# UC Merced

## Proceedings of the Annual Meeting of the Cognitive Science Society

**Title**

Making Heads or Tails out of Selecting Problem-Solving Strategies

**Permalink**

https://escholarship.org/uc/item/0x63m8xs

**Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 17(0)

**Authors**

Lovett, Marsha C.
Anderson, John R.

**Publication Date**

1995

Peer reviewed

# Making Heads or Tails out of Selecting Problem-Solving Strategies

**Marsha C. Lovett**
Psychology Department
Carnegie Mellon University
Pittsburgh, PA 15232
lovett+@cmu.edu

**John R. Anderson**
Psychology Department
Carnegie Mellon University
Pittsburgh, PA 15232
ja+@cmu.edu

## Abstract

When solvers have more than one strategy available for a given problem, they must make a selection. As they select and use different strategies, solvers can learn the strengths and weaknesses of each. We study how solvers learn about the relative success rates of two strategies in the Building Sticks Task and what influence this learning has on later strategy selections. A theory of how people learn from and make such selections in an adaptive way is part of the ACT-R architecture (Anderson, 1993). We develop a computational model within ACT-R that predicts individual subjects' selections based on their histories of success and failure. The model fits the selection behavior of two subgroups of subjects: those who select each strategy according to its probability of success and those who select the more successful strategy exclusively. We relate these results to probability matching, a robust finding in the probability-learning literature that occurs when people select a response (e.g., guess heads vs. tails) a proportion of the time equal to the probability that the corresponding event occurs (e.g., the coin comes up heads vs. tails).

## Introduction

People often have multiple strategies available for approaching a given problem, but they must select just one strategy to apply. Much research on this selection process has focused on the influence of problem features—how the "looks" of a problem can influence what strategy solvers choose to apply to it (e.g., Atwood et al., 1980; Reder & Ritter, 1992; Siegler & Shipley, 1995). Another influence on strategy selection, however, is the knowledge solvers have learned about the strategies themselves (e.g., how successful different strategies are or how costly they are to apply). It is clear that, all else being equal, solvers should tend to select a strategy that is more successful, less costly, etc. than the others, but little is known about how solvers actually represent and use such strategy information. In this paper, we describe a detailed, quantitative study of how solvers' histories of success with different strategies impact selection.

While this focus has not received much attention in the problem-solving literature, the influence of success rates on an analogous selection task has. This is the two-choice selection task studied in probability-learning experiments (e.g., Estes, 1964; Humphreys, 1939; Jones, 1971). The basic paradigm of these experiments is to ask subjects to predict which of two outcomes will occur (e.g., a coin coming up heads or tails), where one outcome (say, heads) has probability $p$ and the other (tails)

has probability $1-p$. In these experiments, the success rate $p$ is varied within or between subjects, and selection tendencies (e.g., guessing heads vs. tails) are studied as a function of $p$. A common finding is that, with multiple trials, subjects exhibit *probability matching*—they select each response a proportion of the time equal to its probability of occurring (Estes, 1964). For example, if heads comes up with probability 0.8, subjects will tend to guess heads 0.8 of the time. (Notice that this does not maximize one's expected number of correct guesses.)

Many models have been developed that predict probability matching in such contextually sparse situations (Atkinson & Estes, 1963; Gluck & Bower, 1988; Lordahl, 1970). An interesting question, however, is whether people will exhibit probability matching when the selection task is embedded in the larger context of solving a problem. One might expect that the learning mechanism underlying probability matching is fundamental and applies in a variety of contexts, but it is possible that, when making selections in service of a larger goal, solvers will be more likely to maximize their expected number of solutions by selecting the more successful strategy all of the time.

