An ACT-R Model of the Evolution of Strategy Use and Problem Difficulty

Glenn Gunzelmann (glenng@andrew.cmu.edu)
John R. Anderson (ja+@cmu.edu)
Department of Psychology, Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

Research has shown the importance of strategies in guiding problem solving behavior. The experiment and model presented here provide further specification of how more optimal strategies come to be adopted with experience. Isomorphs of the Tower of Hanoi were used to allow participants to develop a degree of expertise with a novel task. In the solutions, evidence for at least two strategies is apparent. The results suggest that when strategies are not successful in achieving the goal, other strategies may emerge and eventually come to dominate performance in a task. The ACT-R model of this task captures participant performance by using the same strategies to solve the problems and by gradually switching to more effective ones as simple strategies fail in solving the problems.

Introduction

Strategies are ubiquitous in problem solving. Even in novel tasks, participants bring general strategies to bare while searching for the correct solution. As experience is gained with a problem, these strategies often are abandoned in favor of strategies that are more particular to the task (Anzai & Simon, 1979). As the new strategies are discovered and practiced, solving problems within the task becomes easier and solutions become more efficient. Eventually, a strategy or set of strategies may be developed that can produce a correct solution to any problem for a particular task.

Strategic Influences on Problem Solving

Research on strategies in problem solving has taken two general forms. First, some researchers have focussed on the impact of particular strategies on solution times and accuracy (e.g. Altman and Trafton, 1999; Anderson, Kushmerick, and Lebiere, 1993). These researchers accurately model the particulars of strategy execution in participants. They do not, however, track the shifts in strategy use that typically occur as experience is gained with a particular task. Others have looked at the use of different strategies as individuals learn to solve problems (Lovett and Schunn, 1999; Reder & Schunn, 1999; Siegler, 1987). These researchers look at how multiple strategies for a task may coexist. The experiment presented here looks at the evolution of strategy use in solving the Tower of Hanoi problem. In addition, an ACT-R model of the task captures overall participant performance while closely matching the strategies they use and the transitions among them.

The Tower of Hanoi has served as a useful task in problem solving research for a number of years (e.g. Anderson, et. al., 1993; Anzai and Simon, 1979). The task itself consists of three pegs upon which are placed any number of disks. Figure 1 illustrates the elements of the Tower of Hanoi task and the two isomorphs that we will be studying in our experiments. The goal is to change the disk arrangement from some start state into some particular goal state. There are three rules to guide movement through the problem space for the Tower of Hanoi. The first rule states that only one disk may be moved at a time. The second rule indicates that if more than one disk is on a particular peg, then only the smallest of these disks may be moved. The final rule says that a larger disk may not be moved to a peg where there is a smaller disk. This results in the necessity that the disks form a tower structure at all times, with larger disks always being underneath smaller disks.

