Pre-schooler's solution of problems with ambiguous sub-goals

David Klahr

Carnegie-Mellon University

27 February 1984

Abstract

Children (4 to 6 years old) were presented with problems requiring from 4 to 7 moves for solution. The problems were constructed such that the subgoal ordering was ambiguous. Children's performance was consistent with a generate and test strategy that had a 2-move lookahead for a goal state, a no-backup constraint, and some partial evaluation of progress toward the goal.

When faced with problems in unfamiliar domains, adults draw on a small repertoire of processes called "weak methods", including generate-and-test, heuristic search, means-ends analysis, hill-climbing, and planning (Newell, 1969; Laird & Newell, 1983a). The weak methods are usually inadequate and inefficient compared to knowledge-rich, problem-specific methods. Nevertheless, they are extremely general, and they often provide the only basis for intelligent action. Young children also use rudimentary forms of weak methods requiring the use of subgoals, such as means-ends analysis (Klahr, 1978; Klahr & Robinson, 1981, Spitz & Borys, 1984; Spitz, Webster, & Borys, 1982).

Klahr & Robinson (1981) presented pre-school children with two variants of the Tower of Hanoi (TOH). They found that performance declined when subgoal ordering was not self-evident. On the standard "tower-ending" problems, in which all the objects are stacked on a single peg, it is clear that the bottom-most object must get to the goal peg first, then the second from the bottom, and so on. This subgoal sequence is apparent even though the exact move sequence necessary to achieve it is not. On these problems, half of the 6-year-old subjects could solve 6-move problems, and even 5-year olds were able to solve 4-move problems most of the time.

On "flat-ending" problems, in which each peg has one object on it, the proportion of 5- and 6-year-olds who could reliably plan at least four moves ahead dropped from 81% to 40%. Flat-ending problems do not have an obvious order in which disks reach their goal pegs. When the surface form of the problem does not suggest an unambiguous ordering of subgoals, then children have a difficult time applying MEA. Instead, they must use an even weaker one of the weak methods.

This study further investigates how pre-school children behave when confronted with such ambiguous subgoal problems. We address the following questions:

* Do children move haphazardly when subgoals are ambiguous?

* Do children avoid unnecessary backup?

* Do children advance directly toward a goal once it becomes "visible"?

* Are children reluctant to move away from a goal temporarily in order to ultimately reach it?

* Are they easily led down "garden paths"?

The Dog-Cat-Mouse Puzzle

The Dog-Cat-Mouse (DCM) puzzle consists of three toy animals and three toy foods that "belong" to the animals (a bone, a fish and a piece of cheese) arranged on the game-board illustrated in Figure 1. The board has 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 can be moved along the groves, and the foods can be fastened to and unfastened from each of the four corners.

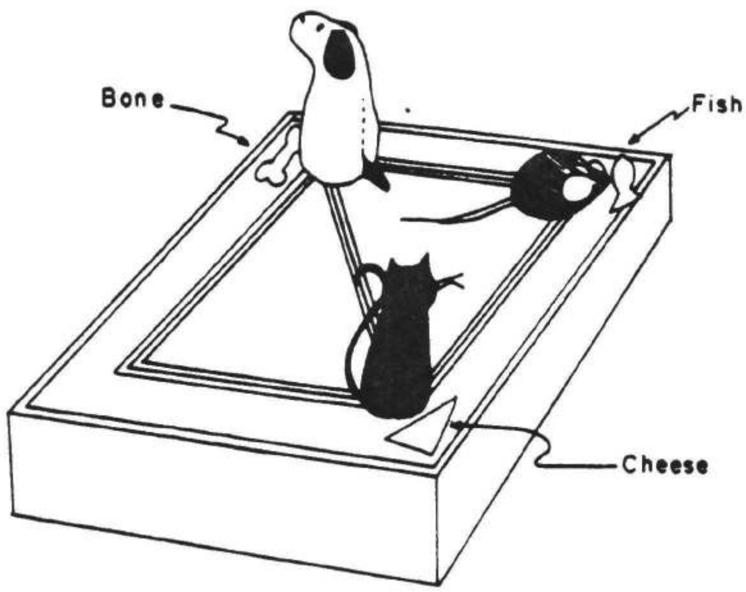

Figure 1 The apparatus for the Dog-Cat-Mouse problem. Each animal must be moved to its favorite food: the dog to the bone, the cat to the fish, and the mouse to the cheese.