To gain a better understanding of the role of strategy-success information in strategy selection and to test the generality of probability matching, we studied how solvers select between two problem-solving strategies in a novel task, the Building Sticks Task (BST) (Lovett, 1994). This provided us with two important opportunities. First, it allowed us to record every success and failure a solver experienced while using the two strategies. We used these trial-by-trial, individual histories to compare an ACT-R model of strategy selection with the class of models that predicts asymptotic probability-matching behavior. Second, it allowed us to manipulate the strategies' success rates from the start of subjects' experience. Thus, we could study the effects of strategy information as it is acquired. Note that, for all the problems we asked subjects to solve, the two strategies appeared equally appropriate and equally likely to lead to a solution; this focused our study on how subjects' history of success with the various strategies influenced their selections, and it made the analogy to probability-learning experiments clearer.[1] A major goal

---

[1] See Lovett & Anderson (1995) for a discussion of how history of success and apparent appropriateness of various choices jointly influence selections in problem solving.

of this research was to capture, in a computational model, how *people* learn from and make strategy selections.[2]

## A model of how people select between strategies in the BST

A theory of how people do this in an adaptive way is part of the ACT-R architecture (Anderson, 1993). We developed a computational model within ACT-R that predicts individual subjects' strategy selections in the BST. In the BST, the goal is to build a *current* stick that is equal in length to the *desired* stick by adding and subtracting multiple *building* sticks (see Figure 1). This goal can be achieved either by selecting a building stick that is smaller than the desired stick and subsequently building up to the desired stick's length (the UNDERSHOOT strategy) or by selecting a building stick that is longer than the desired stick and subsequently cutting down to the desired stick's length (the OVERSHOOT strategy). In the model, each strategy corresponds to a single production whose actions are executed only when its conditions are met in the current situation. Since both strategies can be executed in the initial problem state, the model must select between them on the first step of each problem. Solvers make the same selection implicitly as they solve each problem, and these selections are easily identified (to the experimenter) by the first step.
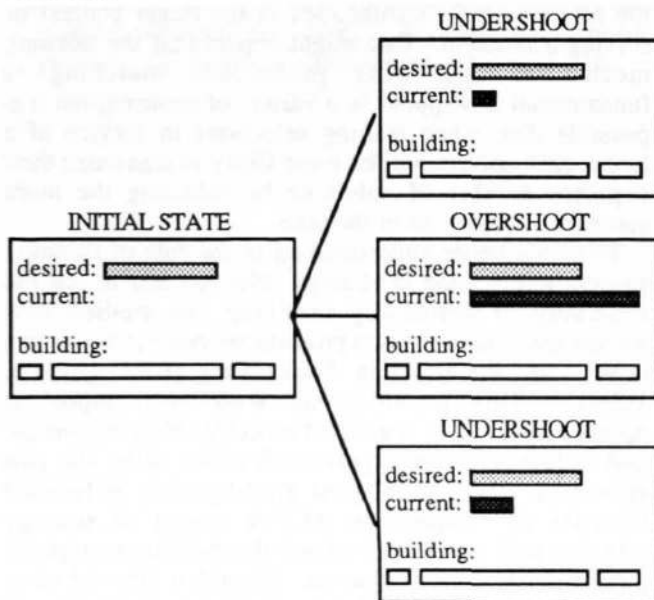


Figure 1. Initial and successor states in the BST.

To select between the two strategies, the model attempts to choose the step that will lead to the highest probability of success (where success is defined as achieving the goal). Since the actual probability of success resulting from a particular step cannot be known in advance, the model estimates the predicted probability

of success (PPS) for each move as a function of the history of success of the production involved in the move. The more often a strategy (i.e., production) has led to success in the past, the higher the model's PPS for moves using that strategy.[3] The PPS of strategy $j$ is estimated as a Bayesian posterior probability of success, $PPS_j = \frac{\alpha_j + s_j}{\alpha_j + \beta_j + s_j + f_j}$, where $s_j$ is the number of successes and $f_j$ the number of failures experienced with that strategy and $\frac{\alpha_j}{\alpha_j + \beta_j}$ is the model's prior for the success of that strategy. Note that the more one uses a particular strategy, the less its PPS depends on the prior and the more its PPS depends on the number of successes and failures experienced with the strategy. We should emphasize that, while this updating formula for PPS is meant to model a part of the process by which people make strategy selections, we do not propose that people are actually calculating these Bayesian updates *per se*. Rather, this formula is meant to capture the changes in knowledge that people learn through experience—changes that are presumably "computed" at the neural level.

