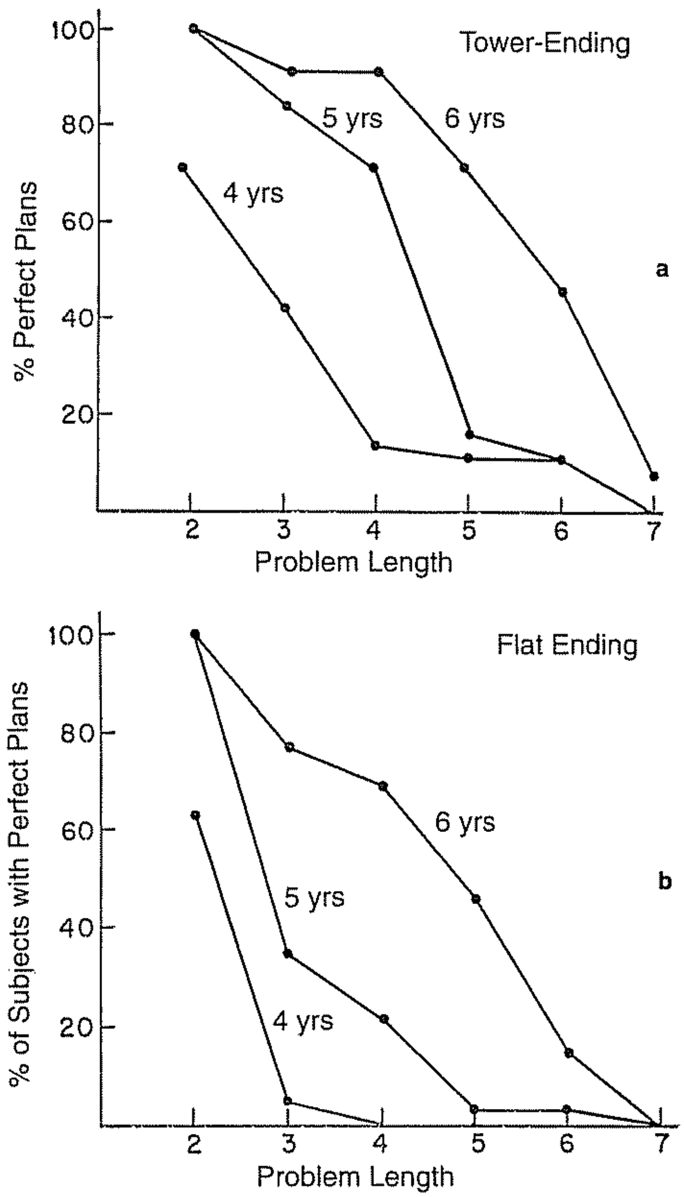

Figure 7.4 Proportion of children at each age level producing perfect solutions: (a) T-end problems; (b) F-end problems.

description of the solution sequence. Furthermore, all intermediate states were different from, but highly confusable with, the two physically present states (the initial and final configurations).

A different picture emerged with F-end problems. One-third of the youngest children could not do anything beyond a one-move problem, and barely one-third of the 5-year-olds could reliably do the three-move problems. Although the 6-year-olds did much better than the two younger groups, their F-end scores were also substantially below their T-end levels.

Summary: Future Orientation in Preschoolers' Solutions to the Tower of Hanoi

Our analysis revealed two important results—one relating to absolute performance, the other to the effects of goal configuration. First, it is clear that many 6-year-olds and some 5-year-olds are able to look ahead six moves into the future, even in a novel domain with arbitrary goals and constraints, as long as the subgoals are easily ordered. This ability appears to result from systematic application of both planning and means-ends analysis, as suggested by the typical protocols shown in Table 7.1.

Second, the relative difficulty of formally equivalent problems (same materials, rules, state space, and path length) depends on the form of the goal configuration. Consider the seven-move, F-end problem from state 15 to state 3: Which can will reach its ultimate destination first? It is not immediately apparent. In constrast, for the T-end problem that goes in the opposite direction (from state 3 to state 15), it is clear that the smallest can will have to reach the goal peg first, then the middle-size can, and so on. More generally, our results suggest that when the surface form of the problem did not suggest an unambiguous ordering of subgoals, children had a difficult time applying means-ends analysis.

How do these results bear on the attributes of future-oriented thinking listed earlier? Although there was no difference between T-end and F-end problems with respect to control, contingency, social dependency, or temporal grain, the two types of problems did differ with respect to effective uncertainty and grain of abstraction. On T-end problems, children were able to minimize the uncertainty that might have been introduced by their own inability to keep track of the future sequence of states, because they could focus on the sequence in which objects had to reach their final position in the goal state. This unambiguous subgoal ordering enabled children to apply both the planning method and means-ends analysis. They could plan to achieve the subgoal sequence (at a slightly abstracted level of analysis) and then "unpack" the plan by applying means-ends analysis in order to execute the detailed move sequence necessary to achieve that subgoal. For F-end problems, the ambiguous ordering of subgoals rendered planning difficult. This in turn made it difficult to deal with the subgoals at a slightly abstracted level (as in planning) and kept the solutions at a very low level. Given children's capacity limitations, this difficulty increased the uncertainty level of move outcomes.

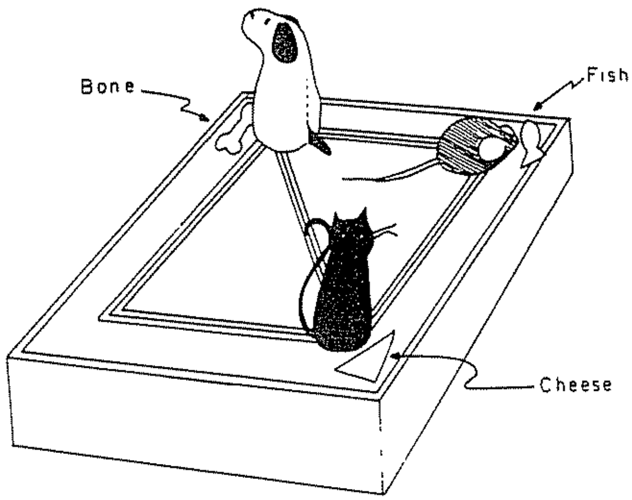

Figure 7.5 The apparatus for the Dog-Cat-Mouse problem. Each animal must be moved to its favorite food: the dog to the bone, the cat to the fish, and the mouse to the cheese. On each move, an animal must move all the way from one corner to another.

Preschoolers' Problem Solving on the Dog-Cat-Mouse Problem

In the second study to be described here, I investigated preschoolers' ability to solve problems in which both planning and means-ends analysis were difficult to apply (see Klahr, 1985, for details). By using a problem that precluded the use of subgoals, I was able to assess the extent to which children used some of the other weak methods listed above. More specifically, I sought evidence that they could use the "hill climbing" method.

