Journal of Experimental Psychology: General
1992, Vol. 121, No. 2, 177-194

0096-3445/92/$3.00

An Adaptive Approach to Human Decision Making: Learning Theory, Decision Theory, and Human Performance

Jerome R. Busemeyer and In Jae Myung
Psychological Sciences, Purdue University

This article describes a general model of decision rule learning, the rule competition model, composed of 2 parts: an adaptive network model that describes how individuals learn to predict the payoffs produced by applying each decision rule for any given situation and a hill-climbing model that describes how individuals learn to fine tune each rule by adjusting its parameters. The model was tested and compared with other models in 3 experiments on probabilistic categorization. The first experiment was designed to test the adaptive network model using a probability learning task, the second was designed to test the parameter search process using a criterion learning task, and the third was designed to test both parts of the model simultaneously by using a task that required learning both category rules and cutoff criteria.

Probabilistic categorization is an important class of decision problems in which stimuli are sampled from a number of categories and the decision maker must decide the category from which each stimulus was sampled. Payoffs depend on both the true category membership and the decision maker's response for each stimulus. Examples are found in all areas of psychology: In perception, auditory or visual stimuli are categorized as signal or noise, and in memory recognition, verbal items are categorized as old or new. In cognition, exemplar patterns are assigned to conceptual categories, and in industrial psychology, job applicants are categorized as acceptable or unacceptable. Finally, in clinical psychology, patient symptom patterns are assigned to disease categories.

For the past 35 years, the general theory of signal detection (Peterson, Birdsall, & Fox, 1954) has served as the most prominent model of probabilistic categorization. It has been successfully applied to all of the areas of psychology mentioned (see Green & Swets, 1966; for perception; Bernbach, 1967, and Wickelgren & Norman, 1966, for memory recognition; Ashby & Gott, 1988, for conceptual categorization; Cronbach & Gleser, 1965, for industrial psychology; and Swets & Pickett, 1982, for medical diagnosis). The core idea is that (a) each stimulus is represented as a point within a multidimensional stimulus space, (b) this multidimensional space is partitioned into response regions, and (c) a stimulus is categorized according to the region within which it lies.

Simple decision rules are normally used to describe how the stimulus space is partitioned.[1] For example, unidimensional stimuli can be divided into two categories by either a cutoff rule (all points above a cutoff go into one category) or by an interval rule (all points inside an interval go into one category). Two-dimensional stimuli can be partitioned into two category regions by either a linear rule (all points above a line go into one category) or an elliptical rule (all points within an ellipse go into one category). Many other decision rules are possible, and consequently, a critical question is, how are decision rules selected?

One answer is that the optimal rule is selected. This is the rule that maximizes expected payoffs, where the expected payoff is defined as the long-run average payoff produced by using the same rule for an infinitely long sequence of trials under a fixed training condition. The optimal rule will vary depending on the training condition, which specifies (a) the distribution of stimuli produced by each category, (b) the prior probability of each category, and (c) the payoffs produced by each stimulus–response category combination.

The optimal rule can be rejected as a descriptive model of human performance on the basis of past research. One robust finding is the conservative cutoff-placement phenomenon—subjects tend to select for the cutoff rule a cutoff criterion that is less extreme than that prescribed by the optimal rule (Green & Swets, 1966, p. 90; Healy & Kubovy, 1981). Another finding is that when two-dimensional stimuli are used, subjects tend to use linear categorization rules even when an elliptical rule is optimal (Ashby & Gott, 1988, p. 51).

An alternative way to understand how decision rules are selected is to consider learning mechanisms. For example, with more training, the subjects in Asbby and Gott's (1988)

This research was supported by National Science Foundation Grant BNS-8710103. We thank Rob Nosofsky and Mark McDaniel for their helpful comments.

In Jae Myung is now at the Department of Psychology, Ohio State University.

Correspondence concerning this article should be addressed to Jerome R. Busemeyer, Psychological Sciences, Purdue University, West Lafayette, Indiana 47907-1364

[1] A simple decision rule for binary categorization tasks with continuous distributions of stimuli is defined rigorously as follows. Let S be the set of all stimulus points that are to be categorized, and let C be the set of category responses. A decision rule, R, is a function that maps points in S onto elements of C. For example, if s is a point in S and c is an element of C, then $R(s) = c$. The inverse image of c, $R^{-1}(c)$, is defined as the set of points in S that map to the element c. The inverse image, $R^{-1}(c)$, is convex if and only if the following is true: Suppose x and y are members of $R^{-1}(c)$. Then $z = (1 - a)x + ay$, $0 < a < 1$, is also a member of $R^{-1}(c)$. A decision rule is simple if and only if the inverse image, $R^{-1}(c)$, is a convex set for at least one of the category responses c in C. This definition covers the majority of previous examples of decision rules used by signal-detection theorists.

experiment may have been able to learn to switch from the suboptimal linear rule to the optimal elliptical rule. A decision-rule–learning model can be used to predict whether the optimal rule will be learned and, if the rule is learned, the amount of training needed to learn it.

There are various lines of evidence for rule learning in probabilistic categorization. Criterion learning models have been used to explain (a) improvements in sensitivity with training (e.g., Swets & Sewall, 1963); (b) sequential effects of earlier trial events on subsequent responses (Atkinson & Kinchla, 1965); and (c) the conservative cutoff-placement phenomenon (Thomas, 1973; see Dusoir, 1980, and Kubovy & Healy, 1980, for reviews). However, nearly all of this past work was limited to the question of how individuals learn the cutoff parameter for the unidimensional cutoff rule. The model we present in this article extends the earlier work by describing not only how the parameters of a rule are learned but also how one decision rule comes to be preferred over others. A simple example helps illustrate the basic ideas involved.

Suppose subjects are asked to make fictitious clinical decisions in which they must decide whether Disease A or B is present on the basis of some unidimensional test score X, and they receive a monetary payoff that depends on the correctness of their decision. The following are four different decision rules (labeled F, G, H, and I) for performing this task:

F: If $X < \theta$, then report A; otherwise report B.

G: If $X > \theta$, then report A; otherwise report B.

H: If $X > -\theta$ and $X < \theta$, then report A; otherwise report B.

I: If $X < -\theta$ or $X > \theta$, then report A; otherwise report B.

Rules F and G are cutoff rules, and Rules H and I are interval rules. Note that each rule contains a parameter, θ, the value of which is unspecified. We distinguish between a general rule, where the parameters of the rule are unspecified, and a specific rule, in which all of the parameters are specified.[2] The primary question is, how do individuals learn to prefer one general rule over others?

This simple example illustrates many of the complexities of decision rule learning. After being informed about the training condition, the decision maker has two different learning tasks. One is to learn the optimal value of the criterion parameter, θ, that maximizes expected payoffs for each general rule. The second is to learn the optimal rule (F, G, H, or I) for a given training condition. We designed the following *rule competition* model to describe how individuals solve these two learning problems.

Rule Competition Model

The rule competition model is composed of two different learning processes: (a) an adaptive network model that learns to predict the payoffs produced by each general rule and (b) a hill-climbing model that searches for the optimal parameters of a general rule.

The rule competition model uses both the adaptive network learning model and the hill-climbing search model on each trial as follows. First, the adaptive network model selects a general rule, and then the hill-climbing model selects a parameter to form a specific instance of the general rule. To help illustrate the model, recall the previous example, where the decision maker considers a set of four general rules {F, G, H, and I}.

First, each of the four general rules is evaluated by the adaptive network model as follows. The decision maker recalls each of the four rules, and each rule is input into the adaptive network for evaluation, one at a time, yielding performance estimates (U_1, U_2, U_3, and U_4) for Rules F, G, H, and I, respectively, which are then stored in a short-term memory. The probability of choosing one rule, for example Rule j, is an increasing function of its estimated performance, U_j, and a decreasing function of the estimated performance of each of the other rules, U_k ($j \neq k$). More specifically, define $P(R_j)$ as the probability of choosing Rule j from the set of n rules. The following ratio model proved to be helpful in using categorization models to predict choice probability (Estes, Campbell, Hatsopoulos, & Hurwitz, 1989; Gluck & Bower, 1988; McClelland & Rumelhart, 1985; Nosofsky, 1987):

$$P(R_j) = \exp(U_j)/\textstyle\sum_k \exp(U_k), \qquad \text{for } k = 1, \ldots, n. \quad (1)$$

Second, the parameter θ for the general rule selected by the previously mentioned adaptive network is assigned a specific value as follows. If a general rule was not used on the previous trial, then the parameter remains unchanged. If a general rule was used on the previous trial, then its parameter is updated following feedback according to the hill-climbing search model.

This completes the general description of the rule competition model. Next, we describe the adaptive network and then the hill-climbing parts of the model.

Adaptive Network Model

The adaptive network model learns to estimate the performance of a general rule for each particular situation. The term *situation* refers to the training condition defined by the distribution of stimuli, prior probabilities, and payoff matrices.

The adaptive network model is illustrated in Figure 1. (The mathematical details are given in the Appendix). The performance of a general rule is estimated in two layers. Initially, information about the rule along with information about the current situation are entered into the top layer of the network. This *situation X rule* input then elicits estimates of the probabilities of each possible payoff at the middle nodes following the first layer. Finally, these probability estimates are inte-

[2] General and specific rules are defined rigorously as follows. First, the domain of a rule given in Footnote 1 needs to be expanded to include parameters. Define T_i as the set of parameter vectors associated with rule i. Then rule i (R_i) is a function that maps stimulus points from S and parameter vectors from T_i onto C. A specific instance of rule i is obtained by fixing the vector taken from the parameter space. In this case, the specific rule is a function that maps points from S onto C.

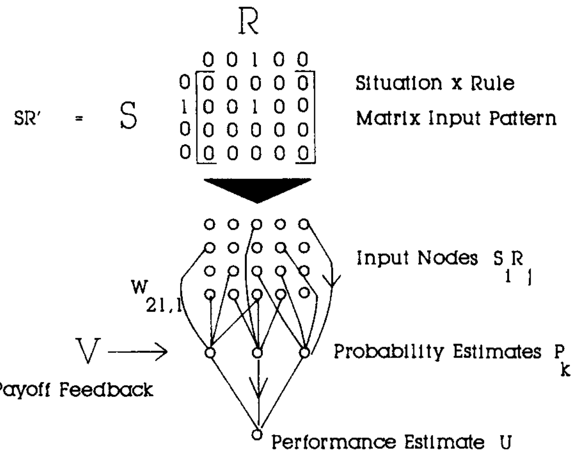

Figure 1. Adaptive network model of decision rule learning. (The top nodes receive inputs from combinations of situation [S] and rule [R] features. The middle nodes compute the probability distribution [P] over payoffs on the basis of the connection weights [W] from the input nodes. The final node computes the estimated performance of a rule [U] on the basis of the connections from the middle nodes. Rule choice is based on the output of the final node. This figure illustrates the input matrix for a simple example with binary valued inputs. In general, the inputs may have continuous activation values. V = payoff feedback.)

grated into a single performance estimate, U, at the terminal node following the second layer.

The weights connecting the first and second layers in Figure 1 are updated according to a Hebb-delta learning rule, which is a generalization of the adaptive network learning models used by Gluck and Bower (1988), Estes et al. (1989), and Knapp and Anderson (1984). (See Appendix for the mathematical details.) Each weight connecting the middle and final nodes represents the subjective worth of a payoff value. The latter weights are not updated during training.

