Available at www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

ELSEVIER

Cognitive Systems Research 4 (2003) 57–76

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Problem solving: Increased planning with practice

Action editor: Gregg Oden

Glenn Gunzelmann*, John R. Anderson

Department of Psychology, Baker Hall, 342-C, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Received 26 March 2002; accepted 26 August 2002

Abstract

Two experiments using two isomorphs of the Tower of Hanoi show that participants increase the amount of planning they do as they learn that it increases problem solving efficiency. In addition, competition among different approaches emerged as participants gained more experience with the task, with an optimal strategy gradually replacing a less effective, though easier one. It is hypothesized that the competition among approaches is mediated by the costs incurred in terms of the number of moves needed to solve the problems. An ACT-R model of participant performance is used to validate this hypothesized mechanism and to examine many of the details of participants' performance. This model corresponds closely to the observed data, from overall performance in terms of number of moves to details of participants' strategy choices and variability in strategy use among individuals.

© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Problem solving; Strategy; Planning; Learning; Computational model; Tower of Hanoi; ACT-R

1. Introduction

Planning is a fundamental part of problem solving behavior. While it is true that reaching the goal can be as simple as choosing moves one at a time to bring the problem closer to a solution (i.e. hillclimbing), many problems can be solved optimally only by planning longer sequences of moves. Despite the potential benefits of planning ahead, participants in experiments generally do not begin by planning long sequences of moves to solve novel problems. There are two potential reasons for this initial lack of planning. Firstly, in many cases simply understanding and representing the problem is challenging, meaning that working memory resources are not available for doing planning. Secondly, individuals may fail to plan ahead because the actual utility of that extra effort is not clear. In this case, individuals will increase their degree of planning only to the extent that it has a notable benefit in terms of solving the problem. If more planning is rewarded with faster solutions or with solutions that require fewer moves, then it is more likely that the degree of planning will increase.

When planning proves beneficial, people will increase the amount of planning they do until one of

*Corresponding author. Tel.: +1-412-268-8115; fax: +1-412-268-2798.

E-mail address: glenng@andrew.cmu.edu (G. Gunzelmann).

1389-0417/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S1389-0417(02)00073-6

two things happens. Firstly, there may be a limit on how much planning is useful in the problem. That is, there may be a point where more planning does not impact solution efficiency. In this case, the cycle will continue only until that optimal level of planning is reached. Alternatively, they may reach the point where further planning would exceed the limit of working memory capacity. To overcome this internal limit on planning, task-specific strategies need to be formed to reduce the cognitive load of planning ahead.

By closely examining shifts in move sequences as a participant gains experience with a task, inferences can be made about how planning behavior is being extended and refined in order to optimize solutions. These transitions may be fairly abrupt (e.g. Anzai & Simon, 1979), or show a noisy shift to more efficient strategies that produce gradual improvement over a number of trials. Individuals may solve one problem quite successfully, only to produce a less efficient solution to a subsequent problem (Lamaire & Reder, 1999; Reder, 1982; Siegler, 1987; VanLehn, 1991). This latter evidence suggests that participants may be uncertain about how much planning to do in particular situations. It is only through experience and practice that this uncertainty can be resolved. As it is, individuals will tend to plan more optimally based on the current problem situation, resulting in improved performance on a variety of measures (solution time, number of moves, error rates, etc.).

The uncertainty about how much planning to do is resolved by evaluating the utility of the extra planning (i.e. is it worth the effort to plan that much further ahead?). There are two primary criteria on which this evaluation can be made. Firstly, the evaluation can be made in terms of the success at solving problems within the task (Lovett & Anderson, 1996; Lovett & Schunn, 1999). On the other hand, planning more moves may not affect the eventual outcome, but rather lead to a better solution. In this case, the evaluation is in terms of the number of moves or solution time (Lamaire & Reder, 1999; Siegler, 1987). In both cases, the utility of additional planning is evaluated in terms of how much it improves the efficiency of the solution. This process can continue for some time before the individual develops a coherent approach that is sufficient for solving the task optimally.

1.1. Planning in the Tower of Hanoi

The Tower of Hanoi has served as a useful task in problem solving research for a number of years (e.g. Hayes & Simon, 1974; Kotovsky, Hayes, & Simon, 1985). The task itself consists of three pegs upon which are placed any number of disks (three in these experiments). The goal is to change the disk arrangement from some given start state into some particular goal state. There are three rules that constrain movement through the problem space of 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 disk may be moved. The final rule is related to the second, and says that a larger disk may not be moved to a peg where there is a smaller disk.

Research on the Tower of Hanoi has identified the sophisticated perceptual strategy (disk subgoaling) described by Simon (1975), or some variant, as particularly common and successful. This strategy is an instantiation of means–end analysis, and starts with the largest disk out of place. In the event that the largest disk cannot be moved directly to its goal peg, a subgoal is created to move the largest blocking disk out of the way. If this disk is blocked, further subgoals are created until a disk can be moved. Eventually, the original (largest) disk can be moved successfully and the next largest disk out of place becomes the focus of the subgoals and planning. This process is repeated until the smallest disk is placed. Not only is this strategy effective for solving standard problems, but it nearly always results in an optimal solution path, even for nonstandard problems. This provides a good example of a strategy that allows working memory limitations to be overcome. In the standard five-disk Tower of Hanoi problem used by Simon (1975), and Anzai and Simon (1979), 16 moves need to be planned in order to correctly place the large disk. It is highly unlikely that an individual would be able to plan so many moves and maintain them in memory without an overarching strategy to organize them. Thus, disk subgoaling provides a global framework for organizing moves and problem solving behavior.

Kotovsky et al. (1985) studied Tower of Hanoi isomorphs where it appears that planning operates more locally. In the problems they used, two moves were sufficient to place the largest disk. Still, their isomorphs are more difficult than the standard Tower of Hanoi task and even this amount of planning (two moves ahead) was difficult for participants in their study. They concluded that planning emerged as participants developed a better understanding of the task. Because managing the representation of their task was challenging, working memory load was likely quite high as participants began. As this load diminished through experience, participants could plan further in the problem to produce better solutions.

In the research presented here, problems similar to those used by Kotovsky et al. (1985) were used. Thus, we assumed that we would see local planning of the sort described by Kotovsky et al. (1985). However, because we use a computerized graphic presentation, there should not be a very high working memory load, even at the start. We believe that this creates a situation closer to the second alternative discussed in the opening paragraph of this article. Specifically, for our participants planning should emerge as a function of its demonstrated ability to improve problem solving, rather than as a function of their ability to plan moves. As participants discover the utility of planning more moves, they will be increasingly likely to engage in that planning. Research by Svendsen (1991) supports the assertion that the emergence of planning in this task may be related to its utility rather than working memory demands. In his research, he had participants solve graphically presented Tower of Hanoi problems using either a mouse- or a command-driven interface. In the command interface it took much longer to execute a move than in the mouse interface, meaning that incorrect moves were more costly in that condition. Svendsen found that participants using the command interface produced solutions in fewer moves than those using the mouse interface. In addition, when participants were switched from a command interface to a mouse interface, their solutions became less efficient. This indicates that people's tendency to plan in the task depends on how much benefit planning has on the efficiency of the solution. As errors become more costly, individuals are more likely to plan further into the problem to avoid them.

The experiments presented here use two isomorphs of the three-disk Tower of Hanoi. The isomorphs are Paint Stripping and Monster Move (see Appendix A and Fig. 1). In the Paint Stripping isomorph, the disks are represented by layers of paint, and the pegs are represented by pieces of furniture. In the Monster Move isomorph, the globes are the disks and the monsters are the pegs. In the standard Tower of Hanoi, the hierarchical ordering of the disks is represented by size (see the rules described above). In the Monster Move isomorph, the ordering of the globes is also based on size, but the relationship is reversed (the large globe corresponds to the small disk). In the Paint Stripping isomorph, this ordering is represented by the darkness of the paint, with darker shades corresponding to smaller disks (see Appendix A for the rules for these isomorphs as they were presented to participants). Two different isomorphs were used to ensure that the results were not due to some particular feature of the task presentation (Hayes & Simon,

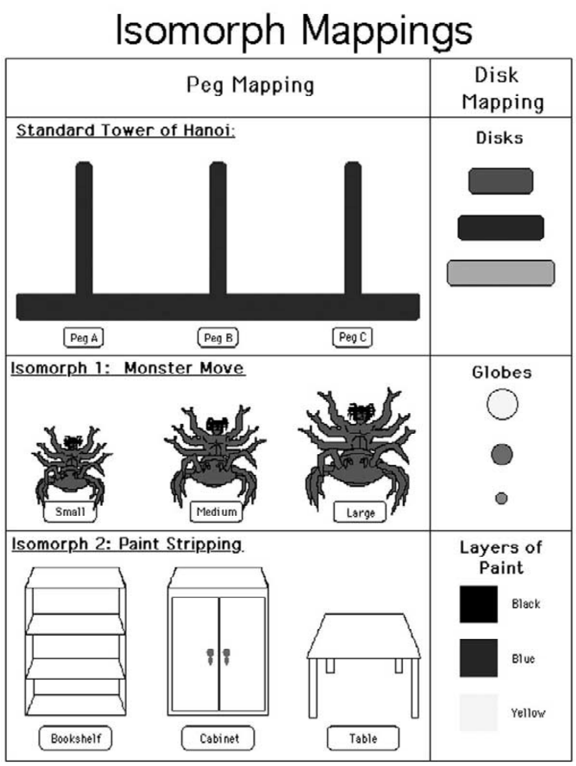

Fig. 1. Mapping of isomorphs used in this study to the standard Tower of Hanoi.