A problem consists of an initial state -- indicated by some arrangement of the animals and a final state -- indicated by some arrangement of the foods.

Problem Set

The state space for the DCM puzzle is illustrated in Figure 2[1]. Each node represents a legal configuration. The label on each arc corresponds to the animal that was moved to get from one state to its neighbor.

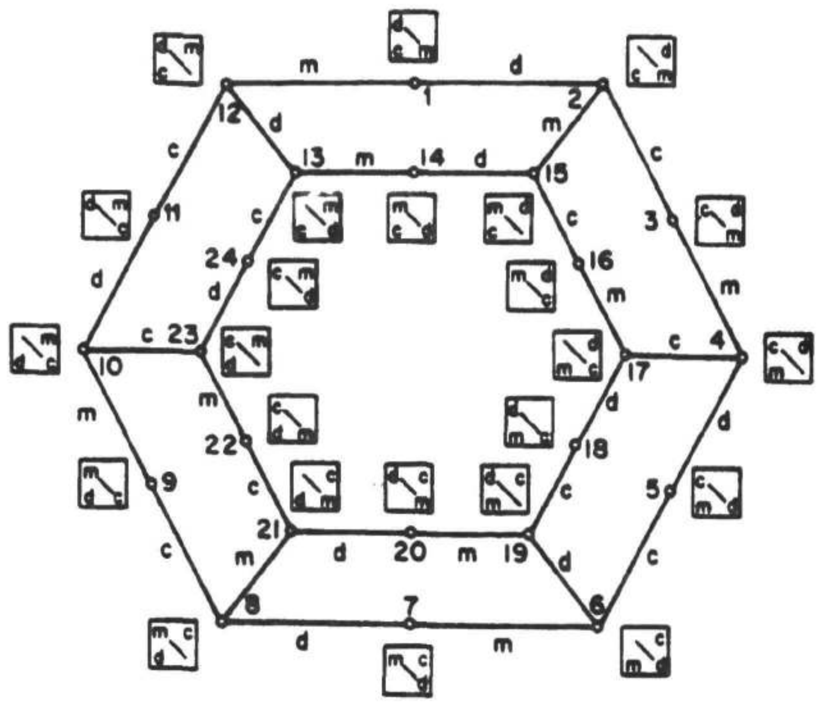

Figure 2 State space for the DCM problem. Each node represents a unique configuration of the three animals.

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

* Rotation problems have both initial and final states on the same hexagon -- either the inner or the outer. They have minimum paths that do not use the diagonal of the game board (Examples: 1-5, 23-17). Permutation problems have initial and final states on different hexagons, and require the use of the diagonal. These problems start and end with different permutations of the three animals, and the permutation order can be changed only by using the diagonal (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.

* If we abstract over the specific identity of the pieces, then there are only two types of nodes: those with open diagonals, having three adjacent states (e.g., 2, 4, 19, 21), and those with closed diagonals, having two adjacent states (e.g., 1, 3, 18, 20). At an open node, there are three possible moves; at a closed node, there are two possible moves.

Problem selection

Problems were designed to vary path length (from 4 to 7), type of initial node (open or closed diagonal), and problem type (permutation or rotation). The eight problems selected are listed in the bottom section of Table 1.[2] In addition, four three-move training problems were used. They are shown at the top of Table 1.

Table 1: Problem Set

		Initial State	Goal State	Path Length	Initial Node	Problem Type
Training Set	T1	1	4	3	closed	rotation
	T2	7	22	3	closed	permutation
	T3	12	9	3	open	rotation
	T4	2	17	3	open	permutation
Problem Set	1	17	21	4	open	rotation
	2	18	8	4	closed	permutation
	3	11	20	5	closed	permutation
	4	10	5	5	open	rotation
	5	13	19	6	open	rotation
	6	24	18	6	closed	rotation
	7	14	7	7	closed	permutation
	8	15	8	7	open	permutation

Method

Subjects

Thirty-nine predominantly middle-class children, ranging in age from 45 to 70 months old, completed this experiment.

