Declarative and Procedural Strategies in Problem Solving: Evidence from the Toads and Frogs Puzzle

Fabio Del Missier (delmisfa@univ.trieste.it)
Department of Psychology, via S. Anastasio 12
Trieste, I-34134 Italy

Danilo Fum (fum@univ.trieste.it)
Department of Psychology, via S. Anastasio 12
Trieste, I-34134 Italy

Abstract

The relationship between theoretically-grounded hints, strategy selection, and solution performance in the Toads and Frogs puzzle, a well-structured problem in which weak methods are difficult to apply, is investigated through an experiment and an ACT-R simulation. The main results show that providing a state specific hint is useful in speeding up the adoption of a hybrid solution strategy, comprising both the retrieval of previous moves and the proceduralization of a domain-specific heuristic that avoids any kind of forward search. The implications of the results for the problem solving theory are discussed.

Introduction

Research work on problem solving has attained to significant success in identifying the sources of difficulty for several kinds of well-structured problems (Newell & Simon, 1972). Working memory limitations (Miyake & Shah, 1999), in particular, play a prominent role in explaining why some problems are "so hard", and many factors have been identified that affect the working memory load. A partial list comprises the execution of legality tests on the operators (Kotovsky, Hayes & Simon, 1985), the number of options to be considered, and the availability of useful external memories (Cary & Carlson, 2001; Zhang & Norman, 1994).

Independently from working memory limitations, problem solvers seem reluctant to engage in a high degree of forward planning (Atwood, Masson & Polson, 1980; Ward & Allport, 1997; Simon & Reed, 1976). People usually recur to heuristic strategies, often relying on weak methods such as hill-climbing or means-ends analysis (Anderson, 1982; Anzai & Simon, 1979; Simon, 1975; Simon & Reed, 1976), and take into account only a limited number of states.

It is however interesting to wonder what strategy would be used in problems requiring a substantial degree of search when weak methods are not directly applicable. Here we try to answer this question by carrying out a first empirical exploration of problem solving behavior in a new kind of task.

First, we introduce the *Toads and Frogs* puzzle (henceforth T&F), a problem we found particularly suitable for the present research because of the peculiar structure of its problem space. We describe then a set of candidate strategies for solving it, and present an experiment designed to study the effectiveness of two types of hints delivered through the interface. Next we summarize the results of an ACT-R simulation aimed at identifying the strategies actually used by participants, and at tracing their development. Finally, we discuss the findings in the light of two main classes of problem solving strategies (memory-based vs. rule-based).

The Toads and Frogs Puzzle

The T&F puzzle (Berlekamp, Conway & Guy, 2001) is a well-structured problem that, to the best of our knowledge, has never been previously utilized in psychological research.

In the variant used here, three toads are placed on the three leftmost squares of a seven-square board while three frogs are placed on the three rightmost squares (Figure 1). The central square is initially empty. The goal of the game is to switch the animals' positions by having the toads occupy the three rightmost, and the frogs the three leftmost squares, respectively. A square can be occupied by only an animal at a time, and an animal can move only into the empty square. Toads can move only rightward and frogs only leftward. There are two possible types of move: a Slide to the next (empty) square and a Jump over an animal of a different type to a two-square apart empty position. A solution can be reached in exactly 15 moves, 9 Jumps and 6 Slides. Two symmetrical solution paths are possible, depending on the animal that is moved first (the solution sequence for the frog-move-first type of problem is presented in Figure 1).

Some of the moves in the solution path are forced (Jump only or Slide only) while in the remaining cases it is necessary to choose between two Slides, or between a Jump and a Slide, or between two Jumps (off the solution path only). Excluding the first move that allows two possible options, there exists a single right move for every solution step, and it is not possible to retract a move when it has been done.