When the model is selecting among multiple moves, it tends to select the move with highest PPS. To include a stochastic component in this selection process, Gaussian noise is added to each PPS value before the selection is made. Thus, the model can be thought of as selecting each move with a particular probability that is a function of its PPS value, relative to the other moves, and the amount of noise in the system. Note that our ACT-R model can exhibit probability-matching behavior, but it can also model other selection tendencies as long as the more successful strategy is selected more often than its competitors.

## Method

### Subjects

Subjects in this experiment were 68 Carnegie Mellon University undergraduates; of these, 49 received course credit for participating, and 19 received $5.00.

### Design

There were eight experimental conditions that differed according to three factors: (1) which strategy (UNDERSHOOT or OVERSHOOT) was designed to be more successful, (2) the relative success rate of the more successful strategy to the less successful strategy (high or low), and (3) whether or not the two strategies were complementary (i.e., whether failure of one implied success of the other). Analyses revealed that the assignment of UNDERSHOOT vs. OVERSHOOT to the more successful strategy did not affect the results, so we collapse this factor and label the four remaining conditions comp-hi, comp-lo, noncomp-hi, and noncomp-lo.

---

[2]Our modeling goals differ from those of related machine-learning work (e.g., Sutton, 1988).

[3]In ACT-R, the estimate of PPS can be influenced by other factors as well, but here it is relatively well modeled as a monotonic function of success rate alone.

Table 1. Solution probabilities as manipulated across conditions.

| Condition | P(MS) | P(LS) | P(N) | Condition | P(MS) | P(LS) | P(N) |
|---|---|---|---|---|---|---|---|
| Comp-hi | | | | Noncomp-hi | | | |
|   MS selected first | .80 | .20 | .00 |   MS selected first * | .80 | .04 | .16 |
|   LS selected first | .80 | .20 | .00 |   LS selected first ** | .64 | .20 | .16 |
| Comp-lo | | | | Noncomp-lo | | | |
|   MS selected first | .60 | .40 | .00 |   MS selected first * | .60 | .16 | .24 |
|   LS selected first | .60 | .40 | .00 |   LS selected first ** | .36 | .40 | .24 |

Note: MS = more successful strategy; LS = less successful strategy; P(MS) = probability MS solves; P(LS) = probability LS solves; P(N) = probability unsolvable. Text refers to * and **.

Subjects were assigned to one of these four "environments" that determined how likely each strategy was to solve problems. The ratios of success were chosen as 80/20 (high) and 60/40 (low). In the complementary conditions, these ratios represented the observed solution rates for the two strategies (e.g., in comp-hi, the more successful strategy solved 80% of the problems and the less successful strategy 20%). In these conditions, whenever the strategy chosen first on a given problem did not lead to a solution, the other strategy would. In contrast, in the noncomplementary conditions, some problems were not solvable. Each strategy had its designated probability of solving the problem (e.g., 80% and 20% for the more and less successful strategies in noncomp-hi). But, if the strategy selected first did not solve the problem (according to that probability), the other strategy only had a chance of solving (according to its designated probability). So, the probability that a problem would be solved by a particular strategy depended on which strategy was selected first.

Table 1 presents the probabilities of success for each strategy in terms of which strategy was selected first. Note that these values are different for the complementary and noncomplementary conditions. In particular, for the latter, the ratios of the two strategies' solution rates are not fixed at 80/20 and 60/40: When solvers select the more successful strategy first, the ratios become more extreme (80/4 and 60/16; see * in Table 1), and when solvers select the less successful strategy first, the ratios become less extreme (64/20 and 36/40; see ** in Table 1).