Suppose a general rule X was applied on trial t, producing the performance estimate $U_X(t)$ before feedback on trial t. After receiving feedback on trial t, a payoff is delivered, and $v(t)$ is the subjective worth of this payoff. On the next trial, $t + 1$, another general rule (Y; note that X may equal Y) is evaluated, which results in a new performance estimate $U_Y(t + 1)$. In the Appendix, we prove that the following simple equation can be mathematically derived from the adaptive network learning model[3]:

$$U_Y(t + 1) = \beta \cdot U_Y(t) + \alpha \cdot s(t + 1) \cdot [v(t) - \gamma \cdot U_X(t)]. \quad (2)$$

Equation 2 states that the new estimate of Rule Y after feedback on trial t is proportional to its estimate before feedback on trial t, $U_Y(t)$, plus an adjustment, $[v(t) - \gamma \cdot U_X(t)]$, multiplied by a gain factor, $\alpha \cdot s(t + 1)$. The adjustment equals the previous prediction error resulting from the use of Rule X on trial t, and the gain is proportional to a similarity measure, $s(t + 1)$. For a fixed situation, $s(t + 1)$ is an increasing function of the similarity between the new rule and the previous rule. For a fixed rule, $s(t + 1)$ is an increasing function of the similarity between the new situation and the previous situation. The constant $\beta > 0$ determines the forgetting rate, the constant $\alpha > 0$ determines the learning rate, and the constant $1 \geq \gamma \geq 0$ determines the weight of the delta rule relative to the Hebb rule. Setting $\gamma = 0$ produces a pure Hebb learning rule, and setting $\gamma = 1$ produces a pure delta learning rule.

Variation in similarity between rules can be illustrated by considering Rules F, G, H, and I, which we described earlier. Rules F and G are both single-cutoff rules, whereas Rules H and I are both interval rules. These similarity relations can be represented by setting the similarity parameter, $s(t + 1)$, equal to non-zero values for related pairs of rules (e.g., F and G) and zero for unrelated pairs of rules (e.g., G and H). Maximum similarity is obtained when the same rule that was applied on the previous trial is evaluated again on the next trial.

Hill Climbing

Consider Rule G, the single-cutoff rule, described in the previous example. For a fixed training condition (i.e., distribution of stimuli, prior probabilities, and payoff matrix), the performance of this general rule is a function of the value of the cutoff parameter θ. If θ is set too high or too low, then poor performance may result even if Rule G is optimal in the general sense. The hill-climbing model is used to search for a parameter value that produces the highest performance for a general rule.

Figure 2 illustrates the basic idea. Hill climbing is guided by the change in payoffs produced by a change in the parameter from two prior applications of a rule. If the previous change in the parameter produced an increase in payoffs (an uphill direction), then the next change is made in the same direction. However, if the previous change in the parameter produced a decrease in payoffs (a downhill direction), then the next change is in the opposite direction.

More precisely, define θ_1 and θ_2 as the parameter values that were used to make category decisions on the last and second-to-last applications of Rule G, respectively. Also define v_1 and v_2 as the values of the payoffs produced on the last and second-to-last applications of Rule G, respectively. The product of differences

$$h = (v_1 - v_2) \cdot (\theta_1 - \theta_2)$$

is called the *hill-climbing adjustment.* It is used to determine the next parameter value as follows:

$$\theta = \theta_1 + \lambda \cdot m(h) + (1 - \lambda) \cdot \epsilon. \quad (3)$$

In this equation, θ is the parameter value to be used on the current trial; θ_1 is the previously used parameter value; $m(h)$ is a bounded increasing function of the hill-climbing adjust-

[3] We are not stating that Equation 2 represents the psychological learning process. The psychological learning process is represented by the Hebb-delta learning rule (see Appendix), and Equation 2 is a mathematical consequence of this process. However, for the purpose of deriving predictions for the experiments reported elsewhere in this article, Equation 2 is sufficient and more convenient.

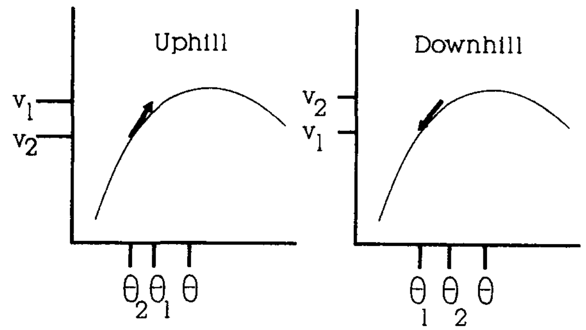

Figure 2. The hill-climbing algorithm. (The left panel illustrates an increase in the parameter producing an increase in payoff, and the right panel shows a decrease in the criterion producing a decrease in the payoff. θ is the value of the criterion to be used on the next application of a rule. θ_1 and θ_2 are the values of the criterion used on the last and second-to-last applications of the rule. V_1 = payoff from last trial; V_2 = payoff from second-to-last trial.)

ment, h; ϵ is a randomly chosen direction; and λ is a weight ($0 < \lambda < 1$).[4]

The function $m(h)$ stabilizes the learning algorithm by squashing extreme adjustments. The simplest example is the step function: $m(h) = c$ if $h > 0$, $m(h) = 0$ if $h = 0$, and $m(h) = -c$ if $h < 0$. The random direction ϵ reduces the likelihood that the search gets stuck on a local maximum. The amount of random search is moderated by the weight $(1 - \lambda)$.

So far, we have discussed the hill-climbing model in terms of the single cutoff rule (Rule G). However, hill climbing is a very general search model, and the same model can be applied to any rule containing a single parameter. Hill climbing can also be used for rules that contain n parameters by defining θ, θ_1, θ_2, and ϵ as n-dimensional vectors. In this case, the function m is applied separately to each of the n coordinates of the hill-climbing adjustment vector.

Outline of Experiments

We designed the following experiments to test the rule competition model and to compare it with various other learning models. Because the rule competition model is composed of two distinct parts, we designed three experiments: Experiment 1 was designed to test only the first part, the adaptive network model; Experiment 2 was designed to test only the second part, the hill-climbing model; and Experiment 3 was designed to test both parts, that is, the combined model that uses both hill-climbing and adaptive network learning models simultaneously.

Experiment 1

The rule competition model is partly based on an adaptive network model that learns the probabilities of the payoffs produced by each decision rule. There was a much earlier work on probability learning, and it is useful to compare briefly the present model with these earlier models.

In the late 1950s, linear learning models were developed to account for probability learning (see Sternberg, 1963, for a review). According to these earlier models, individuals learn to associate category responses to stimuli. These stimulus-response learning models have been largely rejected as overly simplistic because of their limited computational capabilities (e.g., Anderson, 1976; Minskey & Papert, 1969).

The present model differs from these earlier models in a fundamental way. According to the rule competition model, individuals learn to associate payoffs with decision rules (category responses are computed from these decision rules). Because there are no major limitations on the complexity of these decision rules, the current model generalizes and extends the computational ability of the earlier probability learning models.

The present model uses a Hebb-delta learning rule to describe how individuals learn the payoff probabilities produced by each decision rule. However, there are two alternative learning models that also need to be considered—the *frequency array model* (Estes, 1987) and the *multiple-trace model* (Busemeyer, 1985). It is important to consider these two particular models for several reasons. First, both of these earlier models have successfully predicted choice behavior in decision tasks that require learning from experience. Second, these models are based on learning and memory assumptions that are quite different from those used by the adaptive network model. Finally, these earlier models make predictions that differ at an ordinal level from the predictions of the Hebb-delta model, and the three models have never been directly compared. The following clinical trial probability-learning experiment illustrates the differences among these three models.

The subjects of Experiment 1 were told that a population of patients was suffering from a common set of symptom patterns. The subjects' task was to assign patients from this population to treatment strategies. After each assignment, the effect of the treatment was reported as feedback. The effectiveness (or payoff) produced by each treatment varied from patient to patient, and subjects had to learn the distribution of treatment effects produced by each treatment.

The distribution of payoffs produced by each treatment is illustrated in Figure 3. The horizontal axis is the payoff value, and the vertical axis is the probability density. All three distributions are normal with equal variance, σ^2, but different means: The mean payoff for Treatment 1 is $\mu - \delta$, the mean for Treatment 2 is μ, and the mean for Treatment 3 (the optimal treatment) is $\mu + \delta$. This information was not shown to the subjects, and it had to be learned on the basis of trial-by-trial feedback.

Frequency Array Model

According to this model, individuals accurately store the frequency of each payoff value produced by each treatment strategy, producing a memory array of frequency estimates. The sample mean payoff of a treatment is estimated from this relative frequency distribution.

---

[4] Myung (1990) proved a theorem stating that the mean criterion selected by the hill-climbing algorithm converges to the optimal criterion when certain conditions on the learning rate and the distribution of payoffs are met.

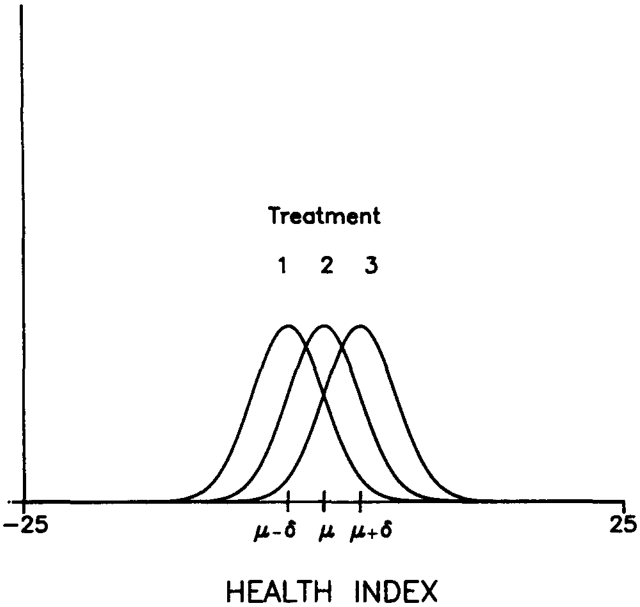

Figure 3. The distribution of treatment effects for patients assigned to Treatments 1, 2, and 3 used in Experiment 1. (All three distributions are normal, with homogeneous variance and equal differences between the means of the center and extreme distributions.)

Define $f_{ij}(t)$ as the estimate of the frequency with which the payoff value v_i was obtained from treatment j after a total of t training trials (the frequency in row i, column j of the memory array). If treatment j is chosen and payoff v_i is obtained on the next trial, then its corresponding frequency is incremented, $f_{ij}(t+1) = f_{ij}(t) + 1$; otherwise it remains unchanged, $f_{ij}(t+1) = f_{ij}(t)$.

The column total, $f_j(t) = \sum_i f_{ij}(t)$, gives the frequency with which treatment j was chosen out of a total of t training trials. The weighted average

$$M_j(t) = \sum_i f_{ij}(t) \cdot v_i / f_j(t)$$

defines the sample mean payoff produced by treatment j during the preceding t training trials. The probability of choosing the optimal treatment is assumed to be given by the ratio rule,

$$P(R_3) = \frac{\exp(\lambda \cdot M_3)}{[\exp(\lambda \cdot M_1) + \exp(\lambda \cdot M_2) + \exp(\lambda \cdot M_3)]}, \quad (4)$$

where λ is a constant.

A key idea of this model is that the probability of choosing the optimal treatment (Treatment 3) is limited by the probabilistic nature of the decision process rather than the learning process. If subjects accurately track the frequencies, then as training progresses, the sample mean, $M_j(t)$, rapidly converges to the population mean, μ_j (cf. Kleyle & De Korvin, 1988). In this limiting case, the probability of choosing the optimal treatment is given by

$$P(R_3) = \exp(\lambda \cdot \mu_3)/[\exp(\lambda \cdot \mu_1) + \exp(\lambda \cdot \mu_2) + \exp(\lambda \cdot \mu_3)]$$
$$= \exp(\lambda \cdot \delta)/[\exp(-\lambda \cdot \delta) + 1 + \exp(+\lambda \cdot \delta)].$$

One important implication of this model is that the effect of the variance (σ^2 in Figure 3) on choice probability should be a strictly decreasing function of training, and in fact, the asymptotic choice probability should be solely a function of the mean difference (δ in Figure 3). This hypothesis is inconsistent with previous research by Busemeyer (1985), who found that the effect of variance on choice probability failed to decrease consistently. Instead, there was a substantial effect of variance that remained constant throughout the later stages (after 2,000 trials) of training. Therefore, the frequency array model can be ruled out for this paradigm.