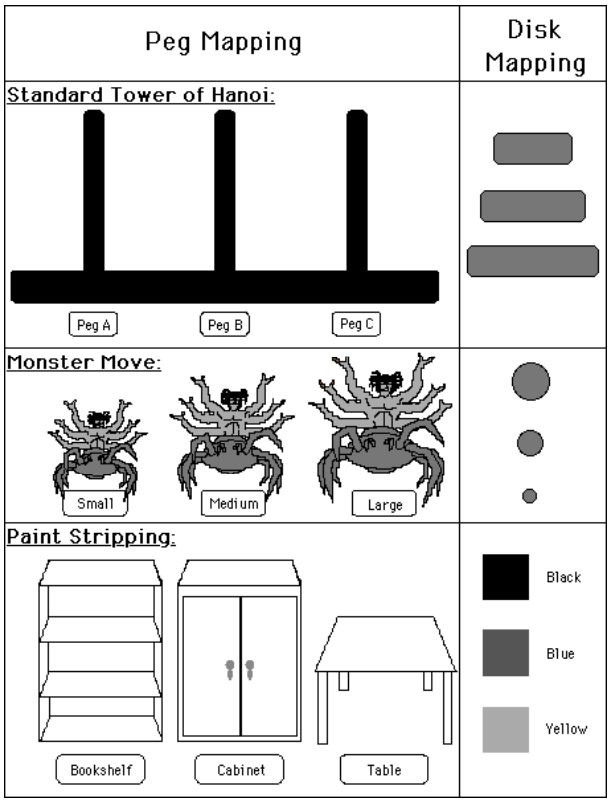

Figure 1: Mapping of the elements of the isomorphs used in this study to the standard Tower of Hanoi.

Research on the Tower of Hanoi indicates the importance of some variant of a disk subgoaling strategy (Anzai & Simon, 1979). In this strategy, subgoals are created to deal with the largest disk out of place. When the largest disk is placed into its goal state, focus is shifted to the next largest disk. This process is repeated until the smallest disk is placed. This strategy is quite effective, usually producing an optimal solution, regardless of the particular problem presented. While disk subgoaling may come to dominate participant solutions in the end, it is not usually the case that participants initially use this strategy. Rather, it tends to emerge as familiarity with the task increasese (Anzai & Simon, 1979). This is particularly the case with isomorphs (Kotovsky, Hayes, & Simon, 1985) where participants tend to start out with some sort of random search or simple hill-climbing strategy and only gradually evolve a preference for subgoaling. Also, subgoaling is not initially as predominant if participants are presented with flat-to-flat problems (problems with start and goal states where one disk is on each peg) rather than with classic tower-to-tower problems. Our research will use isomorphs and flat-to-flat problems, where other strategies often predominate early. However, these are not as effective as subgoaling. We want to study and model the process by which participants come to prefer the disk subgoaling strategy.

Experiment

In the three-disk Tower of Hanoi, there is a particular class of problem states in which there is one disk on each of the three pegs (flat states). There are a total of six of these states, and for each there are exactly two other flat states that are 5 moves away (minimum number of moves). Based on the disk subgoaling strategy, getting to one of these other flat states involves deeper subgoaling (*hard*) than getting to the other (*easy*). The breakdown of the subgoal structure of these problem types is illustrated in Figure 2. It shows that two subgoals need to be formed in the *hard* problems before making the first move, while a single subgoal is sufficient in the *easy* problems. An important feature of these two problem types is that they are otherwise quite similar. They both require 5 moves to solve, utilize the same set of start and goal states, and can be solved optimally using a similar sequence of moves. Because of these interesting characteristics, it is these problems that are used in the current study.

Since the superficial features of a task (cover story) can exert a strong impact on difficulty (Kotovsky, et. al., 1985), two different isomorphs of the Tower of Hanoi are. This is to help insure that any differences found are not simply an artifact of the cover story, but rather involve something more directly related to the task's structure. The Tower of Hanoi isomorphs used in this study are the Monster Move isomorph (Kotovsky, et. al., 1985) and the Paint Stripping isomorph (Gunzelmann & Blessing, 2000). In terms of the isomorphs, the pegs in the Tower of Hanoi are synonymous with monsters in the Monster Move isomorph and with pieces of furniture in the Paint Stripping isomorph. The disks are represented by globes held by the monsters and by layers of paint on the furniture. The relationships among these three isomorphs are illustrated in Figure 1. The relationships among elements are a bit easier to describe in terms of the standard Tower of Hanoi, so the results will be discussed in terms of “disks” and “pegs”.