### Apparatus

Subjects worked individually on Macintosh IIci computers. A cT program (Physics Academic Software, 1992) ran the BST interface, provided initial instructions to subjects, and collected data. Each rectangle in Figure 1 is a sketch of the interface subjects saw.

### Procedure

At the beginning of the experiment, a computer tutorial provided subjects with instructions and practice on how to use the mouse to build sticks. Then, it automatically solved two sample BST problems (one by UNDERSHOOT and the other by OVERSHOOT). The experimental trials included 90 BST problems. In the complementary conditions, subjects were required to work on each problem until they solved it or had taken at least 20 steps. In the non-complementary conditions (because some problems were unsolvable), subjects had the additional option of clicking on a "next problem" button that activated after they took six steps. After the experimental trials, subjects were asked if the experiment had reminded them of any experiments they had learned about in class. One subject was reminded of Luchins's (1942) water jars experiment and the Einstellung effect and so was removed from the analysis.

### Stimuli

All problems were designed so that the two strategies would appear equally appropriate and equally likely to lead to a solution. This neutrality was measured according to a hill-climbing metric tested by Lovett & Anderson (1995). In addition, all problems had three, nearly identical versions: one solvable by UNDERSHOOT, one solvable by OVERSHOOT, and one unsolvable. The three versions of a given problem had the same desired stick but slightly different building sticks (they varied by one or two pixels on the screen). Thus, the three versions allowed a single problem to be switched from being solved by one strategy to the other or neither, merely by adjusting the building sticks sizes ever so slightly. Performing these adjustments with specified probabilities enabled us to manipulate the success rates of the strategies according to the values in Table 1. Adjustments only occurred in how the building sticks added to or subtracted from the solver's current stick. No subject noticed these adjustments or was suspicious about how the interface worked.

## Results and Discussion

In general, we present our results in terms of the percentage of problems on which subjects selected the more successful strategy first, with "more successful" defined by their condition. Figure 2 presents these averages for each block of 15 problems for each condition. We can use these data to test whether subjects' asymptotic behavior approximates probability matching. The two thin, horizontal lines in the figure represent probability-matching behavior for the comp-hi and comp-lo conditions at 80% and 60%. Although the comp-hi selections are somewhat above 80%, observed percentages for both of these conditions in the last three blocks are within 95% confidence intervals of the matching values.
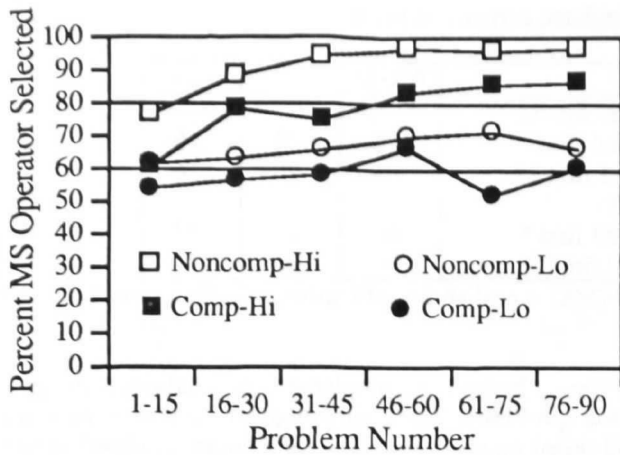
Figure 2. Mean percentage of problems (per 15) on which the more successful strategy was selected.