1977), and three-disk problems were used because of our interest in examining local planning. In these problems, planning requirements for placing disks should be within the working memory limits of participants (a maximum of four moves needs to be planned to place a disk).

There were three primary reasons for using isomorphs of the Tower of Hanoi rather than the standard problem. Firstly, the three-disk version of the standard Tower of Hanoi is quite easy (e.g. Gunzelmann & Blessing, 2000), and participants typically require very little practice with it before recognizing what moves to make. Secondly, a high proportion of undergraduates at Carnegie Mellon University have been exposed to the original Tower of Hanoi task, and have learned algorithms for solving it. These factors would likely have a large impact on performance in the experiment. Thirdly, the isomorphs used actually provide more natural cover stories for the particular problems used. In the three-item (i.e. globes or layers of paint) version of these tasks there is a particular class of problem states in which there is one disk on each of the three pegs (flat states; Fig. 2). The problems used here involve moving from one of these flat states to another. Flat states are natural terminal states in both the painting stripping isomorph (since each piece of furniture is painted) and the monster move isomorph (because each monster has a globe). In contrast, in other problems like the original Tower of Hanoi, tower states (all disks on the same peg) are more naturally terminal states.

In the isomorphs, then, there are a total of six flat states, and for each there are exactly two other flat states that are five moves away (shortest path; ‘Start 1’ and ‘Start 2’ in Fig. 2). There is also one other flat state that is seven moves away (‘Start 3’ in Fig. 2). The problems used here involved transforming one flat state into one of the others five or seven moves away. In addition to flat states being natural terminal states, they are interesting because they promote an easier alternative to an optimal planning strategy. Specifically, the emphasis on flat states should give rise to a planning approach based on transforming one flat state into another (each transformation taking three moves), until the ‘correct’ flat state has been

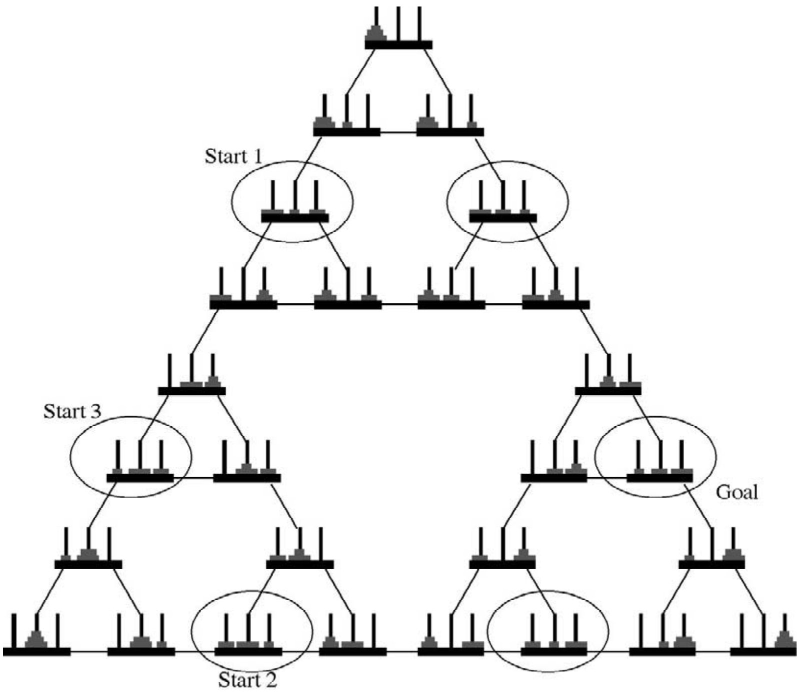

Fig. 2. Problem space representation for the Tower of Hanoi. Flat states are circled.

reached. This strategy is generally successful at producing solutions to the problems but is not optimal. At best, planning in this manner will produce a six-move solution (two sequential flat-to-flat transformations) to the problems that can be solved optimally in five moves. For the seven-move problems used in the transfer phase of experiment 2, three flat-to-flat transformations (nine moves) are needed to reach the solution using the flat-to-flat strategy. We hope to see some competition between the optimal planning strategy and the flat-to-flat planning strategy. While the latter is easier to implement in terms of cognitive effort, the former is more efficient in terms of number of moves.

In order to observe whether local planning emerges gradually, it is useful to follow students as they solve a sequence of problems. In previous studies, participants often received only one or two problems to solve (e.g. Kotovsky et al., 1985). In contrast, these experiments give 12 or 18 problems to each participant. With only a couple of problems, we cannot assess whether a change in solution strategy on a particular trial reflects a permanent change or is just a step along the path of gradual change. These experiments should provide evidence to clarify this issue.

In summary, the experiments presented here were conducted to examine a number of issues relating to planning behavior in problem solving. It is expected that the amount of planning done by participants will increase as they gain familiarity with the task, with solutions becoming more accurate as a result. Because of the nature of the problems, it is expected that this planning will arise in two forms that compete with each other. Specifically, it is expected that a flat-to-flat strategy will emerge in participants' solutions because it generally requires less planning, and also because flat states are salient in the problem. This saliency will make them attractive as states where similarity to the goal state can be evaluated and planning can be done. However, with experience, participants will realize the value of additional planning and focus on the more useful characteristics of the task (i.e. where the globes or layers of paint are) instead of less relevant qualities (i.e. whether or not they are in a flat state). As a result, planning optimal sequences of moves to place globes (layers of paint) should emerge as the dominant approach because it leads to the solution in fewer moves. The presentation of many problems will ensure that the full course of this transition can be examined.

Before discussing the experiments, there is one additional concept that is critical to understanding the research presented here. It has been found that as individuals solve novel problems, two phases are distinguishable in their solutions (Kotovsky et al., 1985). Participants first exhibit an exploratory period in which they make little progress towards the goal state, but then generally produce the solution fairly quickly in the second phase, the final path. These two phases of problem solving are defined in terms of the distance from the goal. The final path consists of the moves made by the individual after being the original distance from the goal for the last time before successfully solving the problem. In experiment 1, participants begin five moves away from the goal state. For these problems, the final path begins the last time the participant is five moves away from the goal before the problem is solved. For the seven-move problems used as transfer problems in experiment 2, the critical distance is seven. Since the exploratory path results in no net progress toward the goal, it is the final path that will be of most interest in these experiments. It is on the final paths that clear indications of successful planning can be found. As participants make their run to the goal, the moves they make and the latencies for those moves will expose what they are planning to do and when they are doing that planning. These data can be used to distinguish among different strategies being used for solving the task. In addition, since this behavior should emerge earlier in participants' solutions as they learn the benefits of planning, the length of the exploratory path should progressively decrease with experience.

2. Experiment 1

The goal of experiment 1 was to investigate how the solutions produced by participants changed as they gained experience solving the five-move flat-to-flat problems. Since these problems can be solved in multiple ways, it is important to determine how experience with the problems affects performance.

This experiment asks participants to solve 12 problems, making it possible to track the emergence of fairly proficient performance on the task. Without training, participants are free to develop any representation of the task that they find useful. As described above, it is expected that a flat-to-flat strategy will appear in solutions as participants recognize that flat states are important in the task. With more experience, however, it is likely that the formal structure of the task will lead participants to more effectively plan moves by focusing more clearly on placing globes (layers of paint) in their goal locations.