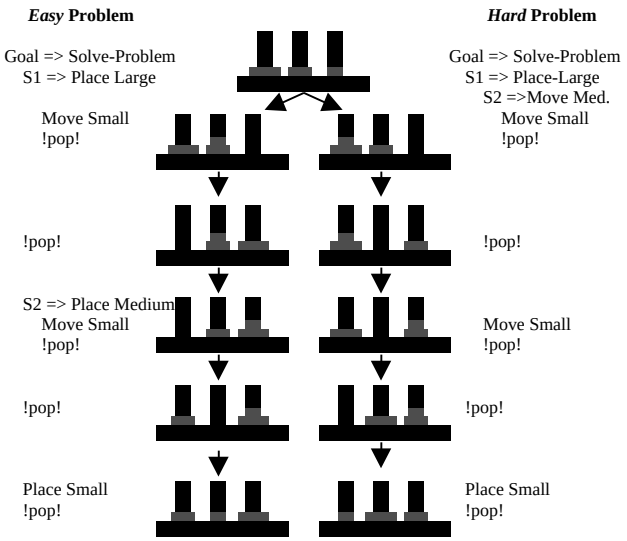

Figure 2. Breakdown of the *easy* and *hard* problem types used in this study.

Since these problems all begin and end with flat states they tend to encourage a particular kind of problem solving strategy in which participants simply transform flat state into flat state, looking for the goal state. One flat state can be transformed into another flat state using a three-move sequence of moving one disk onto a second disk, moving the other disk to where the first disk had been, and then moving the first disk to where the third disk had been. In effect, this switches the location of the first and third disk. At best, such a flat-to-flat strategy will solve the problems in 6 moves (two disk switches each taking 3 moves) rather than the optimal 5. We were interested in seeing how this flat-to-flat strategy would evolve in competition with a disk subgoaling strategy.

Method

The participants were 24 undergraduate students from Carnegie Mellon University. Participants received either course credit (n=7) or were paid (n=17) for their participation in the one-hour experiment.

The entire experiment was completed on a computer. Each participant was given a sequence of three tasks, with the first and third being the same Tower of Hanoi isomorph. The second task was given as a filler task. Before they

began solving the problems for each task, participants were presented with a problem statement (cover story), a set of three rules, and an explanation of how to use the interface. Participants were instructed to solve each problem for each task by reaching the goal state that was presented on the screen. If an error was made while solving any of the problems, a message box appeared restating the rule that had been violated. After each problem, a message box appeared indicating that they had solved it correctly. The same procedure was followed for each of the three tasks.

Participants were randomly assigned to groups based upon, (1) The cover story for the Tower of Hanoi isomorph, (2) The type of problems they completed in the first set of problems (*easy* vs. *hard*), and (3) the type of problems for the second set. Each participant completed 6 problems in each isomorph set. Since the pattern of results was the same for both isomorphs, the data presented here is combined across them.

Results and Discussion

Evidence for the use of a strategy like disk subgoaling comes from problem solutions and the corresponding move latencies. Of all problems, 42% were solved optimally, and an additional 30% had solutions that incorporated an optimal 5-move final path. The move latencies for these solutions support the conclusion that participants were planning and executing 2-moves in sequence, similar to the data reported by Kotovsky, et. al. (1985). That is, move latencies were longer for the first and third moves than for the other three (7.2 versus 2.0 seconds on average), suggesting that more planning occurs before those moves are made. While this does not necessarily mean that participants were using the disk subgoaling strategy specifically, it is reasonable to conclude that they were using a strategy at least quite similar to it.

The evidence also suggest that flat states were particularly attractive to participants as they tried to find a solution. If moves were made entirely at random, it would be expected that participants would arrive at flat states every 4.5 moves (6 of 27 states are flat states). However, the rate was actually every 3.33 moves for participants (ignoring flat-to-flat 5-move final paths), with the minimum distance between flat states being 3 moves (see above). In addition, of the problems that were not solved using a 5-move final path, half had a final path that incorporated only a single additional move, which involved moving through an intermediate flat state en route to the correct solution. The data indicating a preference for flat states is further enhanced by the latency data collected. For *every* participant, move latencies were greater for flat states than for other states. This provides persuasive evidence that more planning occurred in flat states than in other states and that participants were implementing a flat-to-flat strategy. That is, participants seemed to be planning and executing sequences of 3 moves that transformed the problem state from one flat state into another.