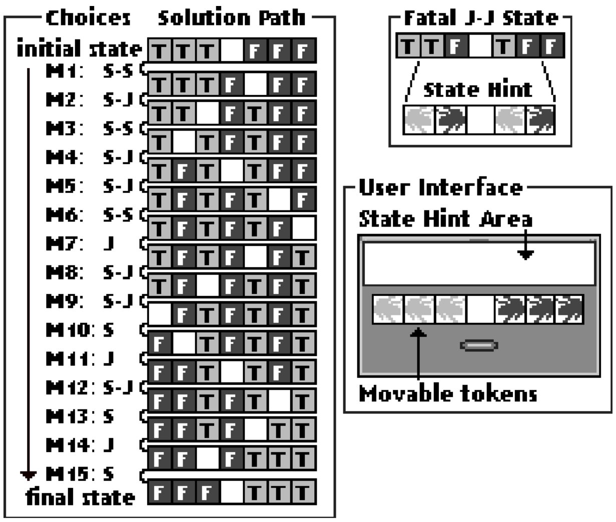

Figure 1: The Toads and Frogs problem

From a normative point of view, it is possible to identify two rules that foster the T&F puzzle solution. First, in deciding between two Slides, avoid the move that brings to a fatal Jump-Jump configuration, i.e., a state in which both legal moves are Jumps (an example is given in Figure 1). Every such configuration lies off the solution path and leads eventually to a dead end. Second, in choosing between a Jump and a Slide, select the Jump.

From a psychological viewpoint, the problem presents some interesting properties. It cannot be solved through hill-climbing, because this strategy does not help in deciding what to do when facing the critical Slide-Slide choices. A pure forward planning approach will not work either, because detecting dead ends would require an unattainable degree of look-ahead. The problem is also hard to solve by means-ends analysis, because of the difficulty to find out useful subgoals. The natural subgoal choice (i.e. putting the most advanced animal to its target location) cannot be used in the first solution steps due to the necessity to plan up to seven moves and to detect ten potential dead-ends. Finally, it is unlikely to find the solution by pure chance (p=0.0039).

Candidate Problem Solving Strategies

From the analysis of some concurrent verbal protocols collected in a pilot study and by combing the problem solving literature, we were able to define a set of candidate solution strategies for the T&F puzzle. We will briefly present the strategies and the hypotheses about the performance measure that can be derived from them.

AVOID BLOCKS (AB) This strategy can be reduced to two rules: (1) Avoid moves that bring to a "blocked" state. A state is considered blocked whenever the moved animal is preceded in its moving direction by an animal of the same type that (a) cannot be moved, and (b) is not placed in its final position. (2) In cases when none of the two legal moves brings to a blocked state, choose randomly between them. This strategy reflects the general heuristic of avoiding bad states, and it hypothesizes chance-level performance in the Slide-Slide choices.

JUMP AND RANDOM (JRND) In the Jump-Slide choices always select a Jump move. In the Slide-Slide choices choose randomly. The "always Jump" rule could be acquired by proceduralizing the hill-climbing heuristic. The strategy predicts a higher error rate in the Slide-Slide decisions in comparison with the Slide-Jump ones.

MOVE PATTERN (MP) Because the sequence of moves to reach the solution is highly patterned, implicitly learn to perform the pattern of moves. For instance, when the first move is the Slide of the most advanced frog [F1], the solution pattern is: [F1] T1 T2 F1 F2 F3 T1 T2 T3 F1 F2 F3 [T2 T3 F3] with the three final moves being on-the-target moves. This strategy (described, among others, by Reber & Kotovsky, 1997) predicts an approximately similar percentage of errors for each class of choices and no difference in the associated decision times.

JUMP AND RETRIEVE (JR) In the Jump-Slide choices, always select a Jump move. In the Slide-Slide choices, try to retrieve the last decision taken in the same situation, using the current state as a memory cue. If the last trial is remembered as a success, repeat the retrieved move, else select the alternative legal move. The strategy constitutes a partial version of the trial-and-error weak method. It predicts higher times (due to retrieval costs) and higher error rates (due to retrieval errors) in the Slide-Slide choices than in the Slide-Jump ones.

JUMP AND SPACE (JS) Always Jump in the Jump-Slide pick. In the Slide-Slide choices, select the move that brings to a state in which there is exactly one interposed square between each neighboring pair of animals like the moved one. This rule implements a domain-specific preference for states in which animals of the same type are regularly spaced. The strategy stems from the verbal protocols of participants stating their desire to reach an "alternating sequence" of animals. They claimed they wanted "to make some space" between the animals of the same type to allow the possibility for the other animals to "jump into". The strategy could be acquired by a perceptual noticing mechanism (Ruiz & Newell, 1989) and by the use of the perceived features as subgoals within a means-ends approach (Anzai & Simon, 1979). It requires: (a) to imagine a state stemming from the execution of a move; (b) to maintain the imagined state in working memory; (c) to perform two distance tests on the imagined state; (d) to select the right move depending on the test outcome. This process is time expensive and can be error-prone, therefore the strategy predicts higher times and error rates in the Slide-Slide decisions than in the Slide-Jump ones.