Procedure

Problems were presented in the order shown in Table 1. Children were told a cover story about animals who wanted to get to their favorite food. 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, then it 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.

Results

For each trial, subjects were assigned a 0/1 score based on the number of moves they made. If the number of moves made on any trial was greater than two more than the minimum path, then the score was 0, otherwise it was 1. Note that the scoring is somewhat more lenient than the actual procedure used to decide whether or not to repeat a trial. Two extra moves were allowed because they correspond to common minor errors in this particular state space:

* On rotation problems, if subjects unnecessarily use the diagonal to move from the inner to outer hexagon, and then move back to the correct hexagon, they will make two extra moves.

* If subjects make a false start, but immediately correct it, then they will make two extra moves.

Each subject was assigned a score based on the percentage of passes (1s) -- on either the first or second presentation of each problem -- across the eight problems. Each problem was assigned a score based on the proportion of subjects passing it. The scores, ranked by subject performance and problem difficulty, are shown in Table 2.

Subjects' performance varied widely: from maximum scores for the three best subjects to almost total failure for the worst subject. Problem difficulty also varied widely: from nearly all subjects passing the easiest problem to about two-thirds of the subjects failing the hardest problems.

The most important result shown in Table 2 is the rank order of the problems. Recall that the problems varied in path length from 4 moves (problems 1 and 2) to 7 moves (problems 7 and 8). Path length is a poor predictor of problem difficulty ($\underline{r}$ = .34). (See also the solid line in Figure 3.) The two easiest problems are also the two shortest, but even though they both have a path length of 4, there is a 20% difference in the proportion of subjects passing them. The next two easiest problems are the two longest (7 moves). The four hardest problems are intermediate in path length, and within that set, there is a large difference between the pairs with the same path length.

In the following analysis, we will show how path length, solution strategy and the structure of the problem space interact to produce this pattern of results.

Strategic analysis

How might children attempt to solve these problems? In this section we will describe a basic strategy and compare it to the subjects' performance. Then we will propose several variations on that strategy, and show that none of them fit the data as well as the basic strategy.

Consider the following procedure -- called Strat2 -- for making moves in the

Table 2: Pass/fail scores (by second trial) for all subjects on all problems. Ranked from best to worst subject and easiest to hardest problem.

Problem Number / Subject	2	1	8	7	3	5	6	4	mean
4	1	1	1	1	1	1	1	1	1.0
12	1	1	1	1	1	1	1	1	1.0
26	1	1	1	1	1	1	1	1	1.0
18	1	1	1	1	1	1	0	1	.88
20	1	1	1	1	1	1	1	0	.88
17	1	1	1	1	1	1	*	0	.85
5	1	1	1	1	1	0	1	0	.75
11	1	1	1	1	1	1	0	0	.75
14	1	1	1	1	0	1	1	0	.75
33	1	1	1	1	0	1	1	0	.75
3	1	1	1	1	0	1	0	1	.75
36	1	1	1	1	0	1	0	1	.75
27	1	1	1	1	1	0	0	1	.75
2	1	1	1	0	1	0	1	1	.75
16	1	1	1	0	1	1	0	1	.75
7	1	1	1	1	0	1	0	0	.63
39	1	0	1	1	0	1	1	0	.63
56	1	0	1	1	1	0	1	0	.63
41	1	0	1	1	1	0	0	1	.63
51	1	1	0	0	1	0	1	1	.63
1	0	0	1	1	1	0	0	1	.50
13	1	1	1	1	0	0	0	0	.50
19	1	1	1	1	0	0	0	0	.50
50	1	0	1	0	1	1	0	0	.50
9	1	1	1	0	0	0	1	0	.50
29	1	1	1	0	1	0	0	0	.50
6	1	1	0	0	0	1	0	1	.50
30	1	1	0	0	1	*	0	0	.43
15	0	0	1	1	0	1	0	0	.38
25	1	0	0	1	0	1	0	0	.38
31	1	0	0	1	0	1	0	0	.38
43	1	1	1	0	0	0	0	0	.38
45	1	1	0	0	0	0	1	0	.38
8	1	1	0	0	1	0	0	0	.38
40	1	1	0	0	1	0	0	0	.38
10	1	1	0	0	0	0	0	0	.25
55	1	1	0	0	0	0	0	0	.25
24	1	0	0	0	0	0	0	0	.13
42	1	0	0	0	0	0	0	0	.13
Problem mean	.95	.74	.69	.59	.51	.50	.34	.33	