For the noncomplementary conditions, 80% and 60% are not necessarily "matching" percentages because here the proportion of problems solved by each strategy varies with the strategies subjects select. Suppose $a$ represents the proportion of problems on which a noncomp-hi subject selects the more successful strategy first. Referring to Table 1, this subject should experience success of the more successful strategy on $.80a + .64(1-a)$ of the problems, on average, and success of the less successful strategy on $.04a + .20(1-a)$ of the problems, on average. A further complication in defining probability matching here is that these two expressions do not add up to 1 since some problems are unsolvable. In other studies of choice, where not all trials lead to success, matching has been defined by the *matching law* (Herrnstein, 1961), which claims that subjects match their *ratio* of responses to the ratio of experienced reinforcements (reinforcements = solutions in our case):[4]

$$\frac{\#A\_Responses}{\#B\_Responses} = \frac{\#A\_Solutions}{\#B\_Solutions}. \qquad (1)$$

Setting the left-hand side equal to $a/(1-a)$ and the right-hand side equal to $(.80a + .64(1-a))/(.04a + .20(1-a))$, we can obtain an equilibrium matching value by solving for $a$. For noncomp-hi, $a \sim .94$, and, for noncomp-lo, $a \sim .69$. Thus, the almost exclusive selection of the more successful strategy among noncomp-hi subjects and the "above 60%" selection among the noncomp-lo subjects both fit closely to these predicted values.

While these global results are consistent with matching, analysis of individuals' selection tendencies reveal some important differences. For example, three of the comp-hi subjects (18%), two of the noncomp-lo subjects (13%), and one of the comp-lo subjects (6%) selected the more successful strategy on 43 or more of the last 45 problems. Under a probability-matching model, the probabilities of such extreme preferences for the more successful strategy

are very low. If we assume that subjects are selecting with probability $p$ equal to the matching value of their condition, the expected probabilities for the three situations above are .003, .00001, and .00000005 (computed from the binomial distribution). Thus, it is very unlikely that these subjects are probability matching; instead, they are likely using an "exclusive" approach—selecting the more successful strategy almost exclusively. Figure 3 presents individual subjects' selection data (proportion of last 45 problems on which the more successful strategy was selected) against the average of their solution experiences (proportion of solved problems solved by the more successful strategy) preceding each of those problems. The line y=x, $R^2$=.60, represents the matching law (and, hence, the prediction of any model that leads to asymptotic probability matching), but the data suggest that the majority of subjects are, in fact, overmatching relative to their experience.
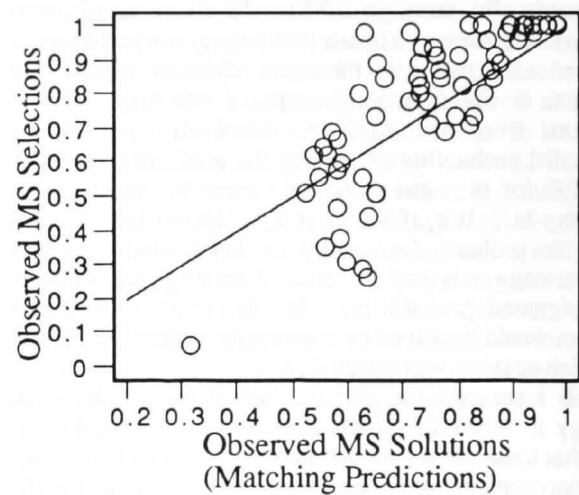


Figure 3. Proportion of the last 45 problems on which the more successful strategy (MS) was selected against the prediction of the matching law, computed for each subject.

Our ACT-R model can be fit to subjects' selection data across all problems. As described above, our model estimates the PPS of each move and then selects the move with the highest PPS value, given some noise is added to each. We set the variance of this Gaussian noise to be $0.05^2$ and obtain the model's predicted probability $p$ of selecting UNDERSHOOT for each problem for each subject. This probability also depends on the individual subject's history of success with UNDERSHOOT and OVERSHOOT preceding that problem. We only varied one model parameter to fit these data, the sum $\alpha+\beta$. The sum $\alpha+\beta$ is used in the model's formula for updating its estimate of each strategy's probability of success. It functions as a "learning rate" for PPS: the larger the sum, the smaller the influence of one success or failure, and the smaller the sum, the larger the influence of one success or failure. The same $\alpha+\beta$ was used for both UNDERSHOOT and OVERSHOOT, with $\alpha$ set to half $\alpha+\beta$. The value 255 for $\alpha+\beta$ minimized the sum of the