Multiple-Trace Model

To account for the extended effect of payoff variance on choice, Busemeyer (1985) proposed the following model. Each time a payoff is experienced following the application of a treatment, a separate memory trace of the event is stored in memory. At the beginning of a choice trial, each treatment strategy serves as a retrieval cue that retrieves a fixed sample of memory traces for each treatment. The payoff values of the traces retrieved from memory are averaged, and the treatment producing the largest fixed sample mean is always chosen.

According to this model, the decision process is deterministic, and random variability only enters through the memory retrieval process. The probability of choosing the optimal treatment is limited by the fixed sample size—if only a limited number of the memory traces can be retrieved, then the sample mean will continue to vary randomly even after extensive training.

Define m_j as the fixed sample mean of n traces sampled from memory for treatment j. The probability of choosing the optimal treatment (Treatment 3) is given by

$$P(R_3) = P[m_3 = \max(m_1, m_2, m_3)].$$

For the payoff distribution shown in Figure 3 (see Bock and Jones, 1968), it follows that

$$P(R_3) = 1 - F(-2z, -z), \quad (5)$$

where F is the standard bivariate cumulative normal distribution function with a correlation equal to .5, and $z = (\delta/\sigma)/\sqrt{(2/n)}$. An important implication of this model is that the probability of choosing the optimal treatment should be a monotonic function of the ratio (δ/σ), and this relation should hold at each trial of training.

Adaptive Network Model

The adaptive network model makes predictions that range between the two extremes of the frequency array and multiple-trace models, depending on the difference between the learning and forgetting rates. The estimated performance for each treatment can be expressed as a weighted average of all the past payoffs, with weights that are an exponentially decreasing function of the lag (the number of trials intervening between the current trial and the trial on which a payoff was delivered).

(See Equation A6 of Appendix). For extremely low learning rates, the weight function decreases very slowly with lag. In this case, extensive training will yield a performance estimate very close to the mean payoff, similar to the frequency array model, and the variance of the payoffs will have almost no effect on asymptotic choice probability. For extremely high learning rates, the weight function decreases very rapidly with lag. In this case only a limited number of the most recent observations receive significant weight, similar to the multiple-trace model. For moderate learning rates, the variance of the payoffs will have a persistent effect on choice probability, but the effect is small relative to the effect of the mean difference (e.g., see Table 1, part B). Although the learning rate may vary widely across subjects, the average rate is expected to be in the moderate range.

Summary

The three models make ordinally distinguishable predictions concerning the asymptotic probability of choosing the optimal treatment. Consider the factorial design illustrated in Table 1, part A, where the cells of this table contain the ratio (δ/σ). The frequency array model predicts that asymptotic choice probability is independent of σ, and consequently, it can be ruled out for this paradigm on the basis of the results reported by Busemeyer (1985). The multiple-trace model predicts that choice probability is a function of the ratio (δ/σ), and this ratio decreases down the main diagonal of Table 1, part A. The adaptive network model predicts that choice probability is a function of both δ and σ but is mainly influenced by δ; note that δ increases down the main diagonal. Thus, the key comparison for Experiment 1 is obtained by focusing on the conditions along the main diagonal in Table 1 where the multiple-trace model (see Table 1, part A) and adaptive network model (see Table 1, part B) predict opposite rank orders.

Table 1
Probability of Choosing the Optimal Treatment for Experiment 1

Standard deviation	Mean difference (δ)			
(σ)	2.0	2.5	3.0	*M*
A. Ratio of (δ/σ)				
3.0	**.67**	.83	1.00	.83
4.5	.44	**.56**	.67	.56
6.0	.33	.42	**.50**	.42
M	.48	.60	.72	.60
B. Probabilities predicted by adaptive network				
3.0	**.67**	.74	.80	.73
4.5	.65	**.70**	.79	.71
6.0	.65	.65	**.81**	.70
M	.66	.70	.80	.72
C. Probabilities observed on last 10 trials				
3.0	**.69**	.84	.85	.79
4.5	.69	**.72**	.84	.75
6.0	.65	.63	**.86**	.71
M	.68	.73	.85	.75

Note. The boldface values are crucial for testing the multiple-trace and adaptive network models.

Method

Procedure. Subjects were asked to imagine that they were physicians and that they had to choose one of three possible treatments on each trial. At the beginning of training for each condition, they had no information about the effectiveness of each treatment, and they were told that they should assume that each treatment was equally likely to be the best. Following the choice of a treatment, they were shown a randomly sampled patient number and the effect of the treatment on that patient only. They were told that the objective was to maximize the sum of the treatment effects over training and that their pay would be proportional to the final sum. The subjects were paid 4¢ per point.

Subjects initially received verbal instructions followed by two brief practice sessions of 5 trials each. After practice, each subject received nine blocks of training (50 trials per block), with a new experimental condition randomly assigned to each block (with the constraint that each condition appeared in each serial position with equal frequency across subjects). Subjects received three experimental conditions, followed by a 3-min break, another three experimental conditions, another 3-min break, and then the final three experimental conditions. The entire experiment lasted approximately 1.5 hr. Subjects were run individually in a quiet room. An IBM personal computer was used to present stimuli and record responses.

Design. The nine problems for the main part of the experiment were constructed by manipulating two factors, mean difference and standard deviation, according to the 3 × 3 factorial design shown in Table 1, part A. Note that standard deviation manipulation produces ratios in the top row that are 2 times larger than those in the bottom row. Also note that the mean difference manipulation produces ratios in the right column that are only 1.5 times larger than those in the left column. On this basis, one would expect the effect of standard deviation to be greater than the effect of mean difference on learning. The mean differences and standard deviations of the treatment effects for the practice sessions were ($\delta = 3$, $\sigma = 1$) and ($\delta = 3$, $\sigma = 3$).

The three treatment means were constant across the 50 trials for a given condition, but they varied randomly across conditions and subjects. The means for Treatments 1 and 3 were always δ units below and above, respectively, the mean for Treatment 2. The mean of Treatment 2 for any given condition was sampled from a uniform [−2, +2] distribution.

Subjects were informed when each condition began and ended. At the beginning of each condition, new labels were assigned to the treatments according to the following procedure. Nine triples of three adjacent keys on the computer keyboard formed a set of labels, and one triple was randomly sampled without replacement for each condition. One letter within each triple was randomly sampled without replacement for each treatment within a condition. The labels used in practice differed from those used in the main experiment. For the reader's convenience, hereinafter we label Treatment 3 as the *optimal treatment*, Treatment 1 as the *worst treatment* and Treatment 2 as the *intermediate treatment.*

Subjects. The subjects were 36 students from Purdue University (graduates and undergraduates majoring in humanities and social sciences) who were paid volunteers. The amount of pay was contingent on performance, as described above, but on the average subjects earned about $5.00 per hour.

Results

Final performance. Recall that the rank order predictions for the multiple-trace model are determined from Table 1, part A, and the predictions from the adaptive network model are shown in Table 1, part B. The cells of Table 1, part C show the observed proportion of optimal choices separately for each of the nine conditions averaged across subjects and the last 10 trials of training. (Each cell proportion is based on $n = 360$ observations.) The margins of this table represent the row and column averages. The key comparison is obtained from the three conditions shown along the main diagonal of Table 1, part C.

As can be seen in the main diagonal of Table 1, part C, the proportions of optimal choices consistently increased from .69 (top left corner) to .86 (bottom right corner) as the ratio (δ/σ) decreased from .67 to .50, and as the mean difference δ increased from 2 to 3. This result was consistent across most subjects, and the difference between the two extreme diagonal cells is significant according to a simple sign test. A total of 25 subjects produced unequal choice proportions under these two conditions (11 subjects produced tied proportions equal to 1.0). Eighteen of these 25 subjects (72%) produced an ordering for these two conditions in the same direction as that observed in Table 1, part C, which is significantly different from chance ($p = .0216$).

More generally, the observed proportions correlated more strongly with the predictions of the adaptive network model (Table 1, part B), than with the predictions of the multiple-trace model (Table 1, part A). The rank correlation with the ordinal predictions of the multiple-trace model in Table 1, part A, equals .67; the rank correlation between the predictions of the adaptive network model in Table 1, part B, equals .93.

One could argue that perhaps the multiple-trace model would make different ordinal predictions for the diagonal cells of Table 1, part C, if the assumptions about the choice rule were changed. In particular, one could assume that the probability of choosing the optimal treatment is given by the ratio rule (Equation 4), with M_j now defined as the n most recent payoffs produced by treatment j. This alternative was evaluated by estimating the two unknown parameters (λ and n) from the learning data for each subject. However, this revised model also failed to predict the observed ordering of the diagonal cells shown in Table 1, part C. Therefore, the problem lies with the use of the fixed sample mean.

Training effects. So far, the frequency array model was eliminated on the basis of previous research by Busemeyer (1985), and the multiple-trace model is eliminated by the results of Table 1, part C. This leaves the adaptive network model as the only model under consideration that cannot be ruled out on the basis of parameter-free ordinal tests. However, we have not yet evaluated the quantitative accuracy of the adaptive network predictions for learning. The following analyses answer this question.

Figure 4 shows the relative frequency with which each treatment was chosen as a function of trial block separately for each condition. The curves with circles indicate Treatment 3, the curves with triangles indicate Treatment 2, and the curves with asterisks indicate Treatment 1. The panels in Figure 4 are organized according to Table 1, and the ratio (δ/σ) is indicated in the upper left corner of each panel.

A 3 (mean difference) × 3 (standard deviation) × 10 (trial block with five trials per block) repeated measures analysis of variance was performed. The main effects of trial block, mean difference, and standard deviation were statistically significant, $F(9, 315) = 53.07$, $MS_e = .09$, $p < .001$; $F(2, 70) = 14.3$, $p < .0001$; and $F(2, 70) = 3.92$, $p < .03$, respectively. The Mean Difference × Trial Block interaction effect was also significant, $F(18, 630) = 1.71$, $p < .05$. No other interaction effects were significant.

To evaluate the quantitative predictions of the adaptive network model, we obtained maximum likelihood estimates of the model parameters separately for each subject. A two-parameter model was fit to the trial-by-trial choices for all 50 trials and all nine conditions. The choice probabilities were computed from Equations 1 and 2, with $s(t + 1) = 1$ or $s(t + 1) = 0$ for each treatment, depending on whether that treatment was chosen on the previous trial. This left only two parameters, α and $\eta = (\beta - \gamma \cdot \alpha)$, which needed to be estimated from the data (see the Appendix, Equation A5, for an explanation of η). The means and standard deviations (pooled across 36 subjects) were (.16, .11) for α and (.71, .31) for η.

The Pearson correlation between the predictions of this two-parameter model and the 270 proportions reported in Figure 4 equaled .98. A plot of the predicted proportions is not shown because the high degree of overlap makes it difficult to discriminate between the predicted and the observed learning curves. Table 1, part B, shows the numerical predictions for the last 10 trials of training. The Pearson correlation between the predicted and observed proportions in Table 1, parts B and C equals .96. In conclusion, the adaptive network model provided very accurate quantitative fits to the observed learning curves.

Discussion

Ideally, theories should be tested on the basis of parameter-free ordinal tests of basic properties. This is the strategy that we followed in comparing the frequency array, multiple-trace, and adaptive network models.