DCM state space:

1. If there is a two-move sequence that can reach the goal state, then make it, otherwise:

2. Generate all candidate moves: (all legal moves, except the piece just moved.)

3. If there is more than one candidate, choose randomly.

4. Go to step 1.

This is a simple generate-and-test strategy, with two constraints: a) Two-move lookahead to the goal state. The lookahead has a very simple evaluation function: the state is either the goal state or it is not. No partial evaluations are made (such as the number of pieces in their goal positions.) b) No immediate backup. In the DCM puzzle, a constraint against moving the same piece twice is equivalent to a prohibition on immediate backup.

We can determine the probability that Strat2 would discover a minimum path solution for each problem by computing the compound probabilities that it will stay on a minimum path.

By applying this analysis to each of the eight problems, we can compare the probability that Strat2 would pass each problem with the subjects' actual performance. For each of the problems in Table 3 -- ranked from easiest to hardest -- we have listed the problem number, the initial and final states, the path length, the probability that Strat2 would find the minimum path solution on a single trial, the probability that Strat2 would be successful if it were given two chances to find the minimum path, and, in the final column, the proportion of subjects passing each problem by the second trial.

Table 3: Subject performance and Model performance

Problem Number	States Initial		Final	Path Length	Strat2 p	Strat2 $2p - p^2$	Proportion Passing
2	18	--->	8	4	.500	.750	.95
1	17	--->	21	4	.333	.556	.74
8	15	--->	8	7	.500	.750	.69
7	14	--->	7	7	.500	.750	.59
3	11	--->	20	5	.375	.609	.51
5	13	--->	19	6	.333	.556	50
6	24	--->	18	6	.250	.440	.34
4	10	--->	5	5	.167	.310	.33

A plot of both the model's and the subjects' likelihood of success for each problem, is shown in Figure 3. Strat2 explains almost 60% of the variance in problem difficulty (r = .767, t = 2.9, df = 7, p < .05). If we eliminate the two 4-move problems, which were much easier for the subjects than for the model, then the correlation between Strat2 and subject performance is r = .95 (t=6.02, df=5, p < .01).

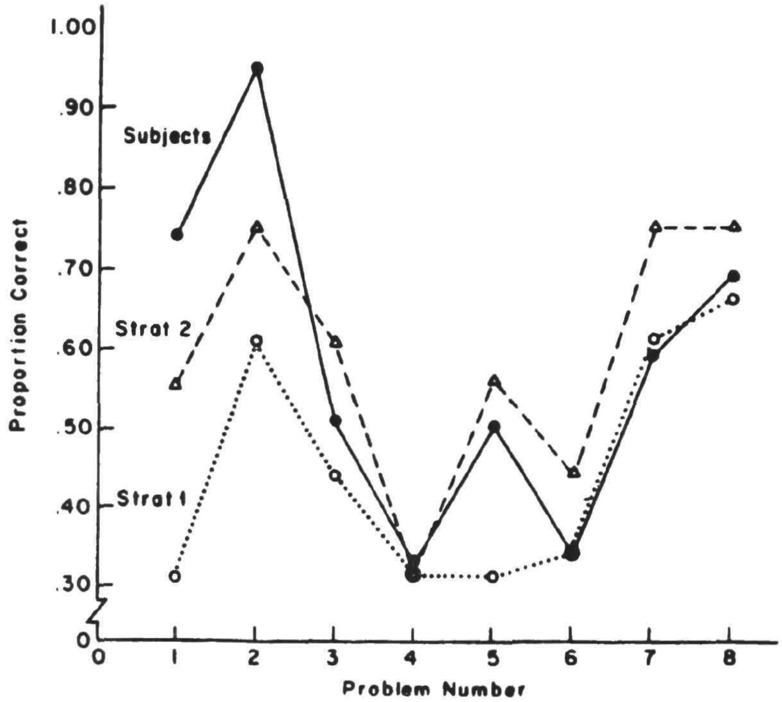

Figure 3 Probability of passing a problem by the second trial for subjects (solid line), Strat2 (dashed line) and Strat1 (dotted line).