2.1. Method

2.1.1. Participants

The participants were 24 undergraduate students (mean age 20.1 years) from Carnegie Mellon University. Participants received either course credit ($n=7$) or \$8 ($n=17$) for their participation in the 1-h experiment. There were 17 males and seven females included in the study.

2.1.2. Materials

Three sets of problems were presented to each participant during the experiment, with six specific problems within each set (18 problems total). The first and third set of problems were either the Monster Move or the Paint Stripping isomorph of the Tower of Hanoi (Fig. 1). Half of the participants were presented with the Monster Move isomorph, while the other half were given the Paint Stripping isomorph. All participants worked with the same isomorph for both sets. The descriptions and rules for the isomorphs are presented in Appendix A. The other task was not directly related to the Tower of Hanoi, but was used to break up the sequence of problems. This filler task was 'Building Sticks' (Lovett & Anderson, 1996), an isomorph of Luchins' (1942) water jug task. All of the tasks, in addition to all instructions for the experiment were presented on Macintosh computers and were completed using only a mouse.

There were three globes in the Monster Move isomorph and three layers of paint in the Paint Stripping isomorph (see Fig. 1 for the mapping of these isomorphs onto the standard Tower of Hanoi). The problems completed by participants were drawn from the set of 12 possible flat-to-flat problems requiring five moves to solve. Two groups of six problems were created from the 12 possible problems based upon how many moves were needed to place the large globe or layer of paint (two or four; see Section 4 for a more complete description). Participants completed all six problems from one of these groups for each problem set in random order. In the first problem set, each group of problems was given to half of the participants. In the second set, half of the participants switched to the group of problems they had not seen, while the other half received the same set of problems again. The different groups of problems did not significantly effect the number of moves needed to solve the problems, $F(1,46)=1.23$, $p>0.25$, so the data below are aggregated over this factor. The differences between these problems will be considered in more detail in Section 4. Also, the data indicated that there were no overall differences between the isomorphs, $F(1,22)=0.01$, $p=0.91$ for overall number of moves. In addition, there was no indication of an interaction between the isomorphs as a function of problem number for number of moves, $F(11,242)=0.81$, $p=0.63$. As a result, the data presented below are collapsed across that factor.

2.1.3. Procedure

Participants completed the entire experiment on a computer. For each task participants were presented with a cover story, a set of rules, and an explanation of the interface. They were able to move back and forth among these screens (using 'continue' and 'back' buttons) for as long as necessary before actually beginning to work on the problems. However, once they began working on the problems, they were unable to return to the description, rules, or interface screens. They were instructed to solve each problem for each task by reaching the goal state that was presented on the screen.

The isomorphs were presented in a graphical display, using the images shown in Fig. 1. The labels below each monster and piece of furniture are buttons that were used to execute moves. A move was made by first clicking on the button below the source monster (piece of furniture). This selected the largest globe (darkest layer of paint) in that location.

This inherently enforced the first two rules for the tasks, since only a single item was selected and it was constrained to be the top item in the hierarchy. The move was completed by clicking on the button below the destination monster (piece of furniture). If the move was legal, the selected globe (layer of paint) was moved to its new location. If the move was not legal, a message box appeared that restated rule 3 and pointed out that the attempted move was illegal. This allowed participants to review the rule that had been violated. At that point, the globe (layer of paint) was deselected, and the participant was able to begin his or her move anew. After each problem, a message box appeared indicating that they had solved it correctly. The same procedure was followed for each of the sets of problems.

2.2. Results and discussion

The problem state and move latency were collected at each point in the solutions, allowing for a careful examination of how participants solved the problems. Following Kotovsky et al. (1985), we partitioned the data into an exploratory portion and a final path (Fig. 3). As can be seen, most of the reduction in the total number of moves comes from a reduction in the length of the exploratory path. For the final paths, recall that if this final path was being planned by the flat-to-flat strategy it should be six moves long, while optimal planning would make it five moves long. So, if these local strategies were used most of the time, it should make sense that final path length wouldn't decrease that much over the course of the 12 problems. In fact, 74% of the final paths were of length five and 19% were of length six (Fig. 4). If participants were choosing randomly, only 11.1% of the final paths would be optimal and only 3.7% would be six-moves long. Fig. 5 contrasts the latency profiles for these two types of final paths. The six-move final paths show long latencies at the first and fourth move where three-move flat-to-flat sequences are being planned. The five-move final paths show an elevated latency for the first move where the bulk of the planning should be done, and a smaller increase for the third move where planning sometimes needs to be done (depending on whether a two-move or a four-move sequence was planned initially; this distinction is mentioned above and will be discussed further in Section 4).

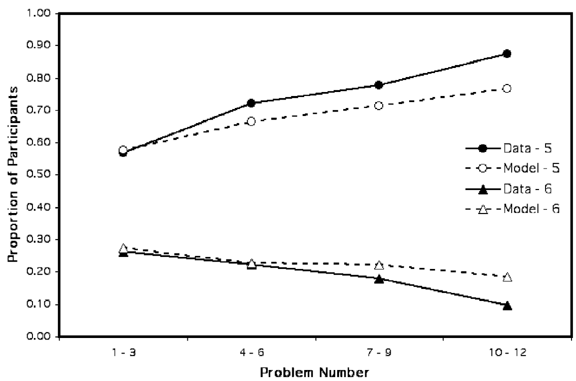

Fig. 4. Proportion of participants that solved the problems with a five-move or six-move final path in experiment 1.

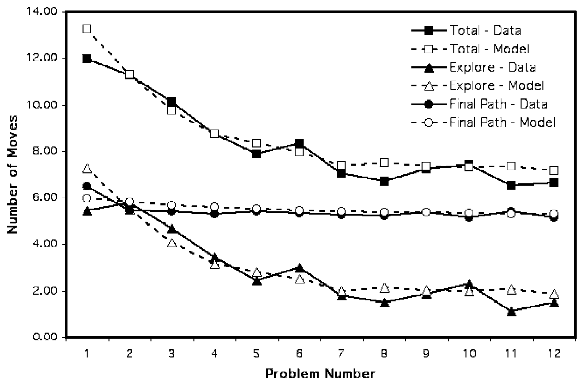

Fig. 3. Number of moves needed to solve the problems in experiment 1.

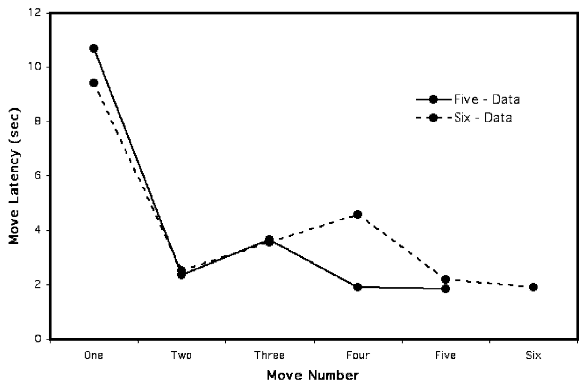

Fig. 5. Move latencies for five-move and six-move find paths.

Experience with the task should have resulted in improvements in two aspects of participant solutions. Firstly, the length of the exploratory path should decrease, meaning that successful planning is being done earlier in the solution process. Indeed, an ANOVA shows that the length of the exploratory path did decrease significantly over the 12 problems, $F(11,253) = 2.69$, $p < 0.01$. This effect is reflected in a significant linear trend of decreasing exploratory path length, $F(1,253) = 23.83$, $p < 0.001$. Besides the exploratory paths, it was also expected that experience with the task would result in an increase in the number of five-move final paths with a corresponding decrease in the number of six-move final paths. In line with these predictions, Fig. 4 plots the proportion of problems incorporating a five-move final path and the proportion with a six-move final path as a function of trial (averaged into quartiles). The data provide evidence that final path length decreased as participants gained more experience with the task, $F(1,253) = 23.35$, $p < 0.01$ for a linear trend.

The final path data provide evidence that flat-to-flat planning was being done in the experiment. In addition, there are other data that support the conclusion that participants treated flat states as somehow special in the task. Firstly, if moves were made entirely at random, it would be expected that participants would arrive at flat states every 4.5 moves (six of 27 states are flat states). However, the rate was actually every 3.42 moves (3.41 to 3.43 with 95% confidence) for participants (excluding moves made on optimal final paths), with the minimum distance between flat states being three moves (Table 1). So, when participants were not planning optimally, they were executing move sequences that brought them to flat states more often than would be expected by chance. Secondly, throughout the experiment move latencies for flat states were over twice as long as move latencies for other states (this excludes the latencies for the first move in each problem), suggesting that more planning was indeed occurring in these states. In fact, for every participant, move latencies were greater for flat states than for other states. This suggests that flat states were indeed a ‘home base’ of sorts where participants could check on progress toward the goal and consider alternatives.