There are two explanations for why participants may have learned such a strategy. First, as stated above, both the start state and the goal state were flat states, immediately drawing participants attention to them as somehow important in the task. Second, in flat states, rule 2 does not apply (if there is more than one disk on a peg, only the smallest may be moved), simplifying the evaluation needed to plan a move. This would reduce the memory load for planning a move, and perhaps allow participants to look further ahead in the problem to plan multiple moves. As such, these states become a “home base” of sorts where participants can regroup and consider alternatives.

It is important to recall that the participants in this study were not given training on strategies for solving any of the problems. Thus, the solution strategies were developed by the participants as they worked through the problems. Still, the *hard* problems took, on average, 1.1 additional moves to solve than the *easy* problems. Though this effect did not reach statistical significance in this experiment, it was fairly robust. It was constant across isomorphs and remained fairly stable across problem number. The fact that this difference did appear, and in the expected direction, lends insight into participants’ representations of the problems and how they sought to solve them. Combined with the evidence for disk-subgoaling, these findings suggest that the added level of subgoaling made it more difficult for participants to successfully plan and execute the moves to solve the hard problems optimally.

There is ample evidence for both a disk-subgoaling and a flat-to-flat strategy in participants’ data. In the next section on the model, we will provide further analysis of these results. Also there is evidence for a shift in these strategies. During the first 6 problems, subjects solved 64% of the problems in the 5 moves dictated by the disk subgoaling and 24% of the problems by going through a pair of flat to flat transformations. For the second 6 problems these percentages were 83% disk-subgoaling and 14% flat-to-flat.

Model

An ACT-R model (Anderson, 1976; Anderson & Lebiere, 1998) of participant performance was developed with the goal of capturing the overall performance of participants while simultaneously matching their strategy use. Based on the data presented above, the model was constructed to use three different strategies as it went about solving the problems presented. These were the disk subgoaling strategy, the flat-to-flat strategy, and a random strategy. The random strategy allows for the unfocussed meandering about the problem space that is particularly characteristic of the early stages of problem solving in a novel task (Kotovsky, et. al., 1985).

Model Design and Mechanisms

The ACT-R model evaluates the success of a strategy by noting whether or not each use of the strategy leads to a solution to the problem. If the evaluation of one strategy becomes increasingly negative, there will be a tendency to switch to other, potentially more effective, strategies. Over time, the model will come to settle on the strategies that are generally more effective. To accomplish this, the model goes through a series of iterations of (1) choosing a strategy, (2) executing the strategy, (3) evaluating the result (i.e. has the problem been solved). The critical stage in this process within the model is strategy selection. At the point in the problem where a strategy needs to be chosen, there are three productions that may fire (one for each of the strategies). The choice of which production fires in ACT-R is governed by the calculation of the "expected gain" (*E*) for each production. In this process, a quantity is calculated for each production to represent how quickly its use is expected to result in satisfying the goal. The production producing the highest value for this quantity is selected and fires. The equation for expected gain (*E*) in ACT-R is:

$$\mathbf{E = PG - C}$$

where *P* is the probability that both the production will succeed and the goal eventually will be achieved, *C* is the anticipated cost (in seconds) of achieving the goal using the production, and *G* is a global variable representing the value (in seconds) of achieving the goal (i.e. how much time is the model willing to spend to solve the problem). The value of *G* was set at 50. While this value is traditionally set at 20 in ACT-R, these problems take longer than that for participants (and the model) to solve. So, this value was raised to accommodate the greater amount of time needed to solve them. For this model, the initial values of *C* were equal for all the strategies. But, as the model performs the task, it adds the cost incurred in executing each strategy to the value of *C* in the strategy-choice production. In contrast, the *P* values were estimated for each production. The equation for *P* is:

$$\mathrm{P} = \mathbf{Successes/(Successes + Failures)}$$

where "successes" and "failures" refer to the number of eventual successes and eventual failures that occurred when this production was used. That is, how often has the goal been achieved and how often has the goal not been achieved when this production has been used? The initial values for the "successes" and "failures" for each of the three strategy choice productions were parameters estimated in fitting the model. While all these values were set, it is really the relative difference in this ratio among the three productions that matters most in the model. In addition, the sum of successes and failures was made to equal 50 for all three. This quantity controls the stability of *E* by influencing how much a single success or failure will affect the calculation of *P*. As the model performs the task, it gains experiences and adds to these values. Each time a strategy is executed either a success (if the problem is solved during the execution of the strategy) or a failure (if the problem is not solved) is added to the calculation of *P*. Thus, as the model gains experience, the rapidity of change and the impact of any single attempt at using a strategy diminishes. That is, the model slowly begins to settle on the most successful strategies. The initial values and the values after learning for the variables (for a single run of the model in one condition of Experiment 1) are presented in Table 2.

Table 1: Initial parameter settings for model and their values after learning.

	Successes	Failures	*P*	*E*
Initial				
Disk-subgoal	25	25	.50	23.95
Flat-toFlat	26	24	.52	24.95
Random	28	22	.56	26.95
12 *hard* trials				
Disk-subgoal	34	29	.54	22.16
Flat-toFlat	29	31	.49	21.35
Random	28	36	.44	20.13
100 *hard* trials				
Disk-subgoal	121	37	.77	24.89
Flat-toFlat	30	42	.42	16.62
Random	28	46	.38	16.76

The values of *P* are not perfectly correlated with the values in *E* is because *E* includes the costs incurred for executing each strategy (*C*). Although disk-subgoaling is more successful at solving the problems, it is more costly (in terms of time) because it involves more planning and more moves per attempt than the others. These offsetting influences maintain the mixture of strategies over the course of the experiment. With more problems, the degree of separation increases, leaving the disk-subgoaling strategy as the preferred strategy. The values for the variables after the model solves 100 *hard* problems are presented at the bottom of Table 2.

The final parameter of importance is a noise parameter that is added to the calculation of *E*. A noise value is produced separately for each production on each cycle of the model. In this model, the value is randomly selected from a distribution with a mean of 0 and a standard deviation of about 1.8. The strategy-choice production selected to fire is the one that has the highest value of *E* after noise has been added to the calculation described above for each.

A major assumption of the model is that most participants would begin the experiment without a clear idea of how to solve the problems. In the model, this is instantiated in the initial values of *E* for the three strategy-choice productions. In particular, the "successes" and "failures" were set such that the random strategy was initially preferred. This was

followed by the flat-to-flat strategy, with the disk-subgoaling strategy least preferred. The reason for this ordering is that the disk-subgoaling strategy is the most sophisticated strategy in the model. Thus, the model tends to begin with the simplest strategy (i.e. make a move and evaluate the result) and moves toward more efficient, though more complicated ones.

There is one additional mechanism that is critical for the fit of the model to the data presented. It is largely responsible for the difference in difficulty found between the *easy* and *hard* training problems and operates similarly to the strategy-choice mechanism described above. At the point in the disk-subgoaling strategy where a second subgoal needs to be formed, there are two productions that may fire. One accurately forms the second subgoal and the problem is ultimately solved correctly from that point. The other gives up on the strategy and goes on to try a different approach (perhaps returning to try disk subgoaling at a later point). The critical parameter for this mechanism is once again the initial difference in the expected gain value for the two productions. For the sake of simplicity, the same sum of successes and failures (50) was used for these productions as for the strategy-choice productions. The estimated "successes" and "failures" for the production to give up were 40 and 10 respectively (P=.80, E=38). For the production that successfully pushed the second subgoal, the values were 38 and 12 respectively (P=.76, E=36). These impact the value of E for these productions similarly to the strategy choice productions presented in Table 2. This means that the model tends to give up initially. However, since this inevitably means that the strategy will fail to achieve the goal (adding a failure to the production that gives up), it learns rather quickly to press on, execute the additional level of subgoaling, and solve the problem successfully (adding a success to the "push-on" production). All of these parameters were estimated to fit the aggregate move data (Figures 3 and 4; averaged into quartiles). The more detailed data on strategy use is examined next.