The Experiment

We performed an experiment to analyze the effect of two interface hints on the performance in the T&F puzzle, and on the acquisition of the solution strategies.

The first hint is motivated by the work of Kotovsky, Hayes & Simon (1985), who suggested that the execution of legality tests on the operators constitutes a major source of difficulty in the Tower of Hanoi isomorphs (the so-called rule-application hypothesis). According to the authors, the working memory load associated with these tests initially hinders the discovery of a solution strategy. Assuming that legality tests could be a source of difficulty also in the T&F problem, we devised an interface hint that completely removes any cost associated with their execution. With the “legality hint” enabled, the movable tokens pop out in the interface, being displayed on squares of a different color. In this way the legality tests are embedded into the interface. This manipulation should free working memory resources, and make them available for problem solving and for the acquisition of a solution strategy.

The second hint is related to the structure of the problem space and, in particular, to our hypothesis that the Slide-Slide choices are the biggest sources of difficulty in the T&F puzzle because they cannot be handled by the weak methods commonly used in problem solving. With the “state hint” enabled, an image pattern representing the common part of the fatal Jump-Jump states is presented in the State Hint Area of the User Interface Window (Figure 1) whenever participants face a Slide-Slide choice. The participants were instructed to avoid executing a move that will bring them in a state corresponding to the hint pattern. According to our hypothesis, the hint should be very effective in removing the main error source, and in helping participants to find a good decision policy in the Slide-Slide choices.

Finally, it is reasonable to expect a synergic interaction between the two kinds of hints, since the working memory unload provided by the first hint should enhance the effect of the state hint.

Method

Participants and Materials The participants were 72 undergraduates students, aged between 19 and 30. The sample was approximately balanced for gender. All the participants had a basic familiarity with computers, and were able to use the mouse. Two different versions of the T&F problem were used by having the first move (Slide a toad vs. Slide a frog) automatically generated by the computer. The first choice actually splits the problem space in two almost completely non-overlapping sub-spaces (there are only a few common states, placed off the solution paths).

Procedure Participants read an instruction document that explained the rules of the T&F, described the task, and showed how to use the interface. Depending on the experimental conditions, the document presented also the available hints. In order to decrease the likelihood of a random solution, we adopted as the learning criterion the attainment of the final state in two consecutive trials. The interface did not allow undoing a previously executed move. After getting stuck, or after a voluntary interruption of a trial, the solver should start again from the beginning. Participants were required to solve both versions of the problem in the order specified by the experimental design. They had allotted a maximum time of 20 min for the first problem and of 10 min for the second one. No limits were placed on the number of trials.

Apparatus A PowerMacintosh G3 was used for the experiment. A program implementing the T&F puzzle was written using MCL 4.3.1 and CLIM2. The program recorded each participant move and the associated time. The interface window was composed by two parts. The upper part was only used to display the state hint. The lower part showed the puzzle board and an “Interrupt” button. To move an animal, it was only required to click on the square containing it. If the move was allowed by the problem rules, the positions of the animal and of the blank square were immediately updated; in case of an illegal move, a warning sound was delivered.

Experimental Design Two between-subjects independent variables (State Hint and Legal Hint availability) were manipulated in a 2x2 factorial design. The 72 participants were randomly assigned to the four experimental groups. We adopted a transfer design, presenting the two different versions of the problem in a counterbalanced order. The hints were available only for the first problem, and the participants were made fully aware of this by the instructions and by the experimenter. The basic dependent variables for each problem were the achievement of the criterion, the total time, and the total number of trials needed to achieve it. More detailed dependent measures were the percentages of errors, and the mean move latency for each choice class.

Results

We will first present the results on the whole data to test the hypotheses about the effectiveness of the interface hints. Then we will report two series of results concerning the participants who reached the criterion in both the problems: the first to assess the transfer, the second to provide detailed performance analyses that will be compared with the predictions of the various strategies.

Hint Effectiveness

Criterion. Table 1 presents the frequency of problem solving outcomes for each hint group and problem.

Table 1: Frequency of problem solving outcomes.