Table 1
Average moves between flat states in both experiments (excludes moves on optimal final paths; values in parentheses are model predictions)

Experiment	Average moves between flat states
Experiment 1	3.4 (3.7)
Experiment 2 (training)	3.4 (3.7)
Experiment 2 (transfer)	3.5 (3.8)

Collectively, the data from experiment 1 provide evidence for two distinct local strategies. The combination of the two proposed strategies account for a large percentage of the problem solutions. Specifically, over 90% of the final paths produced by participants in this experiment were either five or six moves long, with latency profiles matching those predicted by the flat-to-flat and optimal strategies. It is not the case that all of the problem solutions fit into these designations, and a certain amount of apparent randomness is present in the data as well. However, the fact that the data on the whole are so well-explained with only these two strategies provides support for the conclusion that they are being used a great deal by participants in this experiment. In Section 4 these strategies and the transitions among them will be examined more closely.

3. Experiment 2

The findings from experiment 1 provide evidence that participants were using the two anticipated strategies while solving the problems. We decided to do a second experiment to replicate the first experiment, but without a task intervening between the first and second set of six problems. After completing 12 of the five-move problems (thereby replicating the first experiment) we had students do six transfer problems that were more difficult. These were the seven-move problems mentioned in Section 1. While seven moves is an optimal solution, it would take nine moves to reach the solution using the flat-to-flat strategy exclusively. Therefore, the emergence of nine-move final paths would be further evidence for a flat-to-flat strategy. But in these problems the flat-to-flat strategy is at a greater disadvantage to optimal planning, now taking two extra moves. Therefore, because of the greater cost of the flat-to-flat strategy, we expect to see participants abandon it during the transfer phase of the experiment (the seven-move problems). That is, the savings in planning obtained by using the flat-to-flat approach will be overcome by the extra time needed to actually solve the problems, making it a less attractive option than planning a little further ahead.

3.1. Method

3.1.1. Participants

The participants in this study were 60 CMU undergraduate students (mean age 19.8 years) enrolled in a psychology course. There were 42 males and 18 females in the sample. All received course credit for their participation.

3.1.2. Materials

A program similar to the one used in experiment 1 was used to administer the experiment to participants. The entire experiment was conducted on the computer and the program handled all data collection.

3.1.3. Procedure

Participants were randomly divided into two groups based upon cover story (Paint Stripping versus Monster Move). The tasks were presented in the same manner as in experiment 1, and each participant solved 18 problems from either the Monster Move or Paint Stripping isomorph. The first 12 problems were the five-move problems used in experiment 1. The same groups of problems described for experiment 1 were used in experiment 2. In this experiment, however, all participants completed the same group of problems twice during the 12-problem training portion. Again, the order of trials was random. The last six consisted of the complete set of possible seven-move transfer problems, presented in random order. The transition between the two types of problems was not indicated explicitly to the participants. The two types of problems did not produce a significant effect on performance, $F(1,58) = 1.48$, $p > 0.20$. Also, the pattern of data was identical for both isomorphs, $F(1,58) = 2.32$, $p = 0.13$ for a main effect on number of moves, and $F(17,986) = 1.21$, $p = 0.30$ for an interaction between isomorph and problem number for number of moves. Consequently, the data presented here are combined across those factors.

3.2. Results and discussion

The data for the number of moves needed to solve the problems in experiment 2 are presented in Fig. 6. The data from the training portion of this experiment are very similar to the data from experiment 1, providing additional support for the description of the planning behavior exhibited by those participants. Here, 74% of the training problems involved a five-move final path and 19% of the solutions had a six-move final path that involved two sequential flat-to-flat transformations (Fig. 7). So, combining

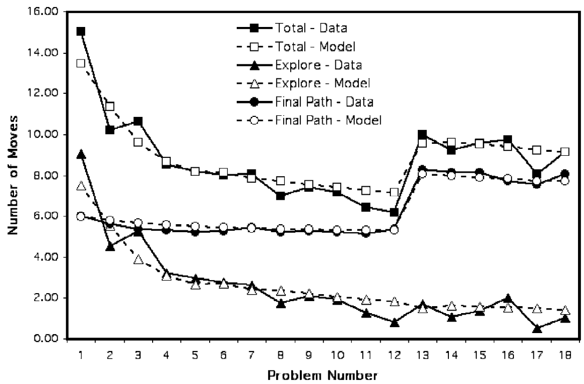

Fig. 6. Number of moves needed to solve problems in experiment 2.

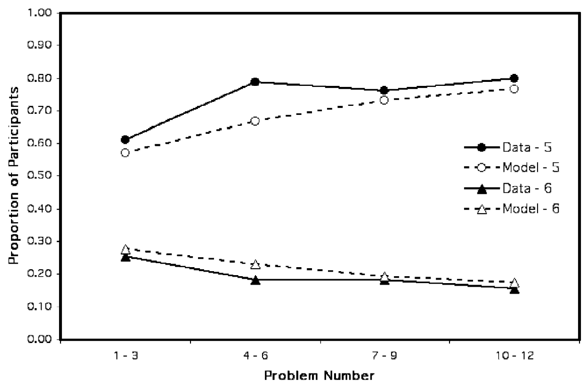

Fig. 7. Proportion of participants who solved the problems with a five-move or six-move final path in the training portion of experiment 2.

the evidence for the two hypothesized planning approaches, it is possible to account for over 90% of the final paths produced by participants in the training portion of experiment 2. These data indicate that problem solutions improved over the 12 problems both in terms of exploratory path length, $F(1,649) = 86.04$, $p < 0.001$ for a linear trend (Fig. 6), as well as final path length, $F(1,649) = 31.16$, $p < 0.001$ for a linear trend (Fig. 7). In addition, flat states were visited more frequently than would be expected by chance (Table 1), and move latencies were again over twice as long for flat states than for other states.

It is possible to go on to analyze the transfer data in a similar manner to determine how participants solved those problems. In the transfer phase, participants continued to spend longer deciding on moves while in flat states and still entered flat states more frequently than would be expected by chance (Table 1). In terms of final paths, these problems require two additional moves to solve, meaning that final paths of seven moves are optimal. Meanwhile, as described above, it takes nine moves to produce a solution to these problems using the flat-to-flat strategy. However, participants also produced some eight-move final paths, which seem to involve a combination of the two approaches. Specifically, a flat-to-flat transformation can be used to get to a flat state where less planning is needed to place the small globe or yellow layer of paint (two versus four moves away). Some solutions involved this transformation followed by an optimal sequence of five moves from that point. Others involved optimally placing the small globe or yellow layer of paint (four moves), moving to a flat state, and then reaching the solution with a flat-to-flat transformation. These two types of eight-move final paths seem to reflect two ways in which flat states were important in the task. The first type seems to illustrate intentional use of flat-to-flat transformations as a means of simplifying the planning necessary to place globes (layers of paint). The second seems to more clearly reflect the attractiveness of flat states in the experiment. In this case, it seems as though that the saliency of flat states caused participants to seize the opportunity to move into one if they were not sure of what to do next. By using solutions that followed the predictions of the flat-to-flat strategy and optimal planning, 90% of the final paths can be explained for the transfer phase of experiment 2 (final paths of seven, eight, or nine). These data are shown in Fig. 8, which also shows that the final paths in the transfer problems became shorter over the course of the six problems, though this trend was only marginally significant, $F(1,295) = 3.35$, $p < 0.10$ for a linear trend.

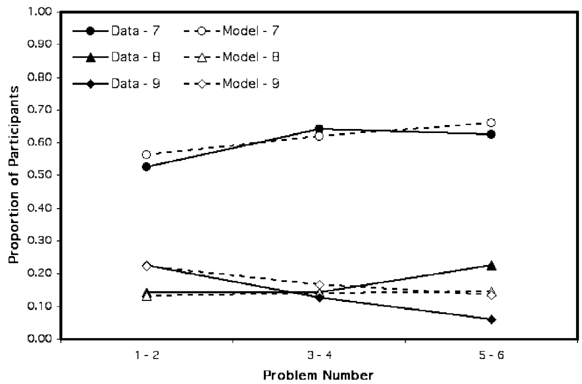

Fig. 8. Proportion of participants who solved the problems with a seven-move, eight-move, or nine-move final path in the transfer portion of experiment 2.