The frequency array model predicts that the effect of payoff variance on choice probability should be a strictly decreasing function of training. Previous research by Busemeyer (1985), however, showed that even after very extensive training (over 2,000 trials), the effect of variance on choice probability remained constant rather than strictly decreasing.

The multiple-trace model was proposed to account for the persistent effect of payoff variance on choice probability. This model predicts that choice probability should be a strictly increasing function of the ratio (δ/σ) at each point of training. The results of the present experiment show ordinal violations of this prediction—the choice proportions increased down the main diagonal of Table 1, part C, rather than decreasing as predicted by the multiple-trace model.

This leaves the adaptive network model as the only remaining candidate of the three models under consideration. The adaptive network model is consistent (at the ordinal level) with the research by Busemeyer (1985) and Estes (1987) and with the results of the present experiment. For moderate values of the learning rate parameter, the adaptive network model predicts that payoff variance will have a persistent effect on choice probability, but the effect of variance is

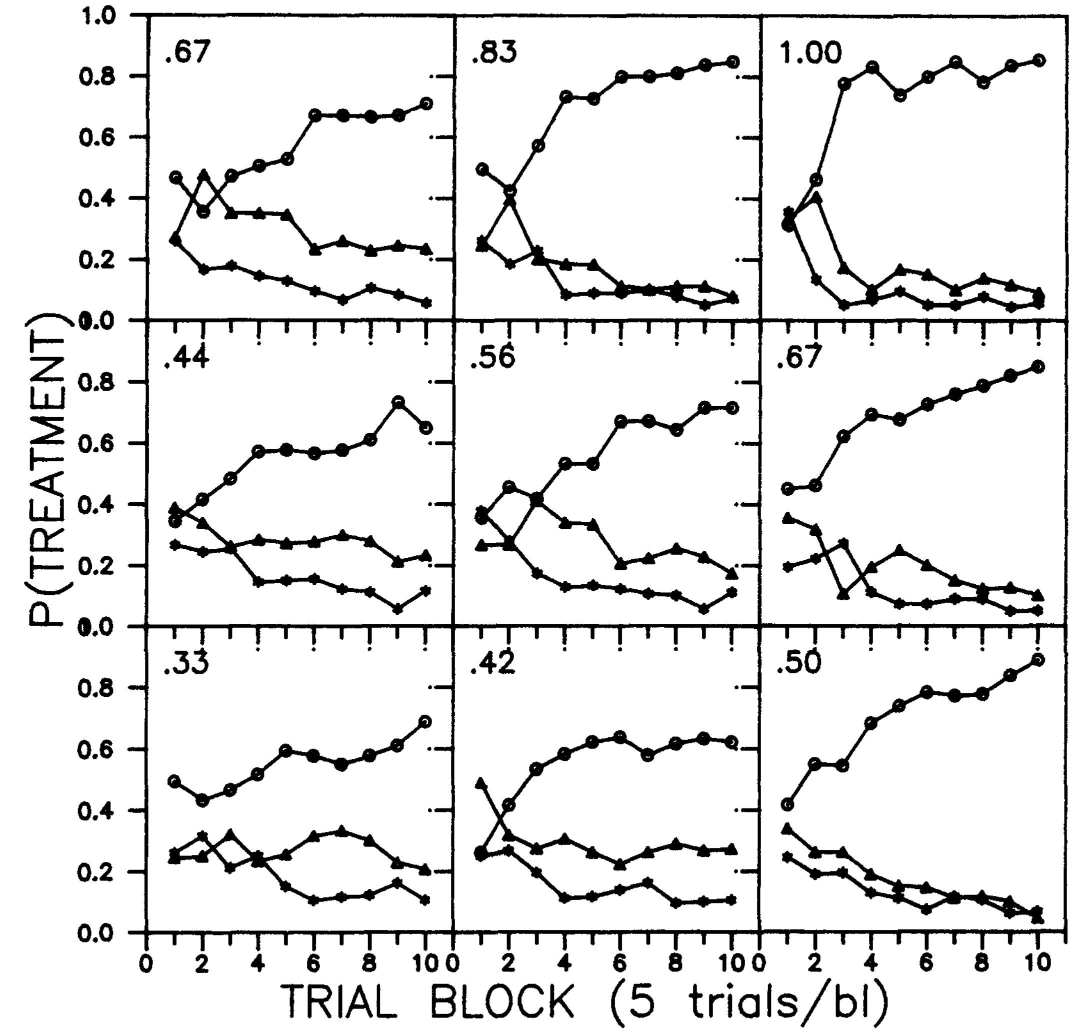

Figure 4. Observed choice proportions (P) plotted as a function of training block (5 trials per block [bl]) with a separate curve for each treatment. (The curves with circles represent Treatment 3 [optimal treatment], the curves with triangles represent Treatment 2, and the curves with asterisks represent Treatment 1. Each row represents the results for a different level of variance. Each column presents the results for a different level of mean difference. The number in the upper left corner is the ratio [δ/σ].)

predicted to be smaller than the effect of the mean difference, so that choice probability was expected to be more strongly correlated with δ than with the ratio (δ/σ).

Although our model-testing strategy was based on parameter-free ordinal tests, we also believe it is important to evaluate the quantitative accuracy of a model using a relatively small number of free parameters. We found that the adaptive network model provided highly accurate quantitative fits to the observed learning curves—96% of the variance of 270 data points was accounted for by a two-parameter learning model.

Experiment 2

The rule competition model is based partly on a general purpose parameter search model called *hill climbing.* An alternative parameter search model called *error correction* was developed by Kac (1962), Dorfman and Biderman (1971), Norman (1972), and Thomas (1973) specifically for the unidimensional, single-cutoff rule. The hill-climbing and error-correction models make qualitatively different predictions regarding the effects of payoff manipulations on criterion learning. Therefore, Experiment 2 investigated how individuals learn the cutoff criterion for a fictitious medical categorization task under different payoff matrix conditions. First we describe the task, and then we describe the predictions for the optimal, error-correction, and hill-climbing models.

Suppose that there are two populations of patients, A and B, and each one has a normal distribution of diagnostic test scores (X) with different means (μ_A, μ_B) but a common variance σ^2. Figure 5 is an example showing the frequency distributions for two populations with means at -15 and 15 and a standard deviation equal to 20.

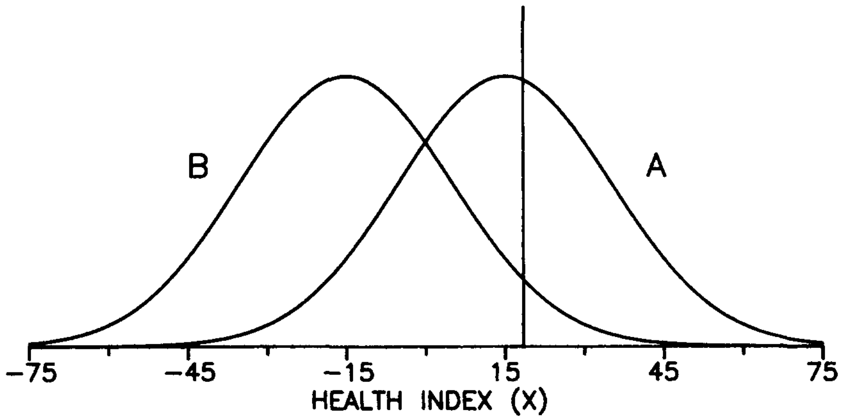

Figure 5. The distribution of health indices for patients from Categories A and B in Experiment 2. (Both distributions are normal with homogeneous variance but different means. The vertical line [x] indicates the location of the optimal cutoff for the payoff matrices used in this experiment.)

On each trial, n patients are sampled with equal frequency from two populations (n is called the *sample size*). The task is to categorize these n patients on the basis of their test scores. Each of the n patients is assigned to a diagnostic category using a single-cutoff rule. Any patient with a score above the cutoff ($X > \theta$) is diagnosed as a Population Type A patient. The cutoff criterion, θ, is illustrated in Figure 5 as the vertical line that intersects the horizontal axis at the score value of 18.

The payoffs are delivered according to a payoff matrix such as that shown in Table 2. A generic payoff matrix is shown on the left, and two special cases are shown in the middle and on the right. For example, if the cutoff rule puts a patient in category A', but that patient actually belongs to Population B, then the decision maker incurs a loss of 400 under matrix 1, 550 under Payoff Matrix 2, and b in general. The superscripts indicate the rank order of each cell according to the payoff value.

Optimal Model

The optimal decision rule for this task can be determined from Figure 6, which plots the expected payoff as a function of the cutoff parameter θ. Two curves are shown, one for each payoff matrix, and Matrix 2 produces a lower expected payoff. As can be seen in Figure 6, the maximum expected payoff is achieved when the criterion is set equal to 18.5. However, note that the mean payoff for any positive criterion produces almost equally good results. This is called the *flat maxima problem* (von Winterfeldt & Edwards, 1982). Hereinafter, we say that the criterion is in the correct region whenever it is positive.

For the generic payoff matrix, the optimal value of θ that maximizes the expected payoff is proportional to $\ln(b + d) - \ln(a + c)$. The quantity $(b + d)$ is the total loss that is incurred when the A' is incorrectly chosen. It includes both the penalty $-b$ that must be paid for the error on that trial and the loss of the opportunity to gain d if the error did not occur. Similarly, $(a + c)$ is the total loss that is incurred when B' is incorrectly chosen. Note that the total loss produced by each type of error is identical for Matrices 1 and 2, and consequently the optimal criterion is the same under both conditions.

Table 2
Payoff Matrices Used in Experiment 2 With Populations A and B

	Generic matrix		Matrix 1		Matrix 2	
Diagnosis	A	B	A	B	A	B
A′	a	−b	50^2	-400^4	100^1	-550^4
B′	−c	d	-100^3	200^1	-50^3	50^2

Note. Data indicate the payoffs corresponding to each combination of diagnostic decision category and disease population. The superscripts indicate the rank order of the cell value within a matrix.

Hill Climbing Versus Error Correction

The rule competition model assumes that subjects use a hill-climbing process to search for the best parameter of a rule. An alternative search model is the error-correction model. The two models make qualitatively different predictions regarding the effects of payoffs.

First, consider the error-correction model. The basic idea is that the new criterion equals the old criterion plus an adjustment, $\Delta\theta$. The adjustment is a product of two factors: the difference in the frequency of each type of error made on the previous trial and the total loss produced by each type of error.