Strat2 can be characterized as a random walk through the state space with two constraints: no immediate backup, and a two-step lookahead for the goal state. We can ask two questions about these constraints. First, how well do subjects adhere to them? Second, how important are they?

<u>The No-Backup Constraint.</u> Children's compliance with the no-backup constraint was assessed by counting the number of times - over both successful and unsuccessful trials - that they moved the same piece twice in succession. Overall, there was a violation rate of 11%. If moves were made without the constraint, we would expect 33% of moves to be double moves.

Removing the no-backup constraint from Strat2 substantially reduces the probability of solution. All the two-way branches become three-way branches, and all the direct connections (e.g., 11-10 in problem 7) become binary nodes. Even with two-move lookahead, such a model would perform far below the average solution rates for our subjects.

Depth of Lookahead. Compliance with the two-move lookahead constraint was assessed by computing the proportion of trials in which subjects move to the goal directly from states that are n moves distant from the goal. Figure 4 shows the actual proportion of minimum path solutions as a function of distance from the goal. Also shown are the proportions predicted by Strat2 and by a random move generator.

Strat2's two-move lookahead predicts perfect performance from up to 2 moves away from a goal, and then a sharp decline. Subject performance is indeed quite good at 2 moves away, but it remains high (nearly 90%) for 3 moves away, rather than dropping as predicted. In fact, about 40% of the subjects exhibited perfect performance once they were 3 moves away from the goal.

Given this relatively good performance from 3 moves away, it is reasonable to consider an alternative to Strat2 that differs only in having three-move, rather than two-move lookahead to the goal. Strat3 would produce very high likelihoods of success within two trials, ranging from .97 and .94 for problems 8 and 7, to lows of .56 for problems 1, 4, and 5. Not only does Strat3 produce unacceptably high solution rates, but also, it only explains about 5% of the variance in subjects' solution rates.

If we degrade the two-move lookahead to a one-move lookahead, then we get a model that explains only 26% of the variance.

All-or-none Evaluation. Associated with Strat2's two-move lookahead is an all-or-none evaluation function. If the children were using a partial evaluation function that was sensitive to some -- but not all -- of the pieces being in their goal positions, then we should see two kinds of biases in their move patterns. One bias would show up as a tendency to favor moves -- early in the solution -- that increase the number of pieces in their goal locations. For example, in Problem 2 (18 --> 8), a first move of the cat increases the evaluation function, while moving the dog does not. The dog is also off the minimum path. 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 Problem 4 (10 --> 5), the minimum path move is the mouse, which does not increase the evaluation function. Only the cat increases the partial evaluation function, and it is preferred on 66% of the trials, even though it is off the minimum path. Similarly, on Problem 5 (13 -->19), the non-minimum move of the dog is preferred on 61% of the trials.