GROUP	Total Failure	Criterion Problem 1 Only	Criterion Problem 2 Only	Criterion Problem 1 & 2
Control	5	0	6	7
Legal Hint	1	2	3	12
State Hint	1	2	0	15
Both Hints	3	4	2	9

In the first problem, a higher number of participants in each hint group reached the criterion in comparison with the control group (no hint). The Fisher Exact test showed significant differences between the control group and the Legal Hint group (p=0.0205), the State Hint group (p=0.0005) and the group with both hints (p=0.0461), respectively. No significant differences were found in the second problem. Contrasting the frequency of criterion attainment in both problems with the frequency of any other outcome, only the State Hint group resulted significantly better than the control group (Fisher Exact test, p=0.0076).

Times. A 2x2 ANOVA on the aggregate problem times yielded only the significant main effect of the State Hint (F(1,68)=8.37, MSE=50.98, p<0.01). The participants without the state hint had to devote more time to the problems (State Hint: M=15.08 min, No Hint: M=19.95 min).

Number of trials. A 2x2 ANOVA on the aggregate problem trials showed a significant interaction between State Hint and Legal Hint availability (F(1,68)=4.78, MSE=216.65, p<0.05). Only the main effect of the State Hint was significant (F(1,68)=10.00, MSE=216.65, p<0.01). The Tukey HSD test showed significant differences between the State Hint group and the control group (p<0.01) and between the State Hint group and the Legal Hint group (p<0.05). The State Hint group had the lowest mean (M=15.89), followed by the control group (M=27), the Legal Hint group (M=30.39), and the group with both hints (M=34.44).

Transfer

Times. We computed a 2x2 ANOVA (Problem x State Hint availability) on the problem solution times. The analysis showed a two-way interaction between Problem and State Hint (F(1,41)=9.20, MSE=9.04, p<0.01). The Unequal N Tukey HSD test underlined a significant difference between the participants with the state hint and those without it, but only in the first problem (p<0.05). In both the conditions, the time to criterion significantly decreased from the first to the second problem (State Hint: n=24, $M1$=7.26 min, $M2$=4.56 min, p<0.05; No Hint: n=19, $M1$=10.85 min, $M2$= 4.19 min, p<0.001).

Number of trials. A 2x2 ANOVA (Problem x State Hint availability) showed a two-way interaction between Problem and State Hint (F(1,41)=9.96, MSE=25.86, p<0.01). The Unequal N Tukey HSD test highlighted the significant difference between the participants with the state hint and those without it, again only in the first problem (p<0.01). Only in the condition without the state hint, the number of trials significantly decreased from the first to the second problem (State Hint: n=24, $M1$=8.42, $M2$=8.12; No Hint: n=19, $M1$=16.63, $M2$=9.37, p<0.001).

Detailed Performance The choices of each participant in the non-forced moves were categorized depending on their location: the first Slide-Slide choice (SS-1), the second Slide-Slide choice (SS-2), or a Jump-Slide decision point. Then, the percentage of errors for each choice point was computed, according to the participant's number of transitions for that state. Finally, given the low value of the percentages for each of the six Jump-Slide choices, an overall mean error percentage (JS-M) for each participant was computed.

Errors. A 2x2x2 ANOVA (State Hint availability x Problem x Error type) yielded a significant three-ways interaction (F(2,82)=3.30, MSE=251.3, p<0.05). There was no significant difference on the JS-M error percentage between the State Hint and No State Hint conditions and between the first and the second problem (Unequal N Tukey HSD test). The JS-M error percentages were significantly lower than the SS error percentages in each State Hint condition and problem (least significant difference: p<0.05; JS-M means between 2 and 0). The difference on the SS-1 error percentages between the State Hint and No State Hint conditions was significant only in the first problem (M-h=18, M-nh=43, p<0.01). A similar result was obtained for the SS-2 error percentages, but the difference only approached significance (M-h=26, M-nh=47, p=0.059). Some single sample t-tests, carried out to evaluate the null hypothesis that the error percentages of the SS choices were not different from random performance (50%), did not allow us to reject the hypothesis only in the condition without state hint of the first problem. In the other cases the percentages were significantly lower (least significant difference: p<0.05; means ranging from 18 to 35).