Recall that a sample of n stimuli is categorized on each trial. Define $n_{A'}$ as the number of stimuli incorrectly assigned to category A′ on the previous trial and $n_{B'}$ as the number incorrectly assigned to category B′. For the unidimensional cutoff rule, the difference, $\Delta f = n_{A'} - n_{B'}$, determines the direction of adjustment. The total loss corresponding to the more frequent type of error determines the size of the adjust-

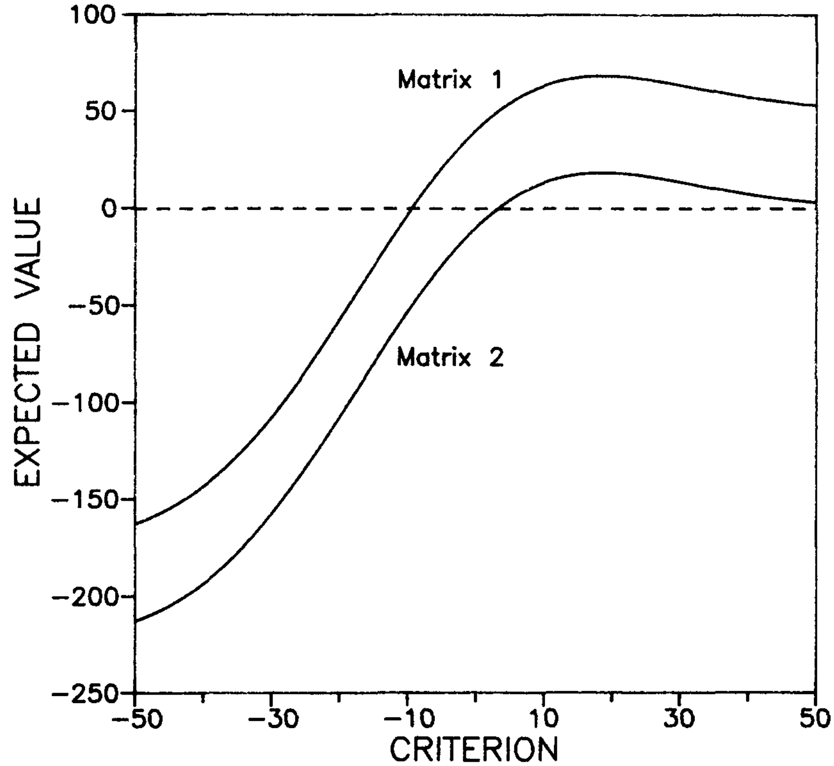

Figure 6. Expected payoff produced by using the cutoff rule plotted as a function of the criterion value separately for Payoff Matrices 1 and 2 for a sample size of 1.

ment. This leads to the following model for the adjustment:

$$\Delta\theta(t) = 0 \text{ if } \Delta f = 0,$$
$$= -m(a + c) \text{ if } \Delta f < 0,$$
$$= m(b + d) \text{ if } \Delta f > 0,$$

where m is a monotonically increasing function of the loss.[5]

For example, consider Matrix 1. If B′ was incorrectly chosen more frequently on the previous trial, then the amount of decrease depends on the total loss for that error, 50 + 100 = 150. If A′ was incorrectly chosen more frequently on the previous trial, then the amount of increase depends on the total loss for that error, 400 + 200 = 600. Because the latter change is larger, this search model will tend to move the criterion in the positive direction above zero.

One a priori prediction from the error-correction model is that there should be no differences produced by using Matrix 1 or Matrix 2 on criterion learning. This is because the total loss produced by each type of error is identical for the two matrices.

Now consider the predictions of the hill-climbing model with regard to the effects of payoffs. In contrast to the error-correction model, the hill-climbing model predicts that Matrix 1 will produce more rapid learning than Matrix 2. For Matrix 1, correctly choosing B′ produces the payoff with the highest rank, which tends to move the criterion in the positive direction. For Matrix 2, correctly choosing A′ produces the payoff with the highest rank, which tends to move the criterion in the negative direction.

Recall from Figure 6 that the criterion is in the correct region whenever it is positive. Table 3 shows the probability that the hill-climbing model (Equation 3) selects a criterion in the correct region for Payoff Matrices 1 and 2 as a function of training block and sample size.[6] For the small sample size condition ($n = 1$), the predicted rate of learning for Matrix 1 is greater than that for Matrix 2. In fact, there are almost no signs of learning for Matrix 2 for the small sample size. However, for the large sample size condition ($n = 15$), the difference in rate of learning between the two payoff matrices is eliminated. In summary, the hill-climbing model predicts a three-way interaction effect between payoff matrix, sample size, and training block.

[5] A slightly different error-correction model is obtained by assuming that n independent adjustments are made separately for each of the n category decisions made on a single trial. In this case, the sum of the n individual adjustments equals $\Delta\theta(t) = \alpha[n_{A'}\, m(b + d) - n_{B'}\, m(a + c)]$. However, this model predicts that 150 trials of training with sample size $n = 1$ is equivalent to 50 trials of training with sample size $n = 3$. The results of Experiment 2 indicate that this prediction is incorrect.

[6] The predictions were generated from Equation 3 by averaging over 1,500 simulations for each condition. The mean criterion for the first two trials was set to zero so that the probability of selecting a criterion in the correct region equaled .5 on the first two trials for all conditions. The effect of the random direction was eliminated by setting $\lambda = 1$. The S-shaped function $m(h)$ was set equal to the step function, and the step size was set at $c = 5$.

Table 3
Proportion of Trials That the Criterion Was in the Correct Region for Experiment 2

	Payoff matrix			
	1		2	
Training block	Predicted	Observed	Predicted	Observed
	$n = 1$			
1–10	.53	.44	.50	.52
11–50	.55	.60	.51	.40
51–150	.58	.63	.51	.39
	$n = 3$			
1–10	.54	.63	.53	.47
11–50	.64	.86	.57	.74
51–150	.73	.91	.61	.84
	$n = 15$			
1–10	.60	.63	.60	.61
11–50	.82	.92	.79	.85
51–150	.90	.87	.90	.94

Note. The correlation between the 18 predicted and 18 observed proportions equals .84.

The present experiment also permits a direct test of the two learning models on the basis of a sequential analysis of the trial events. The error-correction model predicts that the direction of the adjustment on any trial should be only a function of the difference in frequency of each type of error, Δf, from the preceding trial. In particular, the adjustment should be independent of the events from two or more trials back. In contrast, the hill-climbing model predicts that the direction of the criterion adjustment depends on the product of the change in payoffs and the change in criterion values from the two preceding trials (see Equation 3).

Method

Experimental tests of the error-correction and hill-climbing models were conducted using a version of the probabilistic categorization task called the *cutoff report technique*, developed by Kubovy and Healy (1977). It may be helpful to point out the differences between this task and the standard category response task. For the category response version, the trial begins by randomly sampling a stimulus, and the subject chooses the category (perhaps by using a cutoff, but this is not directly observed). For the cutoff report version, the process is reversed. The subject picks a criterion, which is observed, and the computer uses this criterion to categorize the stimulus. The main advantage of the cutoff report technique is that the trial-by-trial criterion adjustments are observable, which is crucial for the direct tests of the learning models.

Although the exact relation between the two tasks may be debatable (see Dorfman, 1977), experiments by Kubovy and Healy (1977) indicate that the two tasks produce indistinguishable results in terms of the estimated placement of the criterion. In any case, we are not claiming that the two tasks are equivalent but rather that the cutoff report task is worthy of study because it is an important decision-rule-learning problem that naturally arises in many applications, such as medical diagnosis and personnel selection.

Procedure. At the beginning of the experiment, subjects were asked to imagine that they were physicians charged with the task of assigning patients to diagnostic categories on the basis of clinical (central nervous system [CNS] activity) test scores. Half the subjects were told that Category A patients had a disease, and the scores for these patients were normally distributed, with $\mu = 15$ and $\sigma = 20$; Category B patients did not have the disease, and the scores for these patients were normally distributed, with $\mu = -15$ and $\sigma = 19$.[7] The other half of the subjects were given the same instructions except that the assignment of disease to categories was reversed. This information was also illustrated graphically by using a line interval with the 50th percentile located at the center, the 2nd percentile score typed at the left endpoint, and the 98th percentile score typed at the right endpoint. This graphical information was displayed on a sheet of paper during the entire experiment.

Subjects were given the following instructions concerning how to respond on each trial. At the beginning of each trial, they were asked to pick a criterion cutoff (an integer ranging from −50 to 50). All patients with scores greater than the cutoff were placed in Category A', and all patients with scores less than or equal to the cutoff were placed in Category B'. The subjects were told that the computer randomly sampled patients with equal frequency from each category and classified the patients according to the chosen criterion. After this they were given a feedback table including (a) the trial number, (b) the criterion selected on that trial, (c) the payoff matrix, (d) the number of patients correctly and incorrectly diagnosed for each category, (e) the net payoff (sum of gains and losses) for that trial, and (f) the total of the net payoffs summed across trials.

At the beginning of the first session, subjects were given 14 practice trials with two different symmetric payoff matrices. Then they were told that a new payoff matrix would be used during the first session. A 5-min break was given after completing the first session, which lasted 30 min. Subjects were told that a new payoff matrix would be used during the second session. All subjects were tested individually in a quiet room. An IBM-XT was used to present stimuli and collect responses.

Design. Two experimental factors were investigated, payoffs and sample size. The first factor was manipulated by paying subjects according to Payoff Matrix 1 or 2, which were described earlier. The second factor, sample size, refers to the number of patients (n) randomly sampled and categorized by the computer on each trial. Three sample sizes were used: $n = 1$, $n = 3$, and $n = 15$.

Each subject experienced only one sample size condition, and 12 subjects were randomly assigned to each of the three sample sizes ($n = 1$, $n = 3$, and $n = 15$). Each subject received two training sessions (150 trials per session), one session with Matrix 1 and one with Matrix 2. However, the direction of the bias produced by each payoff matrix was reversed across sessions by reversing the assignment of labels to rows and columns. For example, if A was assigned to the first row and column of Matrix 1 in the first session, then B would be assigned to the first row and column of Matrix 2 in the second session. This would cause the optimal criterion to be located at 18 in the first session and at −18 in the second session. Six subjects within each sample size condition received an 18 optimal criterion condition during the first session and a −18 optimal criterion condition during the second session, and the other 6 received the opposite order. Three subjects within each subgroup of 6 received Matrix 1 during the first session followed by Matrix 2 during the second session, and the other 3 received the opposite order. Altogether, this produced 3 (sample size) × 2 (order of direction of bias) × 2 (order of payoff matrix) = 12 groups with 3 subjects per group. For purposes of analyses, the sign of the observed criterion was multiplied by −1 for sessions on which

[7] The standard deviations used in Experiment 2 were not exactly homogeneous. This was done to encourage subjects to attend to both the mean and standard deviation of each distribution. This small difference in standard deviations is inconsequential, because the optimal criterion with unequal variances equals 18 rather than 18.5.

the optimal criterion was originally −18, making the correct region positive for all conditions.

Subjects. The subjects were 39 students from Purdue University (primarily humanities and social science majors) who volunteered for pay. The amount of pay was contingent on performance (each point was worth 25¢/n, and all subjects started the experiment with 10,000 points). On the average, subjects were paid approximately $6.00. The data from 3 subjects (1 from each sample size condition) were dropped because of subjects' failure to follow instructions.

Results

Proportion of trials in the correct region. Recall that the error-correction model predicted no effect of payoff matrices on the probability of selecting a criterion in the correct region, while the hill-climbing model predicted greater learning for Matrix 1 under the small ($n = 1$) sample size condition. Table 3 shows the observed proportion of trials that the criterion was in the correct region for each payoff matrix as a function of trial block and sample size.

First consider the results for the small ($n = 1$) sample size. For Matrix 1, the proportion increased substantially across training, but for Matrix 2, the trend is in the opposite direction. For the larger sample sizes, the proportion increased substantially at about the same rate for both matrices. These results are consistent with the predictions of the hill-climbing model, and they are inconsistent with the predictions of the error-correction model.

A 3 (sample size) × 2 (payoff matrix) × 3 (training block) analysis of variance with repeated measures on the last two factors produced a significant three-way interaction effect, $F(4, 66) = 2.88$, $MS_e = .031$, $p = .0293$. A two-way repeated measures analysis was conducted separately for each sample size. The training block by payoff matrix interaction effect was significant for the small sample size, $F(2, 22) = 5.52$, $MS_e = .0308$, $p = .0114$, but only the training block effect was significant for the larger sample sizes, $F(2, 22) = 21.57$, $MS_e = .032$, $p = .0001$, for $n = 3$; $F(2, 22) = 12.67$, $MS_e = .048$, $p = .0002$, for $n = 15$.

Sequential analyses. A direct test of the hill-climbing and error-correction models can be performed by the following sequential analysis of trial events. According to the product rule used to define h in Equation 3, the hill-climbing model predicts a crossover interaction effect of previous change in criterion and change in net payoff on the current change in criterion.