The Dog-Cat-Mouse (DCM) puzzle consists of three toy animals (a dog, cat, and mouse) and three toy foods that "belong" to the animals (a bone, a fish, and a piece of cheese). The animals and the foods were arranged on the game board illustrated in Figure 7.5. The board had four grooves running parallel to each side of the square and a diagonal groove between the upper left and lower right corners of the square formed by the four outside grooves. The animals could move along the grooves, but they could not be removed from the board. The foods could be fastened to and unfastened from small patches of Velcro glued to each of the four corners. A problem consisted of an initial state—indicated by the placement of each animal in a corner of the puzzle, and a final state—indicated by some arrangement of the bone, fish, and cheese. The goal of the problem was to move each animal to its corresponding food.

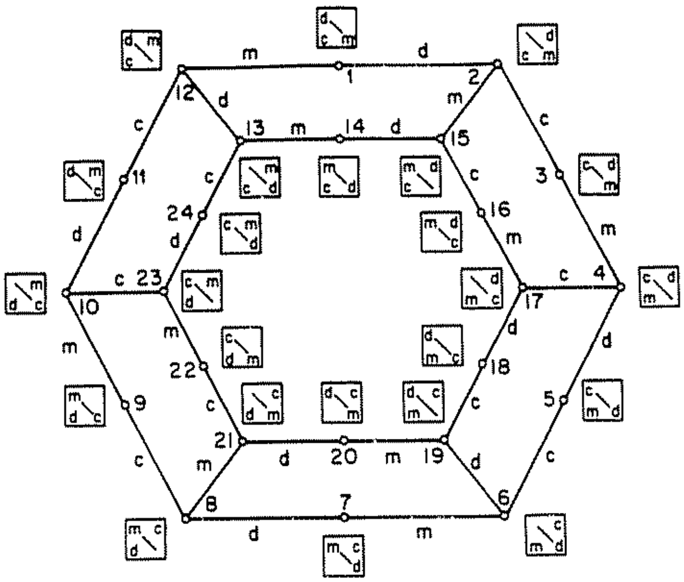

Figure 7 6 State space for the Dog-Cat-Mouse problem Each node represents a unique configuration of the three animals Each arc is labeled with the piece that is moved to change states

This puzzle was chosen for several reasons First, and most important, it has ambiguous subgoal ordering: the order in which the animals will reach their foods is not at all obvious. Second, it has easily remembered rules and a natural way to represent the goal state. Third, the puzzle has a sufficiently wide range of levels of difficulty. Finally, it is novel, so children are unlikely to have encountered similar puzzles.

The state space is illustrated in Figure 7 6. Each node represents one of the legal configurations of the three animals at the corners of the game board. Each arc label corresponds to the animal that was moved to get from one state to its neighbor For example, traversing the state space between nodes 1 and 2 requires a move of the dog. All problems are defined in terms of their initial states (determined by the arrangement of the animals) and their final states (determined by the arrangement of the foods). For example, the problem shown in Figure 7.5 starts at node 11 and ends at node 20

Several properties of the state space are relevant to our subsequent discussion:

Path length The minimum number of moves required to get from the initial state to the final state. Problems vary in path length from one move to seven moves. (Example: 1–20 has a path length of seven.)

Problem type. Some problems (*rotation* problems) do not change the cyclical ordering of the three objects, while other problems (*permutation* problems) require a change in the ordering by using the diagonal link on the problem board (also represented as a change from the outer to the inner hexagon in Figure 7.6) Permutation problems start and end with different permutations of the three animals (e.g., D-C-M-D . . . versus D-M-C-D) (Examples: 1–15, 22–3). Permutation problems generally have several minimum paths. For example, the minimum path from node 1 to node 19 could cross from the outer to the inner loop at nodes 2, 4, or 6.

The effect of problems that differ along these attributes depends on the processes subjects use to solve them. If they use a lot of forward search, then on average longer problems should be more difficult, and for equal-length problems, those starting with open nodes (three possible first moves) should on average be more difficult than those with closed nodes (two possible first moves). Permutation problems should be easier for two reasons: first, they usually have several minimum paths, and if subjects are moving randomly they are more likely to find one; second, if subjects *are* able to formulate subgoals, then a very useful one would be to fix the permutation (use the diagonal) and then rotate to the goal.

Problems, Subjects, and Procedure

Eight problems varying in path length (from four to seven) and problem type (permutation or rotation) were used. They are listed in the bottom section of Table 7.2. In addition, four three-move training problems were used to familiarize the children with the rules of the game. They are shown at the top of Table 7.2. Forty children from the Carnegie Mellon University Children's School participated (mean = 4 years, 10 months, *SD* = 6.3 months). Children were tested individually in a small playroom, adjacent to their regular classrooms, that was equipped with videotape recording facilities. After being brought into the room, the children were presented with the DCM puzzle in the context of the following cover story.

> This is a game about three hungry animals, and your job in the game will be to make sure that each animal gets its favorite food. I have a dog here who loves to chew on bones—would you please give the dog his bone? I have a cat who loves to eat fish, and I have a mouse who loves cheese [subject distributes food]. In this game I will mix up the animals and the food, and you will have to move each animal to its favorite food. There are three important rules about how you can move the animals:
>
> First of all, the animals always sit in the corners next to these circles. They can move along these blue lines—around the outside or up the middle, backward or forward—but they always have to stop in a corner by a circle. That means they can never stop in the middle of a line like this.

TABLE 7.2 Problem Sets, Structural Variables, Subject Performance, and Model Performance for Dog-Cat-Mouse Puzzle

Structural Variables					Performance Variables	
Problem number	Initial state[a]	Goal state[a]	Path length	Problem type	Problem difficulty[b]	Model's performance[c]
Training set						
T1	1	4	3	Rotation		
T2	7	22	3	Permutation		
T3	12	9	3	Rotation		
T4	2	17	3	Permutation		
Problem Set						
1	17	21	4	Rotation	.640	544
2	18	8	4	Permutation	950	.869
3	11	20	5	Permutation	.436	.510
4	10	5	5	Rotation	179	278
5	13	19	6	Rotation	184	399
6	24	18	6	Rotation	.263	352
7	14	7	7	Permutation	436	740
8	15	8	7	Permutation	.590	.800

[a]See Figure 7 6
[b]Proportion of children finding minimum path by second attempt
[c]Probability of model's finding minimum path within two attempts

The second rule is that only one animal can be in a corner at a time This is because my mouse is afraid of my cat, my cat is afraid of my dog, and believe it or not, this big dog is afraid of mice So they never sit together in one place, and you must never move an animal into a corner where another one is already sitting

The third rule is easy to remember—they always move one at a time While the dog moves the cat and the mouse wait, and while the mouse moves the dog and the cat wait Let's start with a couple of easy ones, and then they will get harder (Children were not explicitly instructed to minimize the number of moves; nevertheless here, as in many other studies, they appear to spontaneously attempt to produce efficient solutions)

Problems were presented in the order shown in Table 7.2. Children were given two chances to produce a minimum path solution to each problem. If a problem was solved in the minimum number of moves, then the next problem in the sequence was presented. If it was solved in more than the minimum number, or if it had not been solved after twice the minimum number of moves had been made, or if the subject gave up, then the same problem was presented a second time. Regardless of whether the second trial produced the minimum path, a longer solution path, or no solution, the next problem in the sequence was then presented.