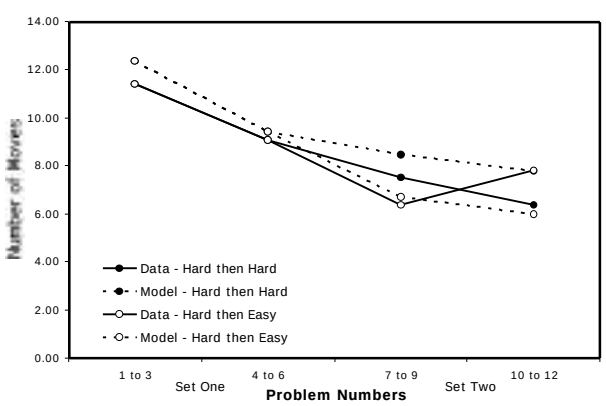

Figure 3: Model fit to moves to solve data where the first set consisted of *hard* problems.

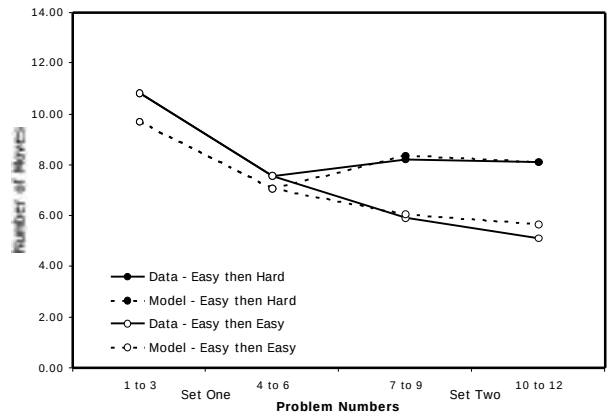

Figure 4: Model fit to moves to solve data where the first set of problems was *easy* problems.

Fit to the Strategy Data (Final Paths)

The rather good fit to the average move data (r^2=.91, mean-deviation=.798) suggests that the model is capturing human performance and learning in the task. However, a more compelling argument for the model comes from the fit of the model to more detailed accounts of the participants' solutions. The best way to examine strategies that participants were using is to look at how they actually solved the problems they were given. From Kotovsky, et. al. (1985) comes the idea of a 2-stage solution process, an initial exploratory phase where no progress is made toward the goal followed by the final path, where generally a rapid and efficient solution is produced. The final path begins when the person is the same number of moves from the goal as he or she was at the start of the problem. For the *hard* and *easy* problems, this means that the final path begins the last time the participant is 5 moves from the goal (via the shortest path) before actually solving it.

Previous research has found that the exploratory moves relate more closely to problem difficulty than length of the final path (e.g. Kotovsky, et. al., 1985). In concert with this finding, the difference in final path behavior between the *easy* and *hard* problems is not large. On the other hand, final path length is informative about the strategies being used by participants as they solve problems. In particular, the final path length of the problems used here can help to differentiate among problems solved using a disk-subgoaling strategy, solutions using a flat-to-flat strategy, and solutions involving a more random sequence of moves. As stated above, the disk subgoaling strategy, if executed correctly, will give rise to perfect solutions and optimal final paths. On the other hand, the flat-to-flat strategy will produce solutions that are slightly less than optimal. For the 5-move problems used here, a flat-to-flat solution would take 6 moves (2 consecutive flat-to-flat transformations).