Column 6 of Table 4 presents the proportion of trials that subjects increased the criterion (given that a change occurred). These proportions were computed separately for each of 12 different conditions defined by the conjunction of three preceding events: (a) the sign of the difference between the frequency of errors for each category (column labeled Δf), (b) the sign of the change in criterion on the preceding trial (the column labeled $\Delta\theta$), and (c) the sign of the change in net payoffs on the preceding trial (column labeled Δv).

Referring to column 6 of Table 4 (Proportion of increase), the unique effect of the hill-climbing adjustment can be seen when Δf and $\Delta\theta$ are held fixed and Δv varies. Compare the first pair of rows within each set of four rows when the criterion was decreased on the previous trial; that is, compare the two rows with minus signs under $\Delta\theta$. The proportion of increases in criterion on the current trial is larger following a decrease in net payoff (minus signs under Δv) as compared with an increase in net payoff (plus signs under Δv). Next compare the second pair of rows within each set of four rows when the criterion was increased on the preceding trial; that is, compare the two rows with a plus sign under $\Delta\theta$. Now the proportion of increases in criterion on the current trial is larger following an increase in net payoff as compared with a decrease in the net payoff. In sum, there is a crossover interaction effect of previous change in criterion and change in net payoff consistent with the hill-climbing principle and contrary to the error-correction model.

Table 4
Sequential Analysis for Experiment 2

Prior event			Frequency of	Proportion of	Proportion of
Δf	Δθ	Δv	event[a]	change[b]	increase[c]
−	−	−	387	.91	.58
−	−	+	363	.71	.47
−	+	−	649	.92	.33
−	+	+	493	.77	.46
0	−	−	190	.92	.74
0	−	+	640	.63	.54
0	+	−	186	.91	.39
0	+	+	657	.68	.49
+	−	−	545	.93	.67
+	−	+	489	.74	.53
+	+	−	344	.89	.41
+	+	+	331	.66	.55

Note. Δf = sign of difference between the number of incorrect A′ and B′ decisions on the previous trial; Δθ = sign of the adjustment in criterion made on previous trial; Δv = sign of change in average payoffs from previous two trials.
[a] Number of trials. [b] Ignores the direction of change. [c] Conditioned on the occurrence of a change.

Discussion

The rule competition model is partly based on a general parameter search model called hill climbing. Past research on criterion learning in probabilistic categorization tasks has only considered the more specialized error-correction learning model. The present research indicates that this model is insufficient and that the adjustment on each trial is also influenced by a hill-climbing mechanism.

There are two lines of evidence supporting this conclusion. Indirect evidence was obtained from the effect of the payoff matrix manipulation—Matrix 2 was more difficult to learn than Matrix 1 for the small ($n = 1$) sample size condition. The error-correction model predicted that this manipulation would have no effect, because the total loss produced by each type of error is the same for both matrices. The hill-climbing model correctly predicted that Matrix 2 would be more difficult, especially for small sample sizes.

Direct evidence for the hill-climbing model was obtained from sequential analyses. As predicted by the hill-climbing model, there was a systematic effect of the change in average payoffs (Δv) from the previous two trials on the criterion adjustment, even when the difference in the frequency of errors (Δf) was held fixed. When A′ errors occur more frequently than B′ errors ($\Delta f > 0$), for example, the error-correction model always predicts an increase in the criterion (this is also true of the version described in Footnote 3); instead, the adjustment systematically decreased when the previous positive adjustment produced a decrease in net pay as predicted by the hill-climbing model.

Experiment 3

In the previous two experiments, we tested the hill-climbing part and the adaptive network part of the rule competition model separately. We designed Experiment 3 to test the combined rule competition model with a more complex decision task that requires using the hill-climbing model in conjunction with the adaptive network model.

The medical classification problem described in the general introduction was used to test the combined rule competition model. On each trial, hypothetical patients were randomly sampled from two populations, A and B. These patients were to be assigned to disease categories on the basis of a diagnostic test score X. Each incorrect category decision resulted in a loss of 25 monetary units, and subjects were instructed to minimize their losses across trials.

Patients were categorized by asking subjects to choose one of the following three categorization rules[8]:

G: If $X > \theta$, then report A; otherwise report B.

H: If $X > -\theta$ and $X < \theta$, then report A; otherwise report B.

I: If $X < -\theta$ or $X > \theta$, then report A; otherwise report B.

Subjects were also asked to select the criterion parameter, θ, corresponding to the chosen rule.

Two experimental factors were manipulated, sample size and stimulus distribution. The number of patients sampled on each trial (i.e., the sample size) was previously shown in Experiment 2 to have a major effect on rate of criterion learning. Therefore, small ($n = 3$) and large ($n = 15$) sample sizes were used in this experiment to manipulate the difficulty of the criterion learning part of the task.

The distribution of test scores was manipulated to vary the nature of the optimal categorization rule. The test scores from each population were normally distributed, but the means and variances were different for each population. Figure 7 illustrates three different pairs of distributions. For all three pairs of distributions in Figure 7, the test scores for Category B were normal with $\mu = 0$ and $\sigma = 20$. The test scores for Category A were also normal, but the mean and standard deviation varied across the three pairs in Figure 7.

Optimal Model

Figure 8 illustrates the expected payoff produced by each rule (as a function of the criterion) for each distribution. For the first pair of distributions (top panel of Figure 7), the test scores for Category A have $\mu = 30$ and $\sigma = 19$. The top panel of Figure 8 shows that Rule G is optimal, and the optimal criterion for Rule G is located at $\theta = 15$.

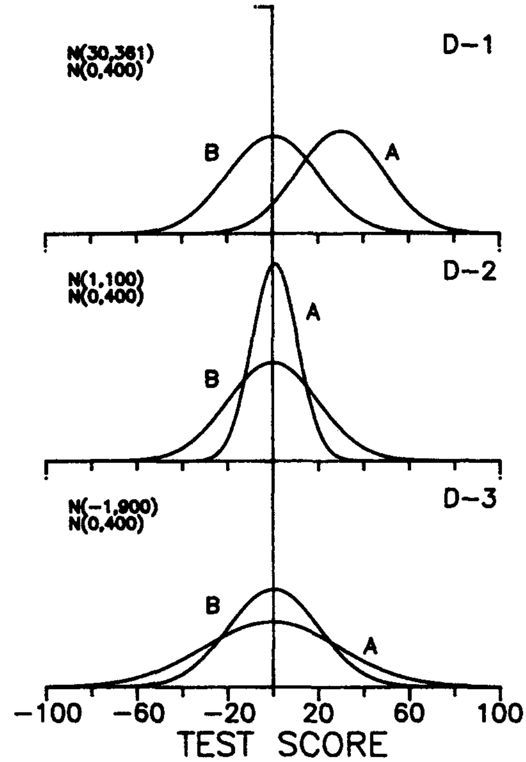

Figure 7. The distribution of health indices for patient samples from Categories A and B in Experiment 3. (Three pairs of distributions [D] are shown, and each pair of distributions is normal [N]. The means and variances are indicated in the upper left corner.)

For the second pair of distributions (middle panel of Figure 7), the test scores for Category A have $\mu = 1$ and $\sigma = 10$. The middle panel of Figure 8 shows that Rule H is optimal, and the optimal criterion for Rule H is located at $\theta = 14$.

For the third pair of distributions (bottom panel of Figure 7), the test scores for Category A have $\mu = -1$ and $\sigma = 30$. The bottom panel of Figure 8 shows that Rule I is optimal, and the optimal criterion for Rule I is located at $\theta = 24$.

Rule Competition Model Predictions

The rule competition model uses both the hill-climbing part and the adaptive network part to generate predictions for the probability of choosing each rule. A stringent test of the model was conducted using a new model-testing methodology. We did not fit the model to the data from Experiment 3. Instead, the learning rate parameters used in Equations 2 and 3 were fixed equal to the estimates obtained from Experiments 1 and 2. Using this method, we could make precise quantitative predictions prior to looking at the results from Experi-

[8] The technique of asking subjects to choose a rule from a set of rules has been used by previous decision researchers to investigate how decision rules are selected on the basis of accuracy–effort tradeoffs (e.g., Christensen-Szalanski, 1978). However, this previous work did not investigate decision rule learning, which is the focus of the present experiment.

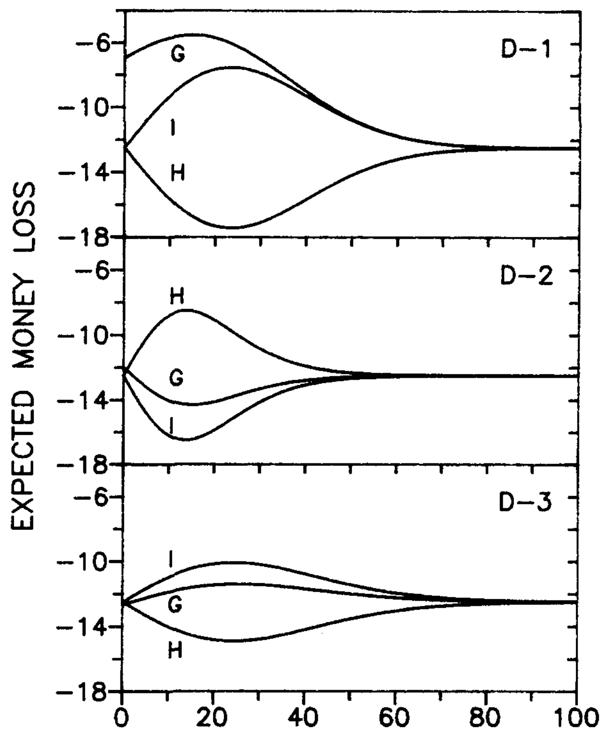

Figure 8. Expected payoff produced by a rule for a sample size of 1 for Experiment 3, plotted as a function of the criterion value with a separate curve for each rule (G, H, or I) and a separate panel for each distribution (D) condition.

ment 3. Table 5 shows the predicted probability of choosing each rule averaged across the last 20 of 100 training trials for each condition of the 3 (distribution) × 2 (sample size) factorial design.[9] On the basis of this table, we made the following predictions:

1. The optimal rule is predicted to be the modal choice under all conditions, but suboptimal rules will continue to compete with optimal rules throughout training. For Distribution 1, the optimal Rule G will compete primarily with Rule I. For Distribution 2, the optimal Rule H will compete with both Rules G and I. For Distribution 3, the optimal Rule I will compete primarily with Rule G.
2. Distribution 1 is predicted to produce the highest probability of choosing the optimal rule under the small sample size condition, but Distribution 2 is predicted to produce the highest probability for the large sample size. For both sample sizes, Distribution 3 is predicted to produce the lowest probability of choosing the optimal rule.
3. Increasing the sample size is predicted to increase the probability of choosing the optimal rule from an average of .57 for the small sample size to an average of .68 for the large sample size (averaged over distributions).

Table 5
Comparison of Predicted and Observed Rule Choices for Experiment 3

Stimulus distribution	Data source	Rule G	Rule H	Rule I
	$n = 3$			
1	Predicted	.63	.04	.33
1	Observed	.69	.08	.23
2	Predicted	.24	.58	.18
2	Observed	.23	.58	.19
3	Predicted	.37	.14	.49
3	Observed	.41	.17	.42
	$n = 15$			
1	Predicted	.66	.01	.33
1	Observed	.72	.04	.24
2	Predicted	.15	.74	.11
2	Observed	.07	.82	.11
3	Predicted	.30	.05	.65
3	Observed	.23	.12	.65

Note. Cells contain the average choice probabilities for the last 20 of 100 training trials. The *predicted* rows contain the predictions from the combined hill-climbing adaptive network model. The predictions were based on learning rate parameters estimated from Experiments 1 and 2, and no parameters were estimated from Experiment 3. The *observed* rows contain the data actually observed. The correlation between the 18 predicted and 18 observed proportions equals .97.