As each problem was presented, the children were reminded to rearrange the animals so that each animal would get its favorite food. The children were allowed to make their own moves; if they attempted an illegal move, they were reminded of the rules.[4]

Scoring and Results

For each problem, subjects were assigned a 1/0 score based on whether they found a minimum path solution by the second presentation. Each subject was assigned a score based on the proportion of passes across the eight problems. Each problem was assigned a score based on the proportion of subjects passing it.

Relative subject performance. Subjects' performance varied widely: the highest-performing subject solved all but one problem, and three subjects failed all but one. Problem difficulty also varied widely, from nearly all subjects' passing the easiest problem to over 80% failing the hardest. The top-ranked subjects tended to fail only the harder problems, and the lowest-scoring subjects passed the easier rather than the harder problems. Although ages were fairly uniformly distributed between 50 and 65 months, age was not correlated with proportion correct.

Relative problem difficulty. The mean problem difficulty (defined as the proportion of subjects finding the minimum path) is shown in Table 7.2. Path length was a poor predictor of problem difficulty. The two easiest problems (1 and 2) were also the shortest, but even though they both have a path length of four, there was a 30% difference in the proportion of subjects passing them. The two next easiest problems (7 and 8) were the longest (seven moves). The four hardest problems were intermediate in path length, and within that set there was a large difference between the pairs with the same path length. Neither path length nor problem type was significantly correlated with problem difficulty.

Both path length and problem type are *structural variables*: features of the problem rather than of the problem-solving process. Even if they are good predictors of difficulty, they leave unstated the underlying processes they affect. But structural variables alone do not cause behavior directly: they are mediated by underlying processes. In situations of even modest complexity, such as the DCM puzzle, there are several plausible processes—or components of weak methods—and their interactions can

4 The most common illegal moves were moving an animal only halfway between two corners, moving two animals to the same corner, or attempting to rearrange the foods rather than the animals. However, illegalities occurred on fewer than 5% of trials and tended to occur only on the training problems.

best be understood by formulating an explicit process model. In the next section I consider some models that might account for these results.

Weak Methods on the DCM Puzzle

State evaluation. Each of the weak methods described earlier has the ability to evaluate the quality of a proposed move. That is, each can look to the future, but they differ in how much information they extract from it and how they use that information about the future to guide their behavior in the present. In generate and test, the evaluation is binary: a state either matches the goal or it does not. In contrast, the hill climbing method uses an evaluation function that gives some measure of how well the current state matches the goal state. For the DCM puzzle, consider an evaluation function—$EV(x, y)$—that computes how many of the pieces in state x are in the same positions in state y. For example, $EV(1, 7) = 0$ because none of the pieces are in matching positions, whereas $EV(24, 5) = 2$ because both the cat and the dog are positioned the same way in the two states (see Figure 7.6).

If children used such an evaluation function, then we would see two kinds of biases in their move patterns: one bias would show up as a tendency to favor moves that increase the function over those that leave it unchanged. For example, in problem 2 (18 → 8) a first move of the cat increases the evaluation function, while moving the dog does not. (A cat move also stays on the minimum path, while a dog move does not.) Over all trials and all subjects, on this problem, the cat was moved 81% of the time. Even more revealing are the "garden path" problems, in which the evaluation function produces a local improvement for moves off the minimum path. In problem 4 (10 → 5), the minimum path move is the mouse, which does not increase the evaluation function. In contrast, moving the cat does increase the partial evaluation function, and it was preferred on 66% of the trials even though it is off the minimum path. Similarly, on problem 5 (13 → 19), the nonminimum path move of the dog was preferred on 61% of the trials.

Another bias would be to prefer moves that leave the evaluation function unchanged over those that reduce it. For example, on problem 3 (11 → 20) the minimum path sequence requires that the dog be temporarily removed from its goal position even though this reduces the evaluation function. On 65% of all trials with problem 3, subjects preferred to move the cat rather than the dog even though this took them off the minimum path.

To determine whether subjects were using hill climbing on this problem, for each subject, I computed an *evaluation sensitivity score*: the proportion of trials on which, if such an evaluation function preferred one move to another, then the subject chose (one of) the preferred alternative(s). All subjects showed a sensitivity to partial evaluation. Evaluation sensitivity scores ranged from .60 to .90 (mean = .69, *SD* = .05). As noted, this sensitivity to local evaluation is not necessarily beneficial, for on garden path problems it moves subjects away from the minimum path. Indeed, evaluation sensitivity scores were *negatively* correlated with overall performance, suggesting that on this set of DCM problems excessive reliance on hill climbing was dysfunctional.

Goal detection. Instead of a multivalue evaluation function, a problem solver could use a simple binary evaluation in conjunction with the capacity to search n moves ahead for the goal. Then we should see perfect performance (no deviations from a minimum path) from n steps away. To assess how far ahead each subject could "see," I computed a *goal detection score* based on the distance from the goal reached directly 100% of the time. For example, if a subject produced minimum path solutions every time he or she was two moves from the goal state but on only 85% of the occasions from three moves away, then the subject would get a goal detection score of 2. Four subjects had goal detection scores of 0, nine had scores of 1, 11 scores of 2, 13 scores of 3, and 2 scores of 4. Overall, two-thirds of the subjects could stay on the minimum path when they were no more than two moves distant from the goal, and one-third could do it even from three moves away.

Strategic analysis. Because subjects appeared to be using a combination of methods, I attempted to capture their behavior by formulating a simple model of how they might approach problems—such as the DCM—in which subgoal ordering is ambiguous. The model has three parameters whose values were empirically determined. First I will describe the model, and then I will justify the parameter settings. The model makes each move according to the following rules:

> If there is an n-move sequence that can reach the goal state, then make it, otherwise:
>
> Generate all candidate moves On all but p_1% of trials, delete the piece just moved from the candidate set (e.g., backup is allowed with probability p_1).
>
> If there is more than one candidate, then compute *EV* between each candidate node and the goal node. Choose the move with the maximum *EV* on p_2% of trials. On $(1 - p_2\%)$ of the trials, or if the *EV*s for all candidate moves are equal, choose among them randomly

The model is an imperfect hill climber The imperfections are that (a) the model occasionally backs up, (b) it moves directly to the goal when

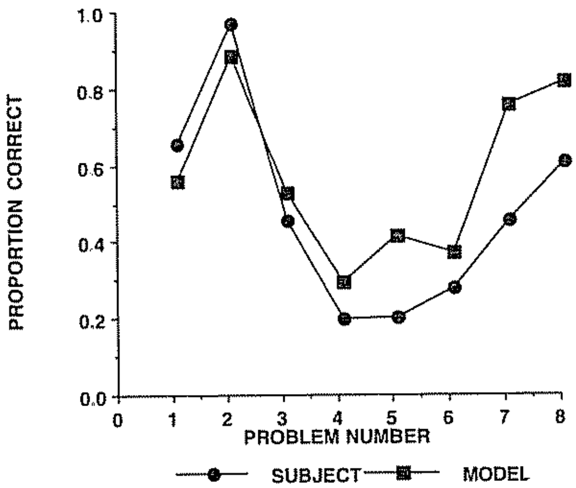

Figure 7.7 Actual versus predicted problem difficulty. Circles show proportion of all subjects passing by second trial. Squares show proportion of 400 cases that model passed by second trial.

it is near, and (c) it does not always choose the move having the maximum *EV*. The imperfections are captured in the model's parameters: n is the depth of goal detection, p_1 is the backup probability, and p_2 is the probability of being affected by *EV*. The values for these parameters were empirically derived: n was set to 2, based on the goal detection analysis, p_1 was set to .10, because 10% of all moves were double moves, and p_2 was set to .69.[5]

The model was implemented as a computer program, and each problem was presented to the program 400 times. Each solution path was scored as a 1 or 0 by the same criteria used for subjects' performance. Then the proportion of minimum path solutions (out of the 400) was computed, and this was converted to a probability of solution by the second attempt. The results are shown in the final column in Table 7.2 and Figure 7.7. The model accounts for over 70% of the variance in problem difficulty.