---

[4]Note that the RHS of this equation does not include unsolvable problems but the LHS does. Thus, the equation reduces to probability matching when all problems are solvable (i.e., #A_Solutions + #B_Solutions = all trials).
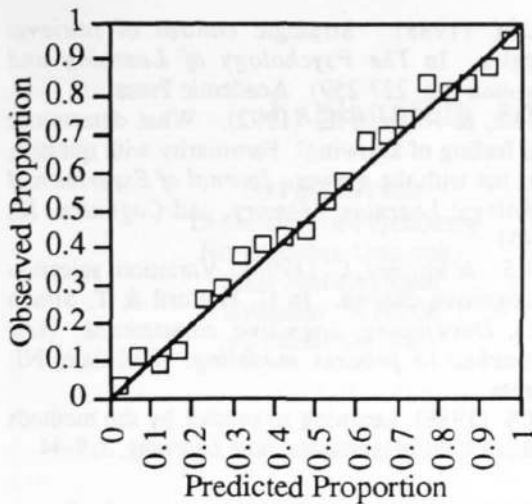
Figure 4. Observed and ACT-R predicted proportions more successful strategy selected across all trials, all subjects.

squared differences between the model's $p$ and the subject's response (UNDERSHOOT or OVERSHOOT) across all subject-problem pairs. Since subjects' responses are a binary variable (UNDERSHOOT or OVERSHOOT) and the model generates a probability, we present the model's fit by "binning" together problems for which the $p$ s are similar. Figure 4 plots the average of the $p$ s in each bin against the observed proportion of UNDERSHOOT selections on the corresponding trials. The line $y=x$, $R^2=.99$, shows that, over all subject-problems, the ACT-R model provides an excellent fit.
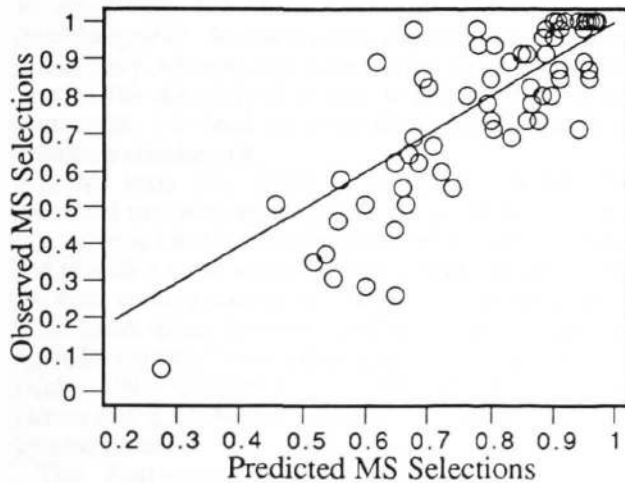


Figure 5. Observed and predicted proportions of more successful strategy selections on last 45 problems, by subject.

Our model can also be compared with the predictions of probability matching (cf. Figure 3) by fitting individual subjects' selections on the last 45 problems. In Figure 5, we plot our model's predicted proportion of selections of the more successful strategy against the observed

proportion of selections of the more successful strategy, averaged over the last 45 problems for each subject. The line $y=x$, $R^2=.66$, shows that the model accounts for the major trend in the data without systematically over- or underpredicting, but there is still much variability. Note that the major difference between Figures 5 and 4 is that Figure 5 displays the model's predictions subject by subject whereas Figure 4 combines the predictions that had similar values, allowing multiple subjects' data to contribute to each bin. A similar difference applies to Figures 3 and 2: Figure 3 presents averages within subjects, and Figure 2 presents averages across subjects. The greater variability and model misfit in Figures 5 and 3 suggest that there are more factors influencing individual strategy selections than most models take into account. By aggregating over subjects, these factors tend to average out.