Method

Procedure. At the beginning of each session, subjects were provided complete information about the distribution of clinical test scores (CNS activity) for the two disease categories (A equals disease present, and B equals disease absent). They were told that both distributions were normal (bell shaped), and they were told the mean and standard deviation of each distribution. In addition, the subjects were shown a line interval indicating the 2nd, 50th, and 98th percentile scores for each distribution. This information was available during the entire session. Students familiar with signal-detection theory (e.g., engineers) could theoretically work out the optimal rule.

Subjects were told to imagine that they were physicians assigned the task of classifying fictitious patients into one of two categories on the basis of clinical test scores, using one of the three previously described rules (G, H, or I). On each trial, the subject first selected one of the three rules and then chose a criterion for the rule. Following this choice, the computer printed a verbal description of the rule with

[9] This note briefly describes the computer simulation used to generate the predictions for Experiment 3. Rule selection was based on Equations 1 and 2, using the same parameters as the mean of the estimates from Experiment 1. The similarity parameter, $s(t + 1)$, was set to one for the rule that was applied on the previous trial and zero otherwise. (We could have estimated the similarity between each pair of rules, which would have improved the fit of the model, but we decided not to do this to make parameter-free predictions.) The estimated performance of each rule on the first trial is unknown and so the initial estimate for each rule was set equal to the payoff expected on the basis of random guessing ($-12.5n$, where n = sample size). The performance estimates for all subsequent trials were determined by Equation 2. The initial value of the criterion was set to 50 (the midpoint of the scale) for the first trial and 40 for the second trial. All subsequent criteria were generated by Equation 3, using the same parameters as Experiment 2 (see Footnote 6). Five hundred simulated subjects were generated by the following algorithm. First, a rule was randomly sampled for each trial according to Equation 1. Second, a criterion value was generated for each rule, using Equation 3. If the criterion fell outside the interval [0, 100], a small random magnitude was added or subtracted to bring it back into the interval. Third, the computer randomly sampled n fictitious patients according to the distributions shown in Figure 7, and categorized the patients according to the rule and criterion chosen for the current trial. Finally, the performance estimates were updated according to Equation 2.

the chosen criterion. At this point, the computer randomly sampled a number of patients and categorized each patient using the rule and criterion chosen by the subject. Finally, the computer printed a feedback table indicating the trial number, rule, criterion, payoff matrix, number of patients correctly and incorrectly classified separately for each category, amount lost on that trial, and total amount of money remaining. (Each subject started with an initial sum equal to 6,700 and 49,000 points for the small and large sample size groups, respectively.)

Subjects were initially given nine practice trials with the following pair of distributions. Category B had a $\mu = 0$ and $\sigma = 20$, whereas Category A had a $\mu = -15$ and $\sigma = 40$. Following practice, they were given the three main sessions with a 5-min break between sessions. Subjects were tested individually in a quiet room. An IBM-XT was used to present stimuli and record responses.

Design. Each subject experienced all three distribution conditions (shown in Figure 7) during three sessions of training with 100 trials per session. One of the three distribution conditions was assigned to each session, and the presentation order was counterbalanced across subjects. Each subject received one of the two ($n = 3$ or $n = 15$) sample sizes during all three sessions.

Subjects. A total of 68 student volunteers from Purdue University participated. All participants were required to have passed an undergraduate statistics course. All subjects were paid contingent on performance (each point was worth .10¢ for the small sample size group and .01¢ for the large sample size group). On the average, subjects earned $10 for the experiment.

Large individual differences were expected on the basis of a student's mathematical background, and so the 68 subjects were divided into two groups—one group of 20 engineering/mathematics students and another group of 48 humanities/social science students. The engineering/mathematics subjects were expected to be able to identify the optimal rule at the very beginning of training, but we did not wish to prevent them from participating for ethical reasons. Eight humanities/social science students were randomly assigned to each of the 2 (sample size) × 3 (presentation order), or a total of six, experimental conditions.

Results

Because of large educational differences between the engineering/mathematics versus humanities/social science groups, separate analyses were conducted on each group. Essentially, the engineering/mathematics group identified the optimal rule within the first few trials and consistently used it throughout training. This result can be interpreted in terms of the present theory as follows: The performance estimate for the optimal rule was relatively large at the very beginning of training for this group because of their prior education.

The results for the humanities/social science group are presented in Table 5 along with the predictions from the rule competition model. Recall that optimal performance is achieved by always choosing Rule G for Distribution 1, Rule H for Distribution 2, and Rule I for Distribution 3.

In agreement with our first prediction, the optimal rule was the modal choice under all conditions, but suboptimal rules continued to compete with optimal rules throughout training. For Distribution 1, the Optimal Rule G competed primarily with Rule I. For Distribution 2, the Optimal Rule H competed with both Rules G and I. For Distribution 3, the Optimal Rule I competed primarily with Rule G.

In agreement with our second prediction, Distribution 1 produced the highest probability of choosing the optimal rule for the small sample size condition, but Distribution 2 produced the highest probability for the large sample size. For both sample sizes, Distribution 3 produced the lowest probability of choosing the optimal rule.

Increasing the sample size increased the probability of choosing the optimal rule from an average of .57 for the small sample size (which is exactly what we predicted) to an average of .73 for the large sample size (which is only 5% higher than our prediction of .68).

The quantitative agreement between the predicted and observed choice probabilities is excellent. The correlation between the 18 predicted and observed choice proportions equals .97. We consider this excellent because no parameters were estimated from the data of Experiment 3.

One might ask how important the hill-climbing part of the rule competition model is for predicting rule choice. To answer this question, we recalculated the predictions with the hill-climbing part of the model "turned off." This resulted in a significant reduction in the squared correlation between the predicted and observed choice proportions: .94 with hill-climbing "turned on," and .54 with hill-climbing turned off, that is, there was a 40% reduction in the percentage of predicted variance. Essentially, the rate of rule learning was substantially reduced by turning off the hill-climbing criterion search. For example, the probability of choosing the optimal rule for Distribution 2 dropped from .58 to .38 under the small sample size, and it dropped from .74 to .40 under the large sample size.

We calculated the statistical significance of the effects of distribution and sample size on proportion of optimal rule choice by calculating a 3 (distribution) × 2 (sample size) × 10 (trial block) analysis of variance, using the proportion of optimal rule choices within each block of 10 trials for each subject as the dependent variable. The following main effects were significant: distribution, $F(2, 88) = 5.69$, $p < .005$; sample size, $F(1, 44) = 5.09$, $p < .03$; trial block, $F(9, 396) = 20.47$, $p < .0001$; and the Sample Size by Trial Block interaction effect, $F(9, 396) = 2.47$, $p < .01$. No other effects were significant at the .05 level.

Discussion

The purpose of Experiment 3 was to test the joint operation of both the hill-climbing and adaptive network parts of the rule competition model in a task that required subjects to learn both rules and criteria. Quantitative, parameter-free predictions were generated from the rule competition model by using the parameters estimated from Experiments 1 and 2 to generate the predictions for Experiment 3. The high correlation (.97) between the predicted and observed probabilities of choosing each rule under each stimulus distribution and sample size condition provides convincing support for the model.

The role of the hill-climbing part of the rule competition model was evaluated by comparing predictions with this part of the model either turned on or turned off. Turning the hill-climbing part of the model off eliminates all criterion learning, and this had the effect of drastically reducing the rate of rule learning. The percentage of variance predicted by the rule competition model was reduced by 40% when the hill-climbing part of the model was turned off. In conclusion, both parts of the model (the hill-climbing model and the adaptive network model) are needed to yield accurate predictions for rule learning.

General Discussion

Extensive training with veridical feedback and monetary incentives does not guarantee that something close to an optimal decision rule eventually will be learned. In fact, the learning process does not seem to converge toward any single rule, but instead it continues to explore a range of feasible rules. Suboptimal performance cannot be explained simply in terms of flat maxima (von Winterfeldt & Edwards, 1982), because this fails to specify how steep the objective function must be and how much training is needed to learn an optimal strategy. A satisfactory answer to the question of what is learned and how fast requires a specific model of the learning process. We proposed and tested a model of decision rule learning called the rule competition model.

The rule competition model consists of two interactive learning processes: (a) an adaptive network model that learns the payoff probabilities produced by each rule and (b) a parameter search model that learns the parameters for each rule. We designed the first experiment to compare the predictions from three different probability learning models, and the results supported the adaptive network model over a frequency array model and a multiple-trace model. We designed the second experiment to compare the predictions from two different parameter search models, hill climbing versus error correction, and the results supported the hill-climbing model, although a mixture of the two search models remains feasible. In the third experiment, we used a task that required both rule and parameter learning, and the rule competition model provided accurate parameter-free predictions for the effects of stimulus distribution and sample size on final performance. In conclusion, these results indicate that the rule competition model provides a simple yet accurate description of the evolution of decision rules on the basis of experience with outcome feedback.

Multidimensional Spaces

In the present experiments, we focused on the learning of decision rules for unidimensional stimuli that involve a single criterion only. However, the hill-climbing model can be used with multidimensional stimuli and rules that involve more than one parameter. For example, the linear categorization rule investigated by Ashby and Gott (1988) requires the decision maker to learn two parameters, the slope and the intercept of a line that divides the stimulus plane into halves. Our Equation 3 can be applied directly to this learning problem by defining θ as a two-dimensional parameter vector containing the two parameters of the linear rule.

In fact, the hill-climbing model was initially designed for multidimensional stimuli (see Busemeyer, Swenson, & Lazarte, 1986) and earlier applications of the hill-climbing learning model have already demonstrated that the model can be used successfully to predict performance from multidimensional learning experiments (see Busemeyer & Myung, 1987; Busemeyer et al., 1986; Nosofsky & Gluck, 1989). We used unidimensional stimuli in Experiment 2 to provide simpler and more direct qualitative tests of the hill-climbing model in comparison with the error-correction model (the latter was developed only for unidimensional stimuli).

Extensions to Other Decision Tasks

So far, we have focused on the problem of learning categorization rules for probabilistic categorization tasks. However, the rule competition model can be applied to any decision task that involves learning to select one rule from a set of mutually exclusive decision rules.

For example, consider a generalization of the probabilistic categorization task called the *deferred decision task* (Busemeyer & Rapoport, 1988; Myung & Busemeyer, 1989). In this task, the decision maker is asked to make a diagnosis on the basis of a sequence of test results. After purchasing and observing each test result, the decision maker can either stop and make a terminal diagnosis or continue sampling by purchasing another test. Two different stopping rules are possible—a critical-difference rule and a counter-race rule. According to the critical-difference rule, the decision maker accumulates the difference in evidence favoring each diagnosis and stops as soon as this difference exceeds a threshold. In the counter-race rule, the decision maker sums the evidence for each diagnosis separately in two counters and stops as soon as one of the counters exceeds a criterion.

The rule competition model could be used to describe how individuals learn to prefer one stopping rule over another. The adaptive network model could be used to learn the estimated performance for each rule, and the parameter search model could be used to search for the best criteria.

More generally, the rule competition model could also be applied to other decision tasks such as multiattribute decisions (e.g., Payne, Bettman, & Johnson, 1988), or information-processing tasks such as attention (Navon & Gopher, 1979; Nosofsky, 1987; Sperling & Dosher, 1986) and motor performance (Meyer, Abrams, Kornblum, Wright, & Smith, 1988). Finally, the rule competition model may be applicable to rule learning in complex hypothesis testing or problem-solving domains with graded and unreliable feedback, and the hill-climbing model may be used to learn connectivity weights in adaptive network models for general objective functions. In conclusion, the rule competition model may provide a very general framework for understanding how humans learn complex rules on the basis of graded and unreliable feedback.