Summary: Future Orientation in Preschoolers' Use of Hill Climbing

Presenting preschoolers with problems having ambiguous subgoal ordering revealed what weak methods they could invoke when means-ends analysis was not useful. One extreme possibility was that they would resort to random trial and error. The other was that they would use a more appropriate weak method. The results of this study support the latter alternative.

5 This equals the mean of the partial evaluation sensitivities described earlier.

The composite model described above embodies two kinds of future orientation. When the short-term future (in this instance two moves away) is certain, fully controlled, and noncontingent, the model moves directly to the goal. Otherwise it computes the difference between the desired future and the locally available options and chooses the best of those. This sensitivity to incremental progress could degrade children's performance (as in the garden path problems). Nevertheless, it is reasonable for children to attempt to use such information.

THINKING AHEAD IN SCIENTIFIC DISCOVERY: DEVELOPMENTAL DIFFERENCES

The two problem-solving studies described so far use arbitrary tasks designed to minimize the influence of children's prior knowledge about the world. At the other extreme of both complexity and the relevance of prior knowledge lies the process of scientific discovery. Although it can be viewed as a type of problem solving, scientific discovery has several features that distinguish it from the simpler types of problems described earlier.

> First, the search process takes place in two spaces: a space of hypotheses about the domain under investigation and a space of experiments in the domain. Scientific discovery requires what we have called a "dual search" in these two spaces (Klahr & Dunbar, 1988)
>
> Second, the goal state is ill defined and complex. At the outset the goal is simply "understand the domain," and the constraints and parameters of that understanding are developed during the discovery process itself. The complexity derives from the fact that the goal state is not a static configuration like those used in TOH or DCM, but a state or rule of nature that can only be inferred from the behavior of a complex, dynamic system.
>
> Third, there is mutuality between the states and operators in each space. Moves from one hypothesis to another in the hypothesis space are effected by "applying" experimental operators and interpreting the results of experimental outcomes. Moves from one experiment to the next in the experiment space are effected by attempts to evaluate the current hypothesis.
>
> Fourth, prior knowledge plays an influential role in scientific discovery, because subjects always come to the task with potentially relevant knowledge about the domain. This prior knowledge influences the kinds of hypotheses that are generated, the strength with which they are held, and the experiments that are conducted to evaluate them.

These considerations led us to design a laboratory investigation of the scientific discovery process. Given this goal, we faced the problem of searching our own experiment space. In doing this search, we imposed some constraints on the kind of task we would use: (a) We wanted a domain in which prior knowledge could influence both the initial hypotheses that subjects might propose and the strength with which they held them. (b) We wanted to allow subjects to design and evaluate their own experiments rather than choose among a set of predetermined alternatives. (c) We wanted a domain in which the mapping between experimental outcomes and hypotheses was not trivial. (d) We did not want to play God with the subjects by telling them whether they had discovered a true hypothesis. Instead, we wanted subjects to decide for themselves when to terminate experimentation. (e) Finally, we wanted a task that was interesting and challenging for a wide range of ages.

Future Orientation in Discovery Microworlds

Scientific discovery requires subjects to think about the future in three senses: (a) they need to be future oriented in their specific predictions about the outcome of the next experiment; (b) they need to consider the future unfolding of the planned sequence of experiments that will be used to evaluate currently held hypotheses; (c) they need to think about the future of their own knowledge states and how they might be changed by the results of their experiments. The studies to be described below investigate the extent to which children and adults can think about the future in these senses.

Laboratory Simulation of Scientific Discovery: BigTrak

The device we used is a computer-controlled toy robot tank called BigTrak. It is a battery-operated, programmable, self-contained vehicle approximately 13″ × 5″ × 8″. The BigTrak keypad interface is depicted in Figure 7.8.[6] The basic execution cycle involves first clearing the memory with the CLR key and then entering a series of up to 16 instructions, each consisting of a function key (the command) and a one- or two-digit number (the argument). The five command keys are: ↑ (move forward), ↓ (move backward), ← (turn left), → (turn right), and FIRE. When the GO key is pressed BigTrak executes the program. For example, suppose you pressed the following series of keys:

CLR ↑ 5 ← 7 ↑ 3 → 15 FIRE 2 ↓ 8 GO.

When the GO key was pressed, BigTrak would move forward five feet, rotate counterclockwise 42° (corresponding to seven minutes on an ordinary clockface), move forward three feet, rotate clockwise 90°, fire (its

6. Figure 7.8 actually shows the keypad from the BT microworld (to be described shortly) based on the BigTrak toy.

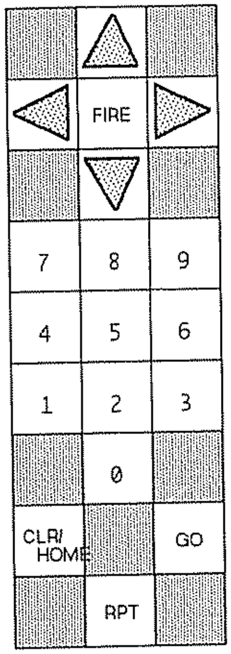

Figure 7 8 Keypad used in BigTrak and BT studies See text for details

"laser cannon") twice, and back up eight feet. Our precedure had three phases.

1. Subjects were introduced to BigTrak and instructed on the use of each basic command. They were also instructed in how to generate verbal protocols. During this phase the RPT key was not visible. Subjects were trained to criterion on how to write a series of commands to accomplish a specified maneuver. This phase corresponded to a scientist's having a basic amount of knowledge about a domain but not understanding all its ramifications.

2. Subjects were shown the RPT key. They were told that it required a numeric parameter (*N*), and that there could be only one RPT *N* in a program. They were told that their task was to find out how RPT worked by writing programs and observing the results. This corresponded to a new problem in the domain: an unresolved question in an otherwise familiar context.

3. Subjects could formulate hypotheses about RPT and run experiments to test them. This required decisions about hypotheses and decisions about experiments. Subjects were never told whether they had discovered how RPT worked. They had to decide when to terminate search.