What follows is a description of a cognitive model, developed in ACT-R 4.0, which accounts for participants' data, both in terms of overall performance and strategy use. The model described here can be accessed online at the ACT-R website (http://act-r.psy.cmu.edu/).

4. ACT-R model

The explanation of the findings in these two experiments seems straightforward. With experience, participants gradually increased their tendency to plan longer sequences of moves, which led to decreases in the length of the exploratory path. In addition, final path length improved as two local planning strategies competed against each other over the course of the problems, with participants gradually switching to the more efficient strategy over the one that was easier to execute. However, it is uncertain whether this explanation can account for all of the data from this experiment. If it can, it is important to understand how these findings may fit with others in the literature on strategy choice in problem solving. To clarify these issues, an ACT-R model was created. By instantiating the explanation in a running model, clear predictions can be made and their relation to the experimental data examined. In addition, the model uses the same learning mechanism used by other researchers to explain strategy learning (e.g. Lovett, 1998; Matessa & Anderson, 2000), and so the model and the explanation fit into the larger literature on strategy choice and learning.

ACT-R is a theory of human cognition that has been instantiated as a running computer simulation (Anderson, 1993; Anderson & Lebiere, 1998). Fundamental to the ACT-R theory is a distinction between declarative and procedural knowledge. Declarative knowledge holds specific facts and information (chunks), such as '$3+4=7$' or 'the small disk is on the rightmost peg'. Procedural knowledge, on the other hand, is the storage for cognitive operators. This knowledge is represented as production rules, or condition–action pairs that specify what action to take when a particular condition is satisfied. So, there may be productions for retrieving a specific chunk from declarative memory or for making a move in the Tower of Hanoi.

At all times, ACT-R maintains a goal and the behavior of the system is directed at working on that goal. This is accomplished in ACT-R through a series of cycles. In each cycle, the current state of the goal is used to identify applicable productions in memory, and a single production from this set is chosen. This production is then executed or 'fired'. Two major types of changes may be made to the goal on each cycle. Either its contents may be changed, or a new goal may become the focus. The latter is achieved by (1) satisfying the current goal (popping on success), (2) giving up on the current goal (popping on failure), (3) creating a subgoal based upon the current goal (pushing), or (4) selecting a new goal to focus on (essentially a pop followed by a push). Through a series of such cycles, ACT-R can produce behavior in many domains that is human-like in a number of important ways (see Anderson & Lebiere, 1998 for a review).

To support the interaction of these symbolic aspects of ACT-R, there are a number of subsymbolic mechanisms that allow it to learn from experience. In declarative memory, these mechanisms relate to the activation of chunks and affect how accessible they are. Since all of the information for this task was always available on the display in these experiments, these activation mechanisms do not play a role in the performance of the model presented here. There are two subsymbolic mechanisms relating to two aspects of procedural knowledge, both of which measure the expected usefulness of each production. Firstly, ACT-R contains parameters that allow the anticipated cost of using a production to be estimated. The value of this quantity essentially answers the question, how much time is likely to be spent achieving the goal from this point if this production is used? The other set of parameters answer the question, how likely is it that the goal will eventually be achieved if this production is used? In the context of the model presented here, the mechanism relating to cost is critical. This aspect of ACT-R is discussed in more detail below.

4.1. Model design and mechanisms

There are two important aspects of ACT-R for the model presented here. Firstly, the model's performance is based largely on the process of selecting a single production from among several applicable alternatives. In ACT-R, this choice is controlled by the calculation of a quantity called 'expected gain' (E). This quantity is calculated for each production on each model cycle and is an estimate of how useful that production is expected to be for achieving the current goal. It is negatively associated to cost in this situation. The production with the highest value of E (or lowest cost) is the one that is chosen on each cycle and fired. The second aspect of ACT-R, utility learning, adjusts the values of E as the model accumulates experience. Each time the goal is popped, the parameters are updated for each production that was used. As experience accumulates, values of E will change according to the experience of the model. This will affect the likelihood that each production will be used, and the model's behavior will change accordingly. The details of these mechanisms and their application in the current model are described below, but first we will describe the options from which the model was selecting.

Based on the evidence gathered from participants, it seems clear that participants were using both optimal planning and flat-to-flat planning. As a result, the ACT-R model performs the task basically by selecting from among these options and evaluating how well they work. However, since there were a number of moves that did not follow any identifiable pattern, the model also can execute random moves as it works toward a solution. This is essentially making moves without planning at all. It is important to note that the model operates on local subgoals related to correctly placing individual globes (layers of paint). That is, the model focuses on placing a single globe (layer of paint), and then shifts its focus to another one that is out of place. The utility of planning in the model is evaluated on the basis of how quickly the focus globe (layer of paint) is placed, not how quickly the whole problem is solved.

The model performs the task in the following way. When a new problem is presented, it is first encoded. The model then focuses on moving the small globe (yellow paint) from its current location to its goal location. When that is accomplished, focus is shifted to the medium globe (blue paint). Because the model focuses on just one globe (layer of paint) at a time, it is possible that the small globe (yellow paint) may be displaced in the process of placing the medium (blue) one. If this occurs, the model simply returns to the small globe (yellow paint) after the other is placed. Once both items have been placed, the single move needed to solve the problem is made. At each point in the problem where the focus globe (layer of paint) is correctly placed onto its goal peg, the parameters in the model are updated (see below).

At almost any point in the problem, the model has three options. The first option is that the model can execute a random move. If the model chooses to execute a random move, it simply selects one of the legal moves available and executes it. While this approach minimizes planning, it tends to produce very poor solutions. The second option is that the model may instead choose to do the planning necessary to correctly place the current focus globe (layer of paint). This can involve from one to four moves, and the likelihood that the effort will be made will vary as a function of how much planning needs to be done. If this option is chosen, the required moves are executed. The third option is to execute moves based on the flat-to-flat strategy. If the current state is a flat state, the model can choose to execute a flat-to-flat transformation, producing a new flat state that is closer to the goal state. If the model is not in a flat state and not in a tower state (all three globes or layers of paint in the same place), it can choose to make the single move that will take it to the nearest flat state. When the model finds itself in a tower state, there are no flat-to-flat moves available. So in tower states the model must either do the planning needed to place the globe (layer of paint) or make a random move. After choosing an approach and making the prescribed move or moves, the model checks to see if the focal globe (layer of paint) has been placed. If it has, the smallest globe (lightest layer of paint) that is still out of place is selected as the focus. If not, it continues to work on the same one. Either way, the model again selects an approach and executes it. This cycle repeats until the problem has been solved.

The critical juncture in the operation of the model lies in the selection of which action to take. There is a production for each of the possible actions (see Table 2). At each point when a decision needs to be made only three of these productions will apply, one for each option (only two will apply in tower states). Productions that do not apply in the current situation will not be considered. The choice of which of the applicable productions fires in ACT-R is determined by the calculation of E for each production. In this calculation, the two subsymbolic quantities relating to procedural knowledge (anticipated cost and probability of achieving the goal) are calculated for each production to represent how beneficial the production's use is expected to be in terms of achieving the goal of placing the focal globe (layer of paint) in its goal location. The production producing the highest value for this quantity is selected and fires. The equation for expected gain (E) in ACT-R is:

$$E = PG - C + \text{noise},$$

where P is the probability that the goal eventually will be achieved if the production is used, C is the anticipated cost (in seconds) of ultimately 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). Because the

Table 2
Initial parameter settings for the choice productions in the model

Production	History (H)	Total cost (TC)	Cost (C) (TC/H)	E ($PG - C$)
Random:				
Random-move	200	1400.00	7.00	43.00
Flat-to-flat options:				
Move to a flat state	2	14.00	7.00	43.00
Place item—flat to flat	2	17.00	8.50	41.50
Ease planning—flat to flat	2	21.00	10.50	39.50
Optimal planning:				
Place focus item—1 move	2	14.00	7.00	43.00
Place focus item—2 moves	2	15.50	7.75	42.25
Place focus item—3 moves	2	17.00	8.50	41.50
Place focus item—4 moves	2	18.50	9.25	40.75

focus globe (layer of paint) always arrives in its goal location eventually, the probability of success is always 1 ($P = 1$). The value of G was set at 50 (s) in this model.[1] However, since P is equal to 1, the particular value does not really impact the model's performance.