To examine the fit of the model to the solutions produced by participants, the length of the final path was determined for each problem solution in both the model and the data. In Table 2, the percentage of problems solved with final paths suggestive of each strategy is indicated by condition. The remaining problems involved longer final paths that did not follow a readily identified pattern.

Table 2: Percentage (%) of problems supporting the disk subgoaling and flat-to-flat strategies by condition.

	Data	Model
Set One		
Easy		
Disk Subgoal	63	68
Flat-to-Flat	25	22
Hard		
Disk Subgoal	67	68
Flat-to-Flat	24	15
Set Two		
Hard (from Hard)		
Disk Subgoal	86	78
Flat-to-Flat	6	13
Hard (from Easy)		
Disk Subgoal	77	82
Flat-to-Flat	19	15
Easy (from Hard)		
Disk Subgoal	83	81
Flat-to-Flat	14	12
Easy (from Easy)		
Disk Subgoal	83	87
Flat-to-Flat	17	12

From the data presented in Table 2, it is clear that the model reproduces much of the strategic richness of participant performance. To test this assertion, a Chi-square test was performed on the participant versus model data. This statistic provides a rough indication of the similarity of the predicted model data to the obtained empirical data. For the data in Table 2, $X^2(11)=7.59$, $P>.05$. This suggests that the model is doing a very good job of modeling the strategies used by participants while solving the problems.

Conclusion

The data and model presented here provide evidence that the failure of simple strategies to reach the goal can lead individuals to switch to the use of more sophisticated strategies for achieving that end. As strategies are attempted, they are evaluated in terms of their success in achieving the goal, but also in terms of the costs associated with executing them. While sophisticated strategies may initially fare poorly, due to a greater cost, in the end they are likely to emerge as the preferred strategy due to their greater likelihood of successfully solving the problem.

The participants in this study were given no instruction on how to solve the problems they were given. Initially, they were unsure of how to maneuver through the problem space, as suggested by the many apparently random moves that were made. But, as they gained experience with the isomorphs, their solutions became increasingly organized, and clear evidence of the two strategies described here emerged. The model accurately captures this aspect of participant solutions, showing a gradual and noisy shift from completely undirected random moves to optimal solutions using a sophisticated strategy. This kind of transition in strategies is certain to appear in other tasks, as it has already been noted in children solving addition problems (Siegler, 1987). The richness of strategy use in participants is an important aspect of problem solving behavior, and one that warrants careful consideration in problem solving research.

References

Altman, E. M., & Trafton, J. G. (1999*). Memory for goals: An architectural perspective*. Proceedings of the Cognitive Science Society (pp. 19-24).

Anderson, J. R. (1976). *The Adaptive Character of Thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (1998). *Atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R., Kushmerick, N., & Lebiere, C. (1993). The Tower of Hanoi and Goal Structures. In Anderson, J. R., *Rules of the Mind* (p. 121-142). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gunzelmann, G., & Blessing, S. B. (2000). *Why are some problems easy? New insights into the Tower of Hanoi*. In Proceedings of the Cognitive Science Society (p. 1029).

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985) Why are some problems hard? *Cognitive Psychology, 17*, 248-294.

Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability, and base-rate neglect. *Journal of Experimental Psychology, General, 128*, 107-130.

Reder, L. M., & Schunn, C. D. (1999). Bringing Together the Psychometric and Strategy Worlds: Predicting in a Dynamic Task. In Gopher, D. & Koriat, A. (Eds). *Cognitive regulation of performance: Interaction of theory and application. Attention and Performance XVII* (pp. 315-342). Cambridge, MA: MIT Press.

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

Spitz, H. H., Minsky, S. K., & Bessellieu, C. L. (1984). Subgoal length versus full solution length in predicting Tower of Hanoi problem-solving performance. *Bulletin of the Psychonomic Society, 22*, 301-304.