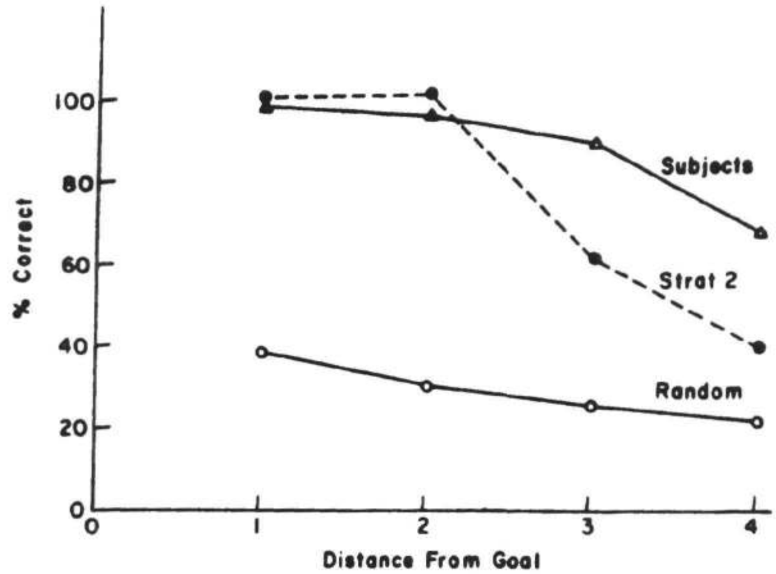

Figure 4 Proportion of minimum path solutions from n moves away, for Strat2, subjects and a random model.

The other bias would be a reluctance to remove pieces from their goal locations -- to reduce the value of a partial evaluation function. This can be assessed on Problem 3 (11-->20), where the minimum path sequence requires that the dog be temporarily removed from its goal position. 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. The all-or-none evaluation function in Strat2 understates the sensitivity of children to partially correct solutions.

Summary of Strategic Analysis

Strat2 explains almost 60% of the variance over all problems and 95% of the variance over the six most difficult problems. Strategies that vary the depth of the lookahead do not do as well. Strat1 explains 26% of the variance, and Strat3 only 5%. Elimination of no-backup from Strat2 yields unacceptably low solution rates.

Strat2 slightly understates children's abilities in two respects. First, the children appear to be capable of some partial evaluation, whereas Strat2 is not. Second, once they are only 3 moves away from the goal state, the children are more likely to find a minimum path solution than is Strat2. Nevertheless, within the space of plausible alternative strategies explored here, Strat2 provides the best account of how children solve problems with ambiguous subgoals.

Discussion

Piaget (1976) concludes from his observations of 5- and 6-year-old children solving conventional TOH problems that they are unable to plan and that "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, studies of pre-schoolers solving a modified version of the TOH (Klahr & Robinson, 1981) show that, although the amount of planning they can do is limited, the procedures they use are highly similar to adult forms.

In this investigation, pre-schoolers were presented with problems having ambiguous subgoals. We discovered that here too, Piaget's characterization does not do justice to young children's abilities. First, as described earlier, even the random component of Strat2 is highly constrained. The avoidance of double moves reveals a rudimentary knowledge about thoroughly useless actions that is not conveyed by the "trial-and-error" view. Second, solutions are not really "arrived at by chance", since there is a lookahead to the goal state, and little deviation from the minimum path, once it is in sight. Third, children use partial evaluations of nearly correct states to guide their choice of moves.

Acknowledgements

This work was supported in part by grants from the National Science Foundation (BSN81-12743) and The Spencer Foundation. Thanks to the parents, the teachers, and particularly the children of the CMU Children's School for their enthusiastic cooperation, and to Margaret Kinsky for her artistry, advice and assistance in creating materials, running the experiments and analyzing the results. Author's address: Department of Psychology, Carnegie-Mellon University, Pittsburgh, PA 15213.

References

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: Lawrence Erlbaum Associates.

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

Laird, J.E. & Newell, A. (1983a). **A universal weak method.** Technical Report, Computer Science Department, Carnegie-Mellon University.

Newell, A. (1969). Heuristic programming: Ill-structured problems. In J. Aronofsky

(Ed.), **Progress in operations research, III** New York: Wiley.

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

Spitz, H.H. & Borys, S.V. . (1984 [in press]). Depth of search: How far can the retarded search through an internally represented problem space. In P.H. Brooks, R. Sperber, & C. McCauley (Eds.), **Learning, cognition and mental retardation** Hillsdale: Lawrence Erlbaum Associates.

Spitz, H.H., Webster, N.A., & Borys, S.V. (1982). Further studies of the Tower of Hanoi problem-solving performance of retarded young adults and nonretarded children. **Developmental Psychology**, 18(6), 922-930.

Notes

[1]The DCM puzzle is nearly identical to the "depth-of-search" puzzle first described by Spitz & Borys (1984).

[2]The problem shown in Figure 1 is Number 3 in Table 1.