## Conclusions

The main result of this experiment is that, while group data suggest that solvers are selecting between strategies by probability matching, individual data suggest that there are a variety of selection tendencies. Individualized predictions based on the matching law showed that many subjects tend to "overmatch" (use the more successful strategy more often than its proportion of solutions) and left 40% of the variance in subjects' asymptotic selection behavior unaccounted for. Our ACT-R model does not necessarily predict matching behavior, but it does predict that solvers will tend to prefer the more successful strategy. This prediction stems from the claim that solvers are choosing moves with the highest PPS, and PPS is estimated by their past successes *and failures* with each strategy. However, the model also assumes there is some noise in this process. So, when one strategy's success rate is much higher than another strategy's, the "better" strategy will likely be selected, but when two strategies' success rates are very similar, one will be selected essentially at random. Thus, according to our model, the degree of preference for one strategy over another depends on the relative number of successes and failures experienced with each strategy and the amount of noise in the system. This stochastic selection based on the relative "strengths" of alternatives is similar to several interactive competition models (Gluck & Bower, 1988; Siegler & Shipley, 1995; McClelland & Rumelhart, 1981).

When the ACT-R model was fit to all subjects' selections across the entire experiment, it provided an excellent fit. This shows that the model is fitting subjects' overall selection tendencies as they develop through the course of the experiment. Similar to the "matching" predictions, however, our model left a large portion of the variance unaccounted for when predicting selections at an individual level. This suggests that there are individual differences influencing selection beyond the variability in individual subjects' experiences. The free parameters in our model may allow us to capture such inter-subject differences.

We are currently investigating individual differences that may be significant influences on subjects' selection behavior. One such difference is subjects' learning rate—how much of an impact one success or failure has on later selections. Another is subjects' assessments of the value of exploration versus exploitation—how much utility is attributed to solving the problem versus experimenting with the various strategies. By incorporating these differences into our model, we hope to gain an even better understanding of the processes involved in strategy selection.

## References

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

Atkinson, R.C., & Estes, W.K. (1963). Stimulus sampling theory. In R.D. Luce, R.R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*. Vol. 2. New York: Wiley.

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

Estes, W.K. (1964). Probability learning. In A.W. Melton (Ed.), *Categories of human learning*. New York: Academic Press.

Gluck, M.A., & Bower, G.H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117*, 225-244.

Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*, 267-272.

Humphreys, L.G. (1939). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology, 25*, 294-301.

Jones, M.R. (1971). From probability learning to sequential processing: A critical review. *Psychological Bulletin, 76*, 153-185.

Lordahl, D.S. (1970). An hypothesis approach to sequential predictions of binary events. *Journal of Mathematical Psychology, 7*, 339-361.

Lovett, M.C. (1994). *The effects of history of experience and current context on problem solving*. Unpublished doctoral dissertation. Carnegie Mellon University, Pittsburgh.

Lovett, M.C., & Anderson, J.R. (1995). *History of success and current context in problem solving: Combined influences on operator selection*. Manuscript submitted for publication.

Luchins, A.S. (1942). Mechanization in problem solving, *Psychological Monographs, 54*, 248.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of the basic findings. *Psychological Review, 88*, 375-402.

Reder, L.M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*, 90-138.

Reder, L.M. (1988). Strategic control of retrieval strategies. In *The Psychology of Learning and Motivation* (pp. 227-259). Academic Press.

Reder, L.M., & Ritter, F.E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 435-451.

Siegler, R.S., & Shipley, C. (1995). Variation, selection and cognitive change. In G. Halford & T. Simon (Eds.), *Developing cognitive competence: New approaches to process modeling*. Hillsdale, NJ: Erlbaum.

Sutton, R.S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning, 3*, 9-44.