The selection of less effective approaches will result in more moves being made on average before the focus globe (layer of paint) is placed. This makes the cost parameter vital in determining how likely the model is to engage in planning. The cost parameter is an estimate of the total time spent from the firing of that production until the goal is achieved. The model keeps a record of all the costs so far and calculates an average cost as its estimate of C:

$$C = \text{total efforts}/\text{history},$$

where total efforts is the sum of all past costs and history is the number of past experiences. Table 2 shows the initial values that were set for each of the productions used in choosing how to solve the problems (Table 2). These values reflect an anticipation of how long the actions are likely to take. They were set so that the initial expected cost increased linearly with the number of moves involved in the approach to be implemented:

$$C = 6.25 + 0.75 \times \text{moves}.$$

The lone exception to this is the production that chooses to execute a flat-to-flat transformation to get a larger globe (darker layer of paint) out of the way (called 'Ease planning' in Table 2). This transformation produces a state where only two moves are needed to place the small globe (yellow layer of paint), whereas four moves would be required to place it before this transformation is made. This application of the flat-to-flat strategy goes against simpler problem solving heuristics (i.e. hillclimbing). That is, when in a flat state that is seven moves from the goal, the large globe (black paint) is in its correct location. So executing a flat-to-flat transformation at that point requires that it be moved away from its goal location. In addition, the resulting flat state has no globes (layers of paint) in their goal locations, making it superficially seem further from the goal. Due to these features, it seems to have been relatively uncommon in participants' approaches and so it was given a higher expected cost which results in a lower likelihood that it will be used.

In addition to cost, each of the critical productions was given a history of successes. This quantity controls the stability of E by influencing how much impact a single use of the production will have. More prior experience reduces the impact of each single use of the production. For all except the random-move production, this value was set at 2, allowing these values to change rather quickly. This

[1] This value is traditionally set at 20 s in ACT-R. However, these problems take longer than that for the model (and participants) to solve. This value was raised to accommodate this fact.

means that the values of C for these productions will quickly come to accurately reflect the actual cost of using them to solve the problems. The separate value that was estimated for the random-move production (200) seems to reflect a rather strong reluctance by participants to exert the effort needed to do any planning at all. As experience accrues, the model obtains information about how much time is needed to place a globe (layer of paint) when each of the available approaches is used. As a result, it becomes more likely to plan longer sequences of moves since planning produces more efficient solutions. As it will also take more moves using flat-to-flat transformations to get the globe (layer of paint) in place, the model also tends to abandon that approach in favor of optimal planning behavior that minimizes the number of moves needed to solve the problems.

The initial settings for these parameters are presented in Table 2, along with the calculated values of E that these settings produce. The actual effort (time) required by the model to execute each move was set to 4 s, based on participants' data. So, the model takes 4 s to make a random move, 8 s to plan and execute two moves to place a globe (layer of paint), 12 s to carry out a flat-to-flat transformation, and so on.

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, noise was set such that the value is randomly selected from a distribution with a mean of 0 and a standard deviation of 1.81 s (or about one-half of a move). 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 from among the applicable alternatives. A major assumption of the model is that most participants would be unlikely to plan ahead very much when they began the experiment. In the model, this is instantiated in the initial values of E for the choice productions. In particular, the initial values of E are lower for approaches that require more planning. Thus, the model generally begins with simple approaches (random moves, placing a globe (layer of paint) in a single move when that option is available, etc.) and moves toward more effective ones (i.e. planning up to four moves in order to place a globe or layer of paint into its goal location).

When a new globe (layer of paint) becomes the focus, a number of actions may be taken before it is placed. Some actions (e.g. a random move) may be taken more than once. When the globe (layer of paint) is finally placed, the parameters for the productions that select actions are updated accordingly. Firstly, for each time the production was used, the history of that production are incremented by one. To update the cost, the time spent placing the globe (layer of paint) is calculated from each point where an action was selected. This total is added to the total efforts, which is then divided by the history to give a new value for C (see equation above). This may be best illustrated using a quick example. If, at the beginning of the problem, the model makes two random moves, it can arrive at a state where three moves are needed to get the small globe (yellow layer of paint) to its goal location. If the model then decides to plan those three moves, the globe (layer of paint) will be placed correctly in five moves (20 s). In this case, the random-move production will have two events added to its history (it was used twice in this scenario) and will have 36 added to its cost parameter (20 s from the first use plus 16 s from the second). Meanwhile, the production that chooses to plan three moves will have its history increased by 1, and its cost increased by 12. In fact, every time that the plan-three production is used it will incur a cost of 12 and will succeed in placing the focus globe (layer of paint). In contrast, the random-move production can experience widely varying costs, and over time this cost is likely to be quite high on average. As the model gains experience it learns that planning is worthwhile because it gets the globe (layer of paint) to its goal location more quickly. Random moves are quite ineffective, while the flat-to-flat approach consistently requires one or two extra moves. Over time, these experiences will lead the model to plan optimally more often.

The initial parameter values (Table 2) were set to match the aggregate move data (Figs. 3 and 6). At this level, the model's performance corresponds quite closely to the data from the participants. In particular, for experiment 1 the correlation between the model and the data is 0.96 (RMSD = 0.57 moves). For experiment 2 the results are similar (correlation = 0.95; RMSD = 0.71 moves). In terms of strategy use at the most abstract level, the model produces comparable data for its flat-state visitation rates (Table 1). Since the actual exploratory moves and final paths produced by participants are more indicative of which strategy they were using, these measures were used as indicators of how closely the model was matching participant performance. These data are presented next.

4.2. Model's fit to exploratory and final paths

The progressive decrease in exploratory path length in the data is closely matched by the model (Figs. 3 and 6). As the model becomes more likely to plan, it follows that planning will tend to emerge earlier in problem solutions. The result is that the model gets on the right track (final path) more quickly as it gains experience, just as the participants do. The second experiment makes this point even more clearly. When participants switch to the transfer problems, exploratory path length continues to decrease, even though overall solution length becomes longer. Since much of the decrease in solution length is produced by a decrease in exploratory path length, it seems that much of the improvement in performance over problems is produced by a reduction in random search.

Looking more closely at the final path data, it is possible to infer what strategies were being used. Optimal final paths are assumed to be the result of optimal planning. Again, support for this conclusion comes from both the accuracy of those solutions as well as from the move latencies produced in these solutions (Fig. 5). For those solutions that did not incorporate a five-move final path, other strategies are likely to have been used. Six-move final paths can be described as using a pair of flat-to-flat transformations executed in sequence. Fig. 4 presents the evidence for the two approaches for experiment 1, illustrating the overall trend toward better solutions with practice. As can be seen, the model does a good job of reproducing both the overall tendency to use each of the strategies, as well as shift that which occurred over the course of the experiment (correlation = 0.996; RMSD = 6.0%). A similar analysis can be made for the first 12 problems in experiment 2 (Fig. 7; correlation = 0.993, RMSD = 5.1%). It is important to note that the proportion of solutions that do *not* fall into these categories is quite small, illustrating that the two strategies account for the vast majority of the data in this respect.

The transfer problems can be interpreted in an analogous way, although the minimum number of moves to solve them is seven rather than five. In these problems, the model continues to make quite accurate predictions in terms of final path length (Fig. 8; correlation = 0.981, RMSD = 4.3%). As can be seen, final path length by the transfer phase is fairly stable across the six problems. The most obvious change across problems is the decrease in the number of nine-move final paths. As indicated above, the flat-to-flat strategy adds two moves to the transfer problems, as opposed to the single move added to the other problems. This extra cost increases the advantage of planning ahead.

Taken together, the model fit to the data from both experiments provide evidence that participant performance in this task can be adequately explained using two straightforward strategies, plus some random exploration. In addition, one complication concerning the five-move problems provides an interesting test of the model. Recall from above that there are two flat states that are five moves away from any given flat state. A careful examination of these two problems reveals that they are not equivalent in a strict sense. For one of those flat states, the small globe (yellow layer of paint) can be placed in two moves (*easy*; 'Start 2' in Fig. 2), while this requires four moves from the other flat state (*hard*; 'Start 1' in Fig. 2). Assuming that participants were focusing initially on the small globe (yellow layer of paint), it is the case that the hard problems require more planning to solve optimally than the easy problems. In turn, those problems should have proven more difficult for participants to solve in the experiment. As it turns out, a quarter (6) of the participants in experiment 1 solved only hard problems and a quarter solved only easy problems, while half (30) of the participants in experiment 2 received only hard training problems and the other half received only easy training problems. Fig. 9(a) shows that participants solving only hard problems tended to require more moves to reach the goal state than participants solving only easy problems. Although this effect was only marginally significant in the data, $F(1,68) = 2.97$, $p = 0.09$, the model produces an effect that is similar in magnitude. This effect arises in the model because of the reluctance to plan longer sequences of moves. When the model solves easy problems, it can get away with planning only two moves at a time.