Move times. A 2x2x2 ANOVA (State Hint availability x Problem x Error type) on the mean move times, showed a significant three-ways interaction (F(3,123)=7.28, MSE=17.57, p<0.001). The analysis comprised also the two kinds of forced moves (Jump and Slide only) and used the aggregated SS decision times. There were significant differences between the first and the second problem, but only for the real choices. The Unequal N HSD test yielded the following results in the conditions with the state hint: p<0.001 for SS ($M1$=10.40 s, $M2$=4.94 s) and p<0.001 for SJ-M ($M1$=5.68 s, $M2$=3.30 s). In the conditions without state hint the results are analogous: p<0.001 for SS ($M1$=7.06 s, $M2$=4.16 s) and p<0.001 for SJ-M ($M1$=5.23 s, $M2$=2.93 s). For both problems, the SS times were significantly higher than the JS times (State

Hint: $p<0.001$ and $p<0.001$; No State Hint: $p<0.001$ and $p<0.05$). The JS times were significantly slower that the forced moves ($p<0.001$ for each condition and problem). The only significant difference involving the State Hint condition was that on the SS decisions in the first problem (*M-h*=10.40 s, *M-nh*=7.06 s, $p<0.05$).

Discussion

The experimental results provided strong support for the effectiveness of the state hint. This hint promoted the achievement of the criterion in the first problem, while its removal did not produce any performance decrease in the second one. A significant support for the effectiveness of the legal hint was not reached, but the limited power of the tests advises caution in the interpretation.

The results on the percentages of errors clearly showed that the main sources of difficulty for the problem were the Slide-Slide choices. The state hint worked by reducing the errors in these decision points in the first problem, thus allowing a faster development of a decision policy.

However, about 60% of the participants were able to reach the criterion in both the problems, thus demonstrating the possibility to acquire an advantageous solution strategy with a sufficient amount of practice. Furthermore, the performance of the successful solvers in the second problem was quite similar across the experimental conditions.

The error and time results were not compatible with the adoption of the MP strategy. The use of the AB or JRND strategies could not excluded in the conditions without the state hint, but only in the first problem. The JR and the JS strategies were the only two strategies potentially in accordance with the evidence on the second problem.

The Simulation

We decided to undertake an ACT-R simulation (Anderson & Lebiere, 1998) to formulate more detailed predictions from the partially supported strategies: Avoid Blocks (AB), Jump and Random (JRND), Jump and Retrieve (JR), and Jump and Space (JS), and to test them against the appropriate subset of data. We implemented the four strategies as separate ACT-R models.

While the JRND strategy leads to a direct implementation, the AB and JS require to mentally simulate the execution of the possible legal moves, and to evaluate the states deriving from them. We implemented the construction and storage of the imagined states via time-costly productions, and we assumed that the state evaluations were always performed errorless. Given that verbal protocol data indicated that detecting a block situation is quite an easy task, we assumed also that the AB model did not need to retrieve the simulated move. Conversely, the JS model, being engaged in more difficult distance tests, had to use the error prone memory retrieval. Thus, the JS model often dictated a move that did not comply with the strategy requirements.

The implementation of the JR strategy required, on the other hand, the memory storage of each Slide-Slide choice, and of the outcome of each trial. When facing a Slide-Slide problem state, the model tries to retrieve the last decision taken in the same situation using the current board configuration as a retrieval cue. It also tries to retrieve the outcome of the last performed trial. If the trial is remembered as a success, the retrieved move will be executed, otherwise the alternative legal move will be carried out. Thus, the JR model sometimes makes the wrong selection due to retrieval errors or to a temporal mismatch between the retrieved move and the trial.

Procedure and Results

We compared the AB and JRND predictions with the data of the first problem without the state hint. Then, we contrasted the JR and JS results with the data of the second problem, in the conditions without the state hint.

For each strategy, we executed a number of simulation runs (2000) sufficient to provide an efficient estimation. The dependent variables were the number of trials to reach the criterion, and the error percentages on the two Slide-Slide choices. The JS-M percentages were not taken into account, because all the strategies predicted few errors in the JS choices and the empirical values are always very close to zero. The experimental data and the simulation results are presented in Table 2 and Table 3 (the number near the strategy label stands for the problem being simulated).

Table 2: Experimental data.

PROBLEM	Trials to Criterion	SS-1 Error	SS-2 Error
FIRST	16.63	43	47
SECOND	9.37	30	35

Table 3: Simulation predictions. (Values within the 95% confidence interval of data are marked with *)

STRATEGY	Trials to Criterion	SS-1 Error	SS-2 Error
AB (1)	13.95*	43*	40*
JRND (1)	13.08	42*	39*
JR (2)	7.56*	34*	37*
JS (2)	13.70	42	39*

The results showed that the AB strategy obtained a slightly better fit than the JRND on the first problem results. The best fit for the performance on the second problem was obtained by the JR strategy. Thus, the most probable explanation is that participants shifted from the AB to the JR strategy as a consequence of their increased experience with the task.