In one of our studies (Dunbar & Klahr, 1989) we used two groups of subjects: Carnegie Mellon University undergraduates and children

TABLE 7.3 Common Hypotheses (in Decreasing Order of "Popularity" or "Plausibility")

'RPT N' tells BigTrak to	Role of N
1 Repeat the entire program N times	Counter
2 Repeat the last step N times	Counter
3 Repeat the subsequent steps N times	Counter
4 Repeat the entire program once	Nil
5 Repeat the last N steps once	Selector
6 Repeat the Nth step once	Selector
7 Repeat the first N steps once	Selector
8 Repeat the entire program $f(N)$ times	Counter

between the ages of 8 and 11 years. Before I describe their performance, here is how the RPT key works. It takes the N instructions preceding the RPT N instruction, and it repeats that sequence one more time. Because this rule is somewhat counterintuitive to the common interpretation of "repeat," it was not easy to discover.

Results: BigTrak Study

Only 2 of 22 children were successful, although 12 of the unsuccessful children were sure they had discovered the correct rule and terminated their experimentation quite satisfied with their discovery. In contrast, nearly all the adults discovered the correct rule, but it was not a trivial task for them. In fact, with respect to average time, number of hypotheses, and number of experiments, the adults were not very different from the children. The explanation for these vastly different success rates must lie at a deeper level. We need to look more closely at the nature of the hypothesis space and the experiment space.

Subjects generated a variety of hypotheses during their experimental phase. The more common hypotheses are listed in Table 7.3 in order of decreasing popularity or plausibility. (Recall that the correct rule is actually number 5.) Hypotheses are classified according to the role they assign to the parameter that goes with the RPT command. In hypotheses 1, 2, 3, and 8, N *counts* the number of repetitions. We call these "counter" hypotheses. In hypotheses 5, 6, and 7, N determines which segment of the program will be *selected* to be repeated. We call these "selector" hypotheses. This distinction between counters and selectors turns out to be very useful in our subsequent experiments. Search in the BigTrak hypothesis space can involve local search among counters or among selectors, or it can involve more far-ranging search between counters and selectors.

How can we characterize the BigTrak experiment space? At one extreme it is enormous: for example, counting only commands, but not

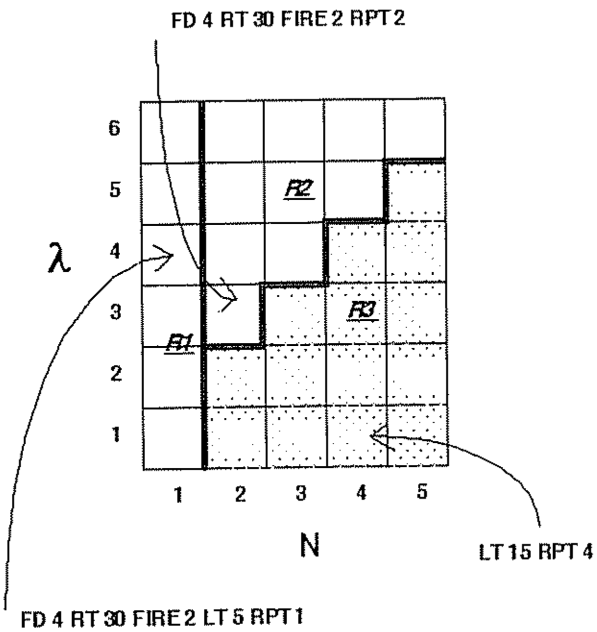

Figure 7.9 Part of the experiment space for the BigTrak studies. Each cell corresponds to a particular combination of program length (λ) and argument for the RPT key (N) Sample programs are shown for three cells: (4,1) (3,2) and (1,4).

their numerical arguments, as distinct, there are over 5^{15} distinct programs that subjects could write. However, we have found that we can adequately characterize the experiment space in terms of just two parameters. The first is λ, the length of the program preceding the RPT. The second is the value of N, the argument that RPT takes. Because both parameters must have values less than 16, there are 225 "cells" in the $\lambda - N$ space. Within that space, we identify three distinct regions: region 1 includes all programs with $N = 1$; region 2 includes all programs in which $1 < N < \lambda$; region 3 includes all programs in which $N \geq \lambda$. A segment of the experiment space, showing the different regions, is depicted in Figure 7.9, together with illustrative programs from each region. Programs from different regions of the experiment space vary in their effectiveness. Note that programs from region 2, where there are more steps than the value of N, are particularly informative.

This analysis of the hypothesis space and the experiment space enabled us to discover a couple of interesting things about how subjects approached this task. By examining the pattern of experiments, we could determine how much of the experiment space subjects searched, and by analyzing their verbal protocols we could classify experiments in terms of where they were in the hypothesis space at the time of each experiment.

We found there were two distinct types of subjects, with fundamen-

TABLE 7.4 Differences between Theorists and Experimenters in BigTrak Study

	THEORISTS	EXPERIMENTERS
Defining property	State selector frame *without* sufficient evidence	State selector frame only *after* sufficient evidence
Time (minutes)	11	25
Total experiments	9.3	18.4
Experiments without hypotheses	0.8	6.1
Comments about experiment space	5.9	0.9
Experiment space cells used	5.7	9.9

tally different strategies. We distinguished the two groups by how much information they had when they changed from a counter frame to a selector frame. If they made the switch without having seen the result of a region 2 experiment, then we called them theorists, because they could not have based their decision on conclusive experimental evidence. On the other hand, if they made the switch from counters to selectors only after running region 2 experiments, we called them experimenters. (This analysis makes sense only for the adults, since so few children discovered the correct rule.)

The two strategies were accompanied by other differences, as shown in Table 7.4. Experimenters took twice as long to discover how RPT worked; they explored much more of the experiment space; and they conducted many more experiments without any active hypothesis. That is, they ran experiments in order to generate a data pattern over which they could induce a frame.

Development of Search Constraints in the Experiment Space

This tendency to suspend the hypothesis-testing mode while attempting to discover some kind of regularity in the data is very common and an extremely important aspect of scientific reasoning. In terms of future-oriented processes, subjects realized they had little control over the outcome of an experiment because they were unable to produce any hypotheses at this point. By switching from hypothesis space search to experiment space search, they were attempting to discover some regularities that would enable them to gain control and certainty over their predictions. This behavior suggested that we needed to find out a lot more about how subjects searched the experiment space and about how different goals might influence that search.

To study these issues, we decided to use the BigTrak paradigm in such a way that we could focus on developmental differences in the heuristics used to constrain search in the experiment space. We knew that subjects at all ages shared *domain-specific* knowledge that biased them in the same direction with respect to the plausibility of different hypotheses. We expected both age and scientific training to reveal differences in the *domain-general* heuristics used to constrain search in the experiment space.

One consequence of domain-specific knowledge is that some hypotheses about the domain are more plausible than others. We explored the effect of domain-specific knowledge by manipulating the role of plausible and implausible hypotheses. Our goal was to investigate the extent to which prior knowledge—as manifested in hypothesis plausibility—influenced how people designed experiments and how they interpreted the results of those experiments.