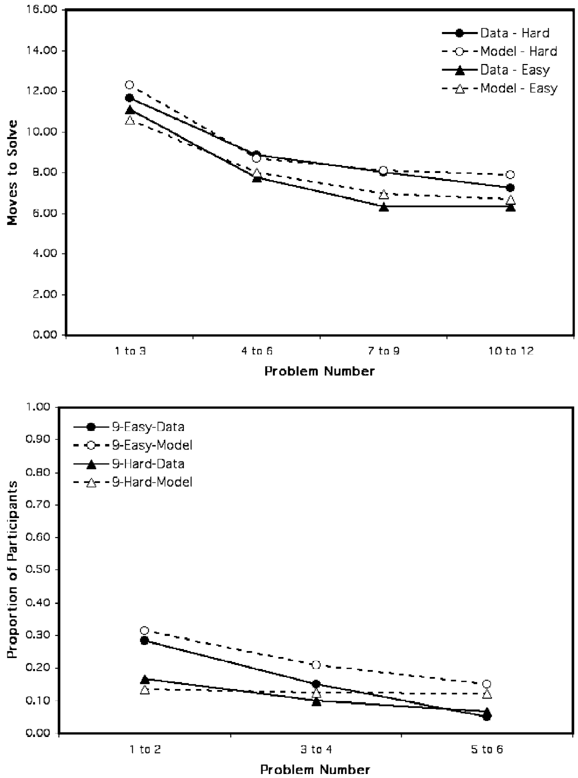

Fig. 9. Evidence for the differential difficulty of the two types of five-move problems used in experiments 1 and 2.

This is not true of the hard problems. So, since the model is less likely to plan four moves than it is to plan two moves, it has a greater tendency to make erroneous moves while solving the hard problems. This results in the hard problems requiring somewhat more moves to solve than the easy problems.

This effect becomes more interesting when the transfer problems in experiment 2 are considered. The two different training conditions produce an interesting effect in the transfer problems. Specifically, participants who had been solving easy problems during training are initially more likely to produce nine-move final paths than those who had been solving hard problems. This difference decreases over time, as participants in the easy training condition produce progressively fewer nine-move paths as they gain experience (Fig. 9(b)). Participants in the hard training condition were able to learn the value of planning four moves at a time while they were solving the training problems. They needed to do this in order to solve those problems optimally. Thus, when they began the transfer problems, they were more prepared to engage in the necessary planning (placing the focus globe or layer of paint in four moves), making it less likely that they would fall back on simpler approaches (i.e. the flat-to-flat strategy). In contrast, those who had been solving easy problems only needed to plan two moves at a time during training (placing the focus globe or layer of paint in two moves). Initially they were not as likely to plan the four moves necessary in order to solve the transfer problems optimally. As a result, they more frequently used the flat-to-flat strategy initially. However, their experience with the transfer problems gradually changed utilities in favor of extra planning, and their final paths improved throughout that phase of the experiment.

Collectively, the data suggest a gradual improvement in performance throughout the experiment. However, it is possible that this gradual improvement is produced by having different participants make abrupt improvements at different times. This is an instance of the general question of whether learning occurs as abrupt, all-or-nothing transitions or as gradual, noisy shifts in performance. As described above, the model's performance is based on the latter assumption. The important question is whether the human participants bear out this prediction. One measure used to address this is a backwards learning curve (Fig. 10; Bower & Trabasso, 1963). This measure is obtained by first identifying the *last* error made by participants. In this case, it is

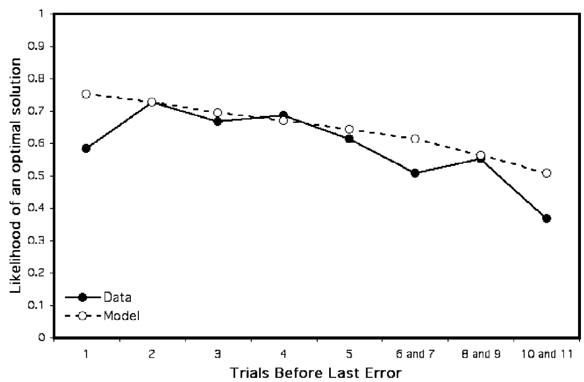

Fig. 10. Backwards learning curve for the five-move problems in both experiments.

considered to be the last non-optimal final path. Then, for each of the previous trials it is possible to determine if the final path was optimal or not. If learning were all-or-none, the likelihood of an optimal final path before the last error should not change as a function of the number of trials before the last error (see Bower & Trabasso, 1963 for a full explanation). However, if participants are gradually increasing their tendency to complete the task optimally as the model predicts, the likelihood that an optimal solution was produced just before the last error should be greater than the likelihood that an optimal solution was produced many trials before. Fig. 10 plots the probability of an optimal final path as a function of the number of trials before the last non-optimal final path, showing that the model's predictions are supported by the data. This provides further evidence that the model is solving the problems in a manner that is quite similar to the participants themselves.

4.3. Variability among individuals

The results presented so far have been the average performance of the model and the participants. However, there was some variation in performance among individuals in the experiment. These differences relate to their propensity to plan in order to solve the problems. This variability can be seen in the lengths of the final paths they produced during the experiments. Specifically, participants differed in terms of how many problems they solved optimally with five-move final paths versus non-optimally with six-move final paths. It is possible to examine the variability of the model on this same measure. For experiment 1 and the training portion of experiment 2, the model captures relatively well the diversity of strategy use in terms of final paths (Fig. 11). Of course, the model shows a much smoother distribution in strategy variation since the data for the model are based on 2000 model runs, whereas there were only 84 participants in the two experiments. In addition, there were some participants who produced optimal final paths for all of the problems, while the model almost never does this. These data suggest that some of the participants were somehow different in terms of the knowledge they brought with them to the task. Indeed, after the experiment some of the participants indicated that they recognized that the tasks were isomorphs of the Tower of Hanoi. Making this connection early on would provide them with information that could influence their planning behavior. The comparison in the transfer phase of experiment 2 simply relates to final paths of seven, eight, or nine (Fig. 12). These data also show that the model does a fair job of capturing the range of participant behavior. The same caveat exists to a lesser extent in these data, with more participants than predicted producing optimal final paths for all six problems.

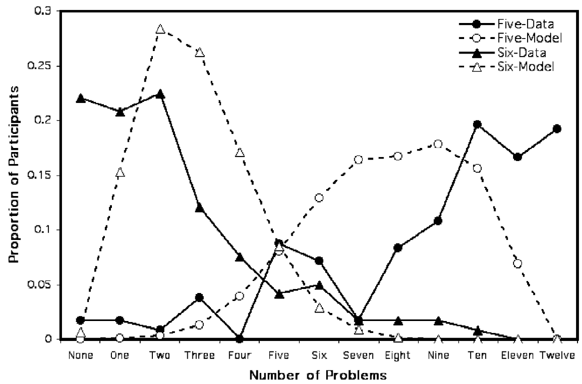

Fig. 11. Proportion of participants solving problems with a five-move and six-move final path with given frequencies for the five-move problems in both experiments.

So, it seems that the model generally does a rather good job in producing variability in performance in comparison to the participants. Within the model, the

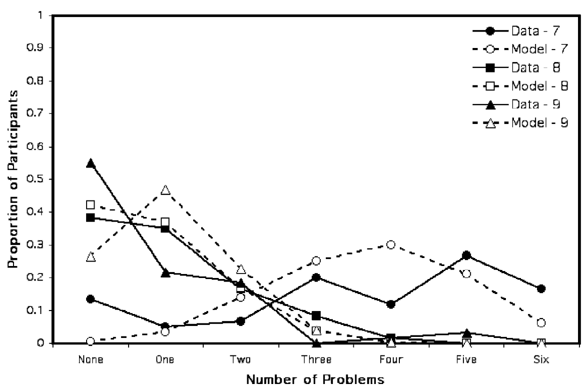

Fig. 12. Proportion of participants solving problems with seven-move, eight-move, and nine-move final paths with given frequencies in the transfer portion of experiment 2.

noise is responsible for the variability in performance, producing both the rough transition in strategy use as well as differences in performance among individual model runs. Thus, the tendency to switch strategies can be seen as a function of when the strategies are used in the process of solving the problems. Specifically, strategy switching is related to when strategies happen to be applied, which will most directly affect the cost of using them in order to place globes or layers of paint (C). This seems reasonable in this experiment, where the tasks were rather easy for the participants. Consequently, individual differences in some cognitive capacity, like working memory, should not be expected to have a huge impact on performance. Rather, the model suggests that in this case variability in performance may simply be the result of making the right choice at the right time. Depending on when the options are chosen, they will result in different costs for placing the globes (layers of paint) correctly. This, in turn, will affect the likelihood that the various approaches will be attempted in the future.