It seems reasonable to assume that the state hint was able to foster the adoption of the JR strategy. A simple memorization of the state hint seems quite implausible, given the move latency data. In the second problem, if the participants were retrieving the state hint and using it to carry out the SS move selection, we should have observed a significant increase in the mean latency. On the contrary, the mean times for the SS moves resulted much lower in the second problem. Furthermore, the whole second problem performance was very similar for the groups with and without the state hint. So, it could be parsimoniously hypothesized that the indirect suggestion of the correct move made available through the state hint could have simply speeded up the natural development of a more general memory-based strategy.

Conclusions

The evidence provided can help us to answer the question that motivated our research on the T&F puzzle. People were able to acquire an effective solution strategy even when hill-climbing or means-ends analysis were not directly applicable. We gained support for a state-avoidance strategy in the initial problem solving attempts, and for a memory-based strategy in later trials. Finally, we demonstrated the effectiveness of a specific type of interface hint.

From a broader prospective, it is worth noting that some findings, obtained in very different settings (Howes & Payne, 2001), seem to bring converging evidence for a kind of memory-based problem solving when weak methods are not applicable or not efficient.

Another meaningful point is that our hybrid memory-based JR strategy was probably derived partly from the weak method of trial-and-error, and partly from the proceduralization of the state-avoidance heuristic. This raises an interesting theoretical issue concerning the ontology of multi-step problem solving strategies. The strategies commonly proposed in the literature seem to belong either to the algorithmic or to the memory-retrieval class. Our work seems to suggest that, in some multi-step situations, people are able to acquire hybrid strategies, relying on memory retrieval to handle problem solving steps where procedural methods don't work. So, in our view, it seems necessary to make a distinction between the strategies that require the intentional usage of memorized instances (Logan, 1988), the ones that keep track of the previous history in a procedural form (Lovett & Anderson, 1996), and the hybrid formulations. This also means that it will be generally necessary to make explicit all the potential candidate solution strategies, and to contrast them in their capacity to fit the data.

References

Anderson, J.R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 396-406.

Anderson, J.R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

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

Atwood, M.E., Masson, M.E. & Polson, P.G. (1980). Further exploration with a process model for water jug problems. *Memory & Cognition*, *8*, 182-192.

Berlekamp, E.R., Conway J.H. & Guy, R.K. (2001). *Winning Ways for Your Mathematical Plays*, v1, A K Peters.

Cary, M. & Carlson, R.A. (2001). Distributing working memory resources during problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 836-848.

Kotovsky, K., Hayes, J.R. & Simon, H.A. (1985). Why are some problems hard?: Evidence from the Tower of Hanoi. *Cognitive Psychology*, 17, 248-294.

Howes, A. & Payne, S.J. (2001). The strategic use of memory for frequency and recency in search control. *Proceedings of the Twenty Third Annual Conference of the Cognitive Science Society* (pp. 425-440). Hillsdale, NJ: Erlbaum.

Logan, G. (1988). Towards an instance theory of automatization. *Psychological Review*, *95*, 492–527.

Lovett, M.C. & Anderson, J.R. (1996). History of success and current context in problem solving. *Cognitive Psychology*, *31*, 168-217.

Miyake, A. & Shah, P. (Eds.) (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge University Press.

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

Reber, P.J & Kotovsky, K. (1997) Implicit learning in problem solving: The role of working memory capacity. *Journal of Experimental Psychology: General, 126, 178-203.*

Ruiz, D. & Newell, A. (1989). Tower-noticing triggers strategy-change in the Tower of Hanoi: a Soar model. *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society* (pp. 522-529). Hillsdale, NJ: Erlbaum.

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

Simon, H. A. & Reed, S.K. (1976). Modelling strategy shifts on a problem solving task. *Cognitive Psychology*, *8*, 86-97.

Zhang, J. & Norman, D.A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*, 87-122.

Ward, G. & Allport, A. (1997). Planning and problem solving using the 5-disk Tower of London task. *Quarterly Journal of Experimental Psychology*, *50*, 49-78.