For this study (Klahr, Fay & Dunbar, 1993) we moved from the original BigTrak toy to a computer microworld called BT, in which a simulated "spaceship" moved around on a computer screen according to instructions entered on a BT keypad (also displayed on the screen).

The study had three phases. The first and third phases were the same as in the previous study. Subjects learned about all the normal keys and were trained to criterion on getting BT to move around the screen. In the second phase, the RPT key was introduced as before. Subjects were told that their task was to find out how RPT worked by writing at least three programs and observing the results. But then we changed the procedure a bit, by suggesting one way that RPT might work. The experimenter said:

> "One way that RPT might work is": [and then we stated one of four hypotheses listed below]. Then we continued with the instructions: "Write down three good programs that will allow you to see if the repeat key really does work this way. . . ."

When subjects had written, run, and evaluated three experiments, they were given the option of either terminating or writing additional experiments if they were still uncertain about how RPT worked. The entire session lasted approximately 45 minutes.

Throughout the study, we used only four rules for BT. Recall that our earlier studies with adults and grade-school children revealed two very "popular" hypotheses about the effect of RPT N in a program. The two popular, or plausible hypotheses were counters:

A: Repeat the entire program N times,
B: Repeat the last step N times.

TABLE 7.5 Design of BT Experiment: Specific Hypotheses for Each Given-Actual Condition

	ACTUAL RULE	
Given hypothesis	Counter	Selector
Counter	B: Repeat last step N times ↓ A: Repeat entire program N times *Theory refinement*	A: Repeat entire program N times ↓ D: Repeat last N steps once *Theory replacement*
Selector	D: Repeat last N steps once ↓ A: Repeat entire program N times *Theory replacement*	C: Repeat step N once ↓ D: Repeat last N steps once *Theory refinement*

In contrast, there were two hypotheses (selectors) that subjects were unlikely to propose:

C: Repeat the Nth step once,
D: Repeat the last N steps once.

We provided each subject with an initial hypothesis about how RPT might work. *The Given hypothesis was always wrong*, but depending on the condition, subjects regarded it as either plausible or implausible. BT was always set to work according to some other rule. We called that the Actual rule. Both the Given and Actual could be either plausible or implausible. In some conditions the Given hypothesis was only "somewhat" wrong, in that it was from the same frame as the way RPT actually worked. In other conditions the Given was "very" wrong, in that it came from a different frame than the Actual rule.

The BT simulator was programmed so that each subject worked with a RPT command obeying one of the two counter rules or two selector rules described above. We used a between-subjects design, depicted in Table 7.5. The Given hypothesis is the one suggested by the experimenter, and the Actual rule is the way BT was programmed to work for a particular condition. Remember, the key feature of this design is that *RPT never worked in the way that was suggested.*

Changing from a hypothesis within a frame to another hypothesis from the *same* frame (from one counter to another counter) requires only a single slot value change. In our microworld, this corresponds to *theory refinement*. In contrast, changing from a hypothesis from one frame to another hypothesis from a *different* frame (from a counter to a selector) requires a simultaneous change in more than one attribute, because the values of some attributes are linked to the values of others. This corresponds to *theory replacement*.

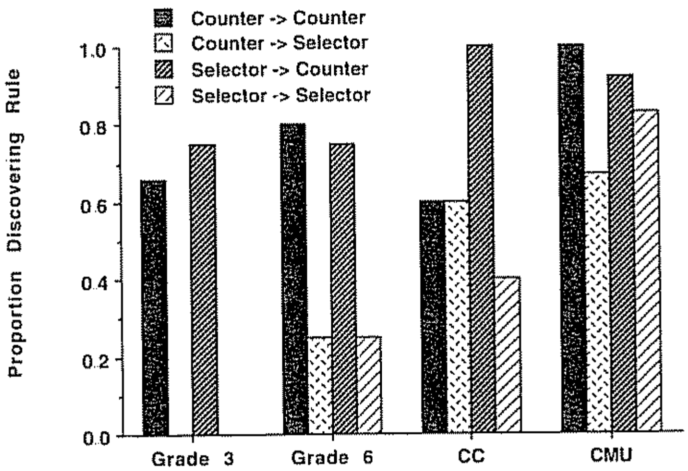

Figure 7 10 Overall success rates: proportion of subjects discovering the rule as a function of grade level and Given-Actual condition

Subjects

We used four groups of subjects: Carnegie Mellon University (CMU) undergraduates, community college (CC) students, "sixth" graders (a mixed class of fifth to seventh graders, mean age 11 years), and third graders (mean age 9 years).

CMUs were mainly science or engineering majors, whereas the CCs had little training in mathematics or physical sciences. Children came primarily from academic and professional families. Most of the third graders had about six months of LOGO instruction (all had at least one month of LOGO). Note that CCs had less programming experience than the third graders.

Results of the BT Microworld Study

As we expected, domain-specific knowledge—as manifested in expectations about what "repeat" might mean in this context—played an important role. Regardless of what the Given hypothesis was, subjects found it easier to discover counters (81%) than selectors (35%). There was also a main effect for group: the correct rule was discovered by 83% of the CMUs, 65% of the CCs, 53% of the sixth graders, and 33% of the third graders. This group effect is attributable to the Actual = selector conditions, in which 56% of the adults but only 13% of the children were successful. In fact, none of the third graders discovered selectors. For counters, adults and children were not so different in their success rates (88% versus 75%).

What about subjects' reactions to the Given hypothesis? Recall that we presented subjects with either plausible or implausible hypotheses

TABLE 7.6 Subjects' Responses to the Given Hypothesis

	ADULTS		CHILDREN	
Category	Counter	Selector	Counter	Selector
1. Accept Given	.70	.60	.71	.33
2. Accept Given *and* propose alternative	.30	.40	.06	.06
3. Reject Given, propose alternative	0	0	.23	.61

in order to determine the extent to which search in the hypothesis space was influenced by plausibility. This is one of the points at which *domain-specific* knowledge (which determines plausibility) might affect *domain-general* knowledge about experimental strategies.

Before running the first experiment, subjects were asked to predict what would happen. Their predictions indicated how well they understood and accepted the Given hypothesis. Each subject's response to the Given hypothesis was assigned to one of three categories: 1, accept the Given hypothesis; 2, accept the Given, but also propose an alternative; 3, reject the Given and propose an alternative.

The proportion of subjects in each category is shown in Table 7.6 as a function of grade level and type of Given hypothesis. In both conditions the adults always accepted the Given hypothesis, either on its own (category 1) or in conjunction with an alternative that they proposed (category 2). Adults never rejected the Given hypothesis. In contrast, no third grader and only two sixth graders ever proposed an alternative to compare with the Given (category 2). Instead, children considered only one hypothesis at a time. When given counters they mainly accepted them, but when given selectors they mainly rejected them and proposed an alternative, which was usually a counter of their own design.

This propensity to consider multiple versus single hypotheses affected the type of experimental goals the subjects set. These goals, in turn, were used to impose constraints on search in the experiment space. We looked at these goals more closely by analyzing both what subjects said about experiments and the features of the experiments they actually wrote.