References

Anderson, J. A. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 33–53.

Atkinson, R. C., & Kinchla, R. A. (1965). A learning model for forced-choice detection experiments. *British Journal of Mathematical and Statistical Psychology, 18,* 183–206.

Bernbach, H. A. (1967). Decision processes in memory. *Psychological Review, 74,* 462–480.

Bock, R. D., & Jones, L. V. (1968). *The measurement and prediction of judgment and choice.* San Francisco: Holden Day.

Busemeyer, J. R. (1985). Decision making under uncertainty: A comparison of simple scalability, fixed sample, and sequential sampling models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 538–564.

Busemeyer, J. R., & Myung, I. (1987). Resource allocation decision making in an uncertain environment. *Acta Psychologica, 66,* 1–19.

Busemeyer, J. R., & Rapoport, A. (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology, 32,* 91–134.

Busemeyer, J. R., Swenson, K. N., & Lazarte, A. (1986). An adaptive approach to resource allocation. *Organizational Behavior and Human Decision Processes, 38,* 318–341.

Christensen-Szalanski, J. J. J. (1978). A mechanism for strategy selection and some implications. *Organizational Behavior and Human Decision Processes, 22,* 307–323.

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2nd ed). Urbana: University of Illinois Press.

Dorfman, D. D. (1977). Comments on "The decision rule in probabilistic categorization: What it is and how it is learned," by Kubovy and Healy. *Journal of Experimental Psychology: General, 106,* 447–449.

Dorfman, D. D., & Biderman, M. (1971). A learning model for a continuum of sensory states. *Journal of Mathematical Psychology, 8,* 264–284.

Dusoir, A. E. (1980). Some evidence on additive learning models. *Perception and Psychophysics, 27,* 163–175.

Estes, W. K. (1987). Application of a cognitive-distance model to learning in a simulated travel task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 380–386.

Estes, W. K., Campbell, J. A., Hatsopoulos, N., Hurwitz, J. B. (1989). Base-rate effects in category learning: A comparison of parallel network and memory storage-retrieval models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 556–571.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117,* 227–247.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Healy, R. A., & Kubovy, M. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 7,* 344–354.

Heath, R. A., & Fulham, R. (1988). An adaptive filter model for recognition memory. *British Journal of Mathematical and Statistical Psychology, 41,* 119–144.

Kac, M. (1962). A note on learning signal detection. *IRE Transactions on Information Theory, IT-8,* 126–128.

Kleyle, R., & De Korvin, A. (1988). Martingale properties of an information feedback loop. *Mathematical Computer Modelling, 10,* 1–11.

Knapp, A. G., & Anderson, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 616–637.

Kubovy, M., & Healy, A. F. (1977). The decision rule in probabilistic categorization: What it is and how it is learned. *Journal of Experimental Psychology: General, 106,* 427–446.

Kubovy, M., and Healy, A. F. (1980). Process models of probabilistic categorization. In T. S. Wallsten (Ed.), *Cognitive processes in choice and decision behavior* (Ch. 13, pp. 239–262). Hillsdale, NJ: Erlbaum.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 14,* 159–188.

Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Smith, J. E. K. (1988). Optimality in human performance: Ideal control of rapid aimed movements. *Psychological Review, 95,* 340–370.

Minsky, M., & Papert, S. (1969). *Perceptrons.* Cambridge, MA: MIT Press.

Myung, I. (1990). *Adaptive learning strategies in a time-series prediction task.* Unpublished doctoral dissertation, Purdue University, West Lafayette, Indiana.

Myung, I. J., & Busemeyer, J. R. (1989). Criterion learning in a deferred decision making task. *American Journal of Psychology, 102,* 1–16.

Myung, I. J., & Busemeyer, J. R. (1992). Measurement-free tests of a general state space model of prototype learning. *Journal of Mathematical Psychology 36,* 32–67.

Navon, D., & Gopher, D. (1979). On the economy of the human information processing system. *Psychological Review, 86,* 214–255.

Norman, M. F. (1972). *Markov processes and learning models.* San Diego, CA: Academic Press.

Nosofsky, R. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 87–109.

Nosofsky, R. M., & Gluck, M. A. (1989). *Adaptive networks, exemplars, and classification learning.* Paper presented at the Thirtieth Annual Meeting of the Psychonomic Society, Atlanta, GA.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 534–552.

Peterson, W. W., Birdsall, T. G., & Fox, W. C. (1954). The theory of signal detectability. *Transactions IRE Professional Group on Information Theory (PGIT-4,* 171–212).

Sperling, G., & Dosher, B. A. (1986). Strategy and optimization in human information processing. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 1–61). New York: Wiley.

Sternberg, S. (1963). Stochastic learning theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 1–120). New York: Wiley.

Swets, J. A., & Pickett, R. M. (1982). *Effective prescriptions for evaluating diagnostic performance.* San Diego, CA: Academic Press.

Swets, J. A., & Sewall, S. T. (1963). Invariance of signal detectability over stages of practice and levels of motivation. *Journal of Experimental Psychology, 66,* 120–126.

Thomas, E. A. C. (1973). On a class of additive learning models: Error correcting and probability matching. *Journal of Mathematical Psychology, 10,* 241–264.

von Winterfeldt, D., & Edwards, W. (1982). Costs and payoffs in perceptual research. *Psychological Bulletin, 19,* 609–622.

Wickelgren, W. A., & Norman, D. A. (1966). Strength models and serial position in short term recognition memory. *Journal of Mathematical Psychology, 3,* 316–347.

(Appendix follows on next page)

Appendix
Mathematical Details of the Adaptive Network Model

$\mathbf{S}(t)$ is defined as a $p \times 1$ situation vector used to represent information about the experimental condition (payoffs, prior probabilities, and stimulus distributions) present on trial t (e.g., $p = 4$ in Figure 1). $\mathbf{R}(t)$ is a $q \times 1$ rule vector representing information about the general rule (e.g., cutoff rule) that was used to make the category decision on trial t (e.g., $q = 5$ in Figure 1). The outer product $\mathbf{S}(t)\mathbf{R}(t)'$ forms a $p \times q$ matrix that can be stretched out to form a $(p \cdot q) \times 1$ input vector, $\mathbf{Z}(t)$, which is the input to the top layer of Figure 1. Binary valued features were used in Figure 1 for simplicity, but generally, a continuous range of feature values can be used to represent different degrees of activation of each feature in the vectors $\mathbf{S}$ and $\mathbf{R}$.

$\mathbf{V}(t)$ is an $r \times 1$ payoff vector representing the actual payoff presented on trial t. Each element of $\mathbf{V}$ represents a different payoff level (e.g., low, medium, and high). $\mathbf{P}(t)$ is an $r \times 1$ probability vector representing the decision maker's estimated probability of each payoff level, and it is the output from the middle layer of the network in Figure 1 (e.g., $r = 3$ in Figure 1). $\mathbf{W}(t)$ is the $(p \cdot q) \times r$ connection weight matrix that connects the input $\mathbf{Z}(t)$ to the output $\mathbf{P}(t)$. $\mathbf{P}(t)$ is obtained from the matrix product:

$$\mathbf{P}(t) = \mathbf{W}'(t)\mathbf{Z}(t). \tag{A1}$$

The weight matrix is updated according to the Hebb-delta rule (Heath & Fulham, 1988; Myung & Busemeyer, 1992):

$$\mathbf{W}(t+1) = \beta \cdot \mathbf{W}(t) + \alpha \cdot \mathbf{Z}(t)[\mathbf{V}(t) - \gamma \cdot \mathbf{P}(t)]'. \tag{A2}$$

Finally, let $\mathbf{u}$ be an $r \times 1$ worth vector that connects each payoff level to the final performance estimate. The estimated performance of a rule being considered on trial t is the inner product:

$$U(t) = \mathbf{u}'\mathbf{P}(t). \tag{A3}$$

Now suppose a rule, Rule X, was applied on trial t, and a new rule, Rule Y, is being considered for trial $t + 1$. In this case, let $\mathbf{Z}(t) = \mathbf{Z}_X$ be the input vector for Rule X used on trial t, and let $\mathbf{Z}(t+1) = \mathbf{Z}_Y$ be the input vector for Rule Y used on trial $t + 1$. Then,

$$\begin{aligned} U(t+1) &= \mathbf{u}'\mathbf{W}'(t+1)\mathbf{Z}(t+1), \\ &= \mathbf{u}'\{\beta \cdot \mathbf{W}'(t) + \alpha \cdot [\mathbf{V}(t) - \gamma \cdot \mathbf{P}(t)]\mathbf{Z}'(t)\}\mathbf{Z}(t+1), \\ &= \beta \cdot \mathbf{u}'\mathbf{W}'(t)\mathbf{Z}(t+1) \\ &\quad + \alpha \cdot [\mathbf{u}'\mathbf{V}(t) - \gamma \cdot \mathbf{u}'\mathbf{P}(t)] \cdot \mathbf{Z}'(t)\mathbf{Z}(t+1), \end{aligned}$$

and substituting $\mathbf{Z}(t) = \mathbf{Z}_X$ and $\mathbf{Z}(t+1) = \mathbf{Z}_Y$,

$$\begin{aligned} U_Y(t+1) &= \beta \cdot \mathbf{u}'\mathbf{W}(t)\mathbf{Z}_Y + \alpha \cdot [\mathbf{u}'\mathbf{V}(t) - \gamma \cdot \mathbf{u}'\mathbf{W}'(t)\mathbf{Z}_X] \cdot \mathbf{Z}_X'\mathbf{Z}_Y, \\ &= \beta \cdot U_Y(t) + \alpha \cdot s(t+1) \cdot [v(t) - \gamma \cdot U_X(t)], \end{aligned} \tag{A4}$$

where the scalars, $U_X(t) = \mathbf{u}'\mathbf{W}'(t)\mathbf{Z}_X$ and $U_Y(t) = \mathbf{u}'\mathbf{W}'(t)\mathbf{Z}_Y$, are the performance estimates of rules X and Y after t feedback trials, the scalar $v(t) = \mathbf{u}'\mathbf{V}(t)$ is simply the worth of the payoff presented on trial t, and the scalar $s(t+1) = \mathbf{Z}'(t)\mathbf{Z}(t+1) = \mathbf{Z}_X'\mathbf{Z}_Y$ is the inner product between the input vectors for Rule X from trial t and Rule Y from $t + 1$. Using Equation A4, the feature values do not need to be specified—only the similarity (inner product) needs to be specified. [Note that $s(t + 1) = 0$ for orthogonal vectors.]

Without loss in generality, we can assume that the input vectors are scaled to have a maximum length equal to 1, that is, $\mathbf{Z}'\mathbf{Z} = 1$. The new performance estimate for the rule used on the previous trial is obtained by setting $\mathbf{Z}(t+1) = \mathbf{Z}(t)$ and $s(t+1) = 1$, and in this case, Equation A4 reduces to

$$\begin{aligned} U(t+1) &= \beta \cdot U(t) + \alpha \cdot [v(t) - \gamma \cdot U(t)] \\ &= (\beta - \alpha \cdot \gamma)U(t) + \alpha \cdot v(t) \\ &= \eta \cdot U(t) + \alpha \cdot v(t), \end{aligned} \tag{A5}$$

where $\eta = (\beta - \alpha \cdot \gamma)$. If the same rule is applied repeatedly under a fixed experimental condition, then for $t > 0$,

$$U(t+1) = \eta^t U(1) + \alpha \cdot \sum \eta^{(t-k)} v(k), \quad \text{for} \quad k = 1, \ldots, t. \tag{A6}$$

Received February 13, 1991
Revision received September 3, 1991
Accepted September 19, 1991 ■