5. Conclusion

The experiments provide clear evidence of a transition toward more planning and to planning that more frequently produced optimal solution paths. With no experience, participants took more moves to solve the problems, and more often used the less effective flat-to-flat strategy to reach the goal state. By the end of each experiment, the length of the exploratory paths had dropped substantially and a greater proportion of the solutions incorporated optimal final paths. Based on previous research (e.g. Svendsen, 1991), it seems that the increase in planning in this task can be characterized as learning about the benefits of planning more so than learning about the task.

The ACT-R model provides an explanation of the transitions in planning and strategy use throughout the experiment. It was constructed to make random moves or to plan using either the flat-to-flat or optimal strategies suggested by the data. As the model learns the relative costs and benefits of planning and of these strategies, it tends to shift from the random and flat-to-flat strategies to an optimal planning strategy. The exploratory paths get shorter because as the model learns, it is more likely to plan moves earlier in the problem, thereby starting down the final path sooner. The learning also allows the model to differentiate between the two local planning strategies, gradually coming to prefer overall efficiency over simplicity of execution. The ACT-R cost mechanism used in the model provides a straightforward explanation for these phenomena. The comparison is made on the basis of how much effort (time) is expected to be needed for each approach. With experience, these estimates come to accurately reflect their costs. The addition of noise to the calculation of E results in a model that produces a noisy shift from less effective to more effective plans, in much the same way as participants in the experiments. This type of shift in performance seems incompatible with theories that posit representational shifts that lead to better solutions. These explanations typically rely on sudden changes, or insight, that lead to abrupt and permanent improvements in performance (e.g. Anzai & Simon, 1979).

The model also predicts the differences found between the two types of five-move problems used in these experiments. If planning begins with a focus on the small globe (yellow layer of paint), one of these problem types is more difficult to solve, as it requires two extra moves in order to place the small globe (yellow layer of paint) in its goal location. This also explains why those who did the easier problems would be initially more dependent on the flat-to-flat strategy in the transfer problems. Since planning two moves was sufficient to find the optimal solution on the easy training problems, the participants who solved them were less likely to plan the longer sequences (four moves) in the transfer problems. In this case, they fell back on heuristic plans like the flat-to-flat strategy.

Not only does the model provide a good fit to the aggregate strategy use data, but also produces variability in performance that approaches that of the participants themselves. Specifically, the model produces variation in strategy use on the final path that nearly captures the range of participant behavior (Figs. 10 and 11). This is an encouraging sign, as one major criticism of computational models has been that they often fail to capture the variability of human behavior, tending only to fit average results (e.g. Roberts & Pashler, 2000).

The research and model presented here complement other research in problem solving and provide evidence pertinent to other interesting research areas. Firstly, the model's fit to the data is based upon a mechanism that has been used previously to explain shifts in behavior over time (e.g. Lovett, 1998; Matessa & Anderson, 2000). As familiarity is gained with a task, more accurate appraisals can be made of how well various approaches will fare. The result is problem solving behavior that becomes increasingly optimal over time. Secondly, this mechanism involves a noisy shift from simpler strategies to more sophisticated approaches, producing gradual improvement over time. Both the data and the model support this type of transition, adding to the literature that has addressed this general issue in several areas (e.g. Bower & Trabasso, 1963; Siegler, 1987). Finally, the model provides an explanation of how individual variability may arise in tasks that do not place a heavy burden on cognitive capacities. Differences in performance among individuals in this experiment appear to be the result of accidents of the participant's behavior. Making the 'correct' random move, or trying a strategy at an opportune moment increases the likelihood that the strategy will be used on a subsequent trial. Over the course of 12 or 18 problems, these random variations in specific moves or problems can accumulate to produce noticeable differences in overall performance.

In conclusion, the experiments and model presented here illustrate a noisy shift from strategies that are easy but relatively ineffective to ones that are more difficult to execute, but also quite effective in solving the problems. The model provides an explanation for this shift that involves comparing the strategies based on the estimated costs of using them to achieve local goals (subgoals). These findings add credence to the idea that such transitions are not produced by all-or-none transitions in strategy choice or planning behavior (Lamaire & Reder, 1999; Reder, 1982; Siegler, 1987). Rather, over a number of trials, better estimates of cost can be made based on the experienced costs. As this happens, greater sophistication gradually supplants random exploration. This idea accounts for both the decrease in exploratory path length and for the transition to more optimal final paths. This suggests that such evaluations of utility are generally applicable features of cognitive functioning.

Acknowledgements

This research was supported by NSF grant #BCS997-5-220. Portions of this research were presented at the Fourth International Conference on Cognitive Modeling in Fairfax, Virginia (July 2001). The authors would like to thank Raluca Budiu, Dick Hayes, and Marsha Lovett for their comments on earlier versions of this paper.

Appendix A. Descriptions and rules for Tower of Hanoi isomorphs used in this study

Monster move isomorph

Description

On a strange planet far from here, there is a race of intelligent, if slightly odd, monsters. They are six-handed and have curious customs including the use of globes in ceremonies. Because of the curious nature of their society, both monsters and globes come in exactly three sizes: small, medium, and large.

These monsters live in a complex culture, where different occasions or events require that each size monster hold a particular size globe. In addition, they have a detailed system of etiquette which determines how the globes may be exchanged. You will be asked to help three monsters prepare for events by telling them how to exchange globes in agreement with their traditions and culture.

Rules

Monster etiquette is complicated and very specific. According to the rules of this culture, globes may only be passed according to the following restrictions.

1. Only one globe may be passed at a time.
2. A smaller globe may not be passed to a monster holding a larger globe.

3. If more than one globe is being held by a single monster, that monster may only pass the largest globe it is holding.

Paint stripping isomorph

Description

A chemical has been developed that allows a person to strip a layer of paint from a painted surface and reuse the paint on something else. It is cheaper, faster and less environmentally damaging than traditional painting techniques. The only drawback is that the paint that has been stripped must be reapplied right away. A single coat of paint can be stripped and reapplied an unlimited number of times without sacrificing the quality or look of the paint.

In this problem you will be helping an individual repaint pieces of furniture. She has heard of this new product and is fascinated by the possibilities. Of course, there are some practical limitations in using this product.

Rules

Whenever using paint, there are some practical issues that must be considered. For this problem, they are as follows.

1. Only one layer of paint may be stripped at a time.
2. If there is more than one layer of paint on a surface, only the darker (visible) layer of paint may be stripped.
3. A lighter shade of paint may not be painted over a darker shade of paint.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

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

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

Bower, G., & Trabasso, T. (1963). Reversals prior to solution in concept identification. *Journal of Experimental Psychology*, *66*, 409–418.

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, Poster session presented at the Annual Meeting of the Cognitive Science Society, Philadelphia, PA.

Hayes, J. R., & Simon, H. A. (1974). Understanding written problem instructions. In Gregg, L. W. (Ed.), *Knowledge and cognition*. Hillsdale, NJ: Erlbaum, pp. 167–200.

Hayes, J. R., & Simon, H. A. (1977). Psychological differences among problem isomorphs. In Castellan, N. J., Pisoni, D. B., & Potts, G. R. (Eds.), *Cognitive theory*, vol. II. Hillsdale, NJ: Lawrence Erlbaum, pp. 21–41.

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

Lamaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition*, *27*, 364–382.

Lovett, M. C. (1998). Choice. In Anderson, J. R., & Lebiere, C. (Eds.), *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum, pp. 255–296.

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

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

Luchins, A. S. (1942). Mechanization in problem solving: the effect of einstellung. *Psychological Monographs*, *54*(6), 95.

Matessa, M., & Anderson, J. R. (2000). Modeling focused learning in role assignment. *Language and Cognitive Processes*, *15*, 263–292.

Reder, L. M. (1982). Plausibility judgments versus fact retrieval: alternative strategies for sentence verification. *Psychological Review*, *89*, 250–280.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*, 358–367.

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

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

Svendsen, G. B. (1991). The influence of interface style of problem solving. *International Journal of Man-Machine Studies*, *35*, 379–397.

VanLehn, K. (1991). Rule acquisition events in the discovery of problem-solving strategies. *Cognitive Science*, *15*, 1–47.