Subjects' verbal protocols contained many statements indicating both explicit understanding of the experiment space dimensions and a general notion of "good instrumentation": designing interpretable programs containing easily identifiable markers. Subjects made explicit statements about both kinds of knowledge. Here are some typical adult statements:

TABLE 7.7 Proportion of Self-Generated Constraints

	CMU	CC	Sixth grade	Third grade
Explicit $\lambda - N$ comments	83	60	53	20
Standard turn units	.92	.95	.71	53
Small arguments	92	85	65	47
Proportion of programs in small experiment-space region	.50	63	.26	31

I don't want to have two of the same move in there yet, I might not be able to tell if it was repeating the first one or if it was doing the next part of my sequence;

I'm going to use a series of commands that will . . . *that are easily distinguished from one another,* and won't run it off the screen;

so I'm going to pick two [commands] that are the direct opposite of each other, to see if they don't really have to be direct opposites but I'm just going to write a program that consists of two steps, *that I could see easily* (emphasis added)

Sixth graders were somewhat less articulate but still showed a concern for both experiment space dimensions and program interpretability. In contrast, third graders rarely made such comments. The proportion of subjects making such comments is shown in the top row of Table 7.7.

At a finer level of detail, good instrumentation was assessed by how well subjects observed three pragmatic constraints: (a) using standard units of rotation, such as 15 or 30 "minutes" (90° and 180°), for turn commands; (b) using small numeric arguments (values <5) on move commands, so that the actions of BT are not distorted by having it hit the boundaries of the screen; and (c) using distinct commands in a program where possible. Programs constrained in these ways produce behavior that is easier to observe, encode, and remember. For both turns and moves, there was a strong effect of grade level.

Yet another interesting difference between the children and the adults was the way adults limited their search to a small "corner" of the experiment space. We looked at the section of the experiment space with λ between 1 and 4, and N between 1 and 3. This corresponds to only 5% of the full experiment space. But we discovered that over half of the adults' experiments occurred within this small area. On the other hand, children's experiments were much more scattered throughout the space. Both what subjects *said* and what they *did* support the conclusion that older subjects—even those with weak technical backgrounds—were better able than children to constrain their search in the experiment space and to design interpretable experiments.

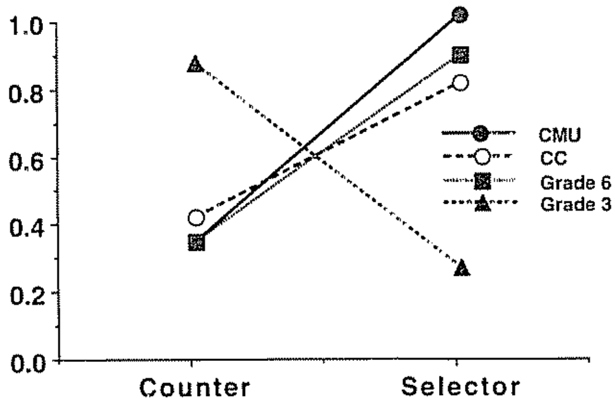

Figure 7.11 Proportion of subjects with $\lambda > N$ on first experiment, by grade level and Given condition.

So what were subjects trying to do here? What were their experimental goals? How can we infer these goals from the kinds of experiments they ran? We reasoned as follows: If the experimental goal is to *determine* which of the program steps are repeated for selector hypotheses, or to discriminate between selectors and counters, then subjects should write programs having more than N steps (i.e., with $\lambda > N$). (In programs where λ is several steps greater than N, it is easy to distinguish among repeats of all steps, first step, last step, and N steps.) On the other hand, if the goal is to *demonstrate* the effect of a counter, then subjects should use larger values of N and (for pragmatic reasons) relatively short programs (programs with $\lambda \leq N$). Figure 7.11 shows the proportion of subjects in each condition whose first programs had $\lambda > N$. Both of the adult groups' responses and sixth graders' responses were consistent with the normative account I just gave. Third graders showed the opposite pattern. More detailed results can be found in Klahr, Fay, and Dunbar (1993).

Heuristics for Constraining Search

We believe that these patterns of constrained search in the experiment space result from a set of domain-general heuristics that are differentially available to children and adults. Based on the present study, we have proposed the following four heuristics:

1. *Use the plausibility of a hypothesis to choose experimental strategy* In this study we found that both children and adults varied their approach to confirmation and disconfirmation according to the plausibility of the currently held hypothesis. When hypotheses were plausible, subjects at all levels tended to set an experimental goal of *demonstrating* key features of the given hypothesis rather than conducting experiments that could discriminate between rival hypotheses.

For implausible hypotheses, adults and young children used different strategies. Adults' response to implausibility was to propose hypotheses from frames other than the Given frame and to conduct experiments that could *discriminate* between them. Our youngest children's response was to propose a hypothesis from a different, but plausible, frame and then to ignore the initial, and implausible, hypothesis while attempting to demonstrate the correctness of the plausible one. Third graders were particularly susceptible to this strategy.

2. *Focus on one dimension of an experiment or hypothesis*. An incremental, conservative approach has been found to be effective in both concept attainment and hypothesis testing. This suggests that in moving from one experiment or hypothesis to the next, or in moving between experiments and hypotheses, one should decide on the most important features of each and focus on just those features. Here the CMU adults stood apart from the other three groups. They were much more likely than any of the others to make conservative moves—that is, to minimize differences in program content between one program and the next.

3. *Maintain observability*. As BT moved along the screen it left no permanent record of its behavior. Subjects had to remember what BT actually did. Thus, one way to implement this heuristic is to write short programs. Adults almost always used it, whereas the youngest children often wrote programs that were very difficult to encode. This heuristic depends on knowledge of one's own information-processing limitations as well as knowledge of the device. Our finding that the third graders did not attempt to maintain observability, whereas the sixth graders and adults did, may be a manifestation, in the realm of experimental design, of the more general findings about the development of self-awareness of cognitive limitations (Wellman, 1990).

4. *Design experiments giving characteristic results*. This heuristic maximizes the interpretability of experimental outcomes. Physicians look for "markers" for diseases, and physicists design experiments in which suspected particles will leave "signatures." In the BT domain, this heuristic is instantiated as "use many distinct commands." On average, about half of all programs in each group did not contain any repeated commands, although because third graders were more likely to use long programs, they were more likely to use repeated commands, which reduced the possibility of generating characteristic behavior.

These, then, are the four heuristics our subjects used to constrain search in the experiment space. As I noted in describing each one, adults and children differed in their use. Adults not only appeared to use each of them, but also they seemed able to deal with their inherent contradictions. In contrast, children either failed to use these heuristics at all or else let one of them dominate. This is not simply a matter of children being unable to understand the *logic* of the discriminating experiment or the difference between testing hypotheses and generating effects. Indeed, Sodian, Zaitchik, and Carey (1991) have shown that even first graders understand these distinctions in simple, two-alternative situations. The problem for young children is how to constrain search in a very large space of hypotheses and experiments.

Thinking About the Future in Discovery Tasks

In these discovery tasks, the subjects' goal is to discover something about the world. The only way they can gather information about the current workings of that world is to enter a program and predict its outcome. The subjects evaluate their understanding about how RPT works by assessing the accuracy of that prediction. Thus, attempts to predict the future behavior of a complex system lead to the discovery of its underlying principles.

Recall the three senses in which scientific discovery invokes future-oriented thinking: (a) predictions about specific experiments, (b) planning an experimental sequence, and (c) understanding that experimental outcomes may, in the future, change current knowledge states. Adults were able to maintain this future orientation in all three senses. Children of middle-school age and beyond appeared to be able to deal with the first sense but had difficulty with the other two. For example, even our youngest children understood that their experiments would be played out in the future, and that their predictions were about things yet to come. However, children were less able than adults to understand that they would be engaged in a series of experiments and that the experimental series itself required an overall plan in which the outcome of one experiment was related to the outcomes of prior and subsequent experiments. Instead, our youngest children were satisfied with demonstrating that they could maintain their current hypotheses by obtaining a particular effect. Finally, although we have only indirect evidence from the studies reported here, it seems that adults were much more aware than children that their own knowledge states would be changed by the results of experimentation.

CONSIDERING THE FUTURE IN PROBLEM SOLVING AND SCIENTIFIC DISCOVERY

In this final section, I return to the framework (see Table 7.8) and apply it to the domains discussed earlier. The rows correspond to various situations in which an individual may have to think about the future, and the columns correspond to the attributes of the framework. The cell

TABLE 7.8 Framework for Future-Oriented Thinking Processes

Domain	Subject of thinking	Uncertainty	Control	Contingency	Abstraction	Social	Temporal extent
Game playing and problem solving	Move in chess or checkers	In principle, no; capacity limits, yes	Partial	Adversarial	No	No	minutes
	T-end problems in TOH	No	Full	None	Probably	No	seconds
	F-end problems in TOH	Yes	Full	None	No	No	seconds–minutes
	Solution path in DCM	Near goal, no; distant from goal, yes	Full	None	No	No	minutes
Scientific reasoning	BT discovery	Yes	Partial	Informational	Several levels	No	minutes–hours
	Collaborative discovery tasks	Yes	Partial	Supportive	Several levels	Yes	minutes–hours

entries indicate the value of the attribute for the situation (at least according to my best guess; this is, after all, a pretty informal process).

Future Orientation in Problem Solving

The entries at the top of Table 7.8 summarize the earlier analyses. For many problem-solving and game-playing situations, the formal properties of the task interact with the limited capacity of the problem solver in ways indicated earlier. For example, for the Dog-Cat-Mouse problem, the level of uncertainty depends on the distance remaining to the goal and the subject's depth of search capacity.

Future Orientation in Scientific Discovery

These attributes can also be considered with respect to the cognitive processes involved in scientific discovery. Scientific reasoning is inherently uncertain, since nature responds to "moves" in ways still unknown (else there would be no discovery to be made). The contingencies are benign, because the physical world does not attempt to deceive or "hide its secrets." In our studies, control over the design of an experiment is complete, but the outcome is not fully determined by the problem solver. The degree of abstraction varies widely and is one of the distinctions between effective and ineffective problem solvers in discovery tasks. Our most effective subjects suppressed much of the detail in experimental variation and thought only in terms of λ and N. Indeed, in most areas of real-world science, skilled performance rests in no small measure on the scientist's ability to work at just the right grain size, neither ignoring relevant detail nor being overwhelmed by irrelevancies. In the tasks reported here, there is no social aspect, but others have used the BT paradigm to study the effects of collaboration on scientific discovery (Teasley, 1992), and the real scientific enterprise has very influential social components.

Other Varieties of Future-Oriented Thinking

All the tasks discussed in this chapter, from the Tower of Hanoi and the Dog-Cat-Mouse to the BigTrak discovery tasks, involve problem solving. But there are other kinds of future-oriented processes that are not easily characterized in those terms. In these concluding paragraphs I will mention a few extreme departures from the set of tasks I have just described. I leave it as an exercise for the reader to extend the framework listed in Table 7.8 to accommodate these situations, as well as the many others described in this volume.

Consider first the kind of tasks presented to infants by Haith, Hazan, and Goodman (1988)[7] in their demonstration that 3.5-month-old infants developed expectations for patterns of alternating visual events. Because the infants were responding to events that were not contingent on the infants' behavior, Haith et al. characterized their future orientation in terms of "expectancies" and "anticipations" about perceptual events. In this most rudimentary form of future orienation, it appears that infants' default assumption is to expect the future to be pretty much like the past: static objects are expected to remain in place, moving objects are expected to continue along their trajectories, and simple systematic patterns of perceptual activity are expected to repeat indefinitely. Can this kind of future orientation be cast as a type of problem solving? Perhaps. Haith et al. imply that infants do have a goal in such situations: "to detect regularities in dynamic events and to develop expectations partly in order to bring their behavior under self-control" (1988, p. 477). Another example of discovering the present by predicting the future? Perhaps.

There are other variations. One can think about the future with no particular goal in mind: consider weather forecasting. One can construct representations for nonexistent states (as in planning) but still lack a future orientation (dreams, reminiscences, musings about missed opportunities and "the road not taken"). For example, consider the following: I am thinking about the feasibility of skiing tomorrow from the top of one mountain (A) to the bottom of a nearby mountain (B). I have a good memory of the network of trails, and I mentally work my way down the

7. These tasks are also described in Haith's chapter in this volume.

slopes, planning tomorrow's day. Now consider a different situation. As I reflect on today's skiing, I recall all the paths I took and think about whether I *could have gone* from the top of A to the bottom of B. What is the difference in these two mental processes? How does the future-oriented plan differ from the memory-oriented reflection? What are the processes and representations that differentiate these situations? In constructing models of the psychological processes involved in future-oriented thinking, it will be important to clarify these issues. I believe that we do not know how to do this at present. Perhaps we will, in the future.

REFERENCES

Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review, 86,* 124–140.

Dunbar, K., & Klahr, D. (1989). Developmental differences in scientific discovery strategies. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 109–143). Hillsdale, NJ: Erlbaum.

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

Friedman, W. J. (1990). *About time: Inventing the fourth dimension.* Cambridge: MIT Press.

Hadley, G. (1962). *Linear programming.* Reading, MA: Addison-Wesley.

Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development, 59,* 467–479.

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

Klahr, D. (1978). Goal formation, planning, and learning by pre-school problem solvers, or "My socks are in the dryer." In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 181–212). Hillsdale, NJ: Erlbaum.

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

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

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

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

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

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

Piaget, J. (1976). *The grasp of consciousness.* Cambridge: Harvard University Press.

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

Siegler, R. S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development, 46* (Whole No. 189).

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

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

Teasley, S. D. (1992, April). Children's collaborations and talk. Paper presented at symposium on Learning through peer collaboration: Conceptual and methodological issues (E. Forman, Chair), conducted at the annual meeting of the American Education Research Association, San Francisco.

Wellman, H. M. (1990). *The child's theory of mind.* Cambridge: MIT Press.

Wickelgren, W. A. (1974). *How to solve problems.* San Francisco: Freeman.