

THEORETICAL AND REVIEW ARTICLES

Comparison of basic assumptions embedded in learning models for experience-based decision making

ELDAD YECHIAM and JEROME R. BUSEMEYER
Indiana University, Bloomington, Indiana

The present study examined basic assumptions embedded in learning models for predicting behavior in decisions based on experience. In such decisions, the probabilities and payoffs are initially unknown and are learned from repeated choice with payoff feedback. We examined combinations of two rules for updating past experience with new payoff feedback and of two choice rule assumptions for mapping experience onto choices. The combination of these assumptions produced four classes of models that were systematically compared. Two methods were employed to evaluate the success of learning models for approximating players' choices: One was based on estimating parameters from each person's data to maximize the prediction of choices one step ahead, conditioned by the observed past history of feedback. The second was based on making a priori predictions for the entire sequence of choices using parameters estimated from a separate experiment. The results indicated the advantage of a class of models incorporating decay of previous experience, whereas the ranking of choice rules depended on the evaluation method used.

Recently, interest has been rising in learning models that are applied to choices from repeated-play games. Recent studies of choice behavior in individual (see, e.g., Busemeyer & Myung, 1992; Erev & Barron, in press; Sarin & Vahid, 1999) and multiplayer (see, e.g., Camerer & Ho, 1999b; Cheung & Friedman, 1997; Erev & Rapoport, 1998; Erev & Roth, 1998; Fudenberg & Levine, 1995; Sarin & Vahid, 2001; Stahl, 1996) games have shown that learning in repeated-choice problems can be summarized by using surprisingly simple mathematical models. The purpose of this article is to provide a systematic comparison of the basic assumptions used to construct models for decision making based on experience from repeated play.

We evaluated four classes of models that were formed by combining two basic assumptions about learning rules with two basic assumptions about choice rules (see Table 1). The learning rules in these models differed according to the manner in which past experience was updated on the basis of new feedback: In one class, called the *interference models*, only the chosen option was updated, and unchosen options remained unchanged; in the

other, called the *decay models*, the chosen option was updated and the unchosen options were discounted by some amount. The choice rules differed according to the manner in which past experience was mapped onto choice behavior: In one class, the option producing the maximum expectation was always chosen (with some guessing allowed); in the other, choices were probabilistically determined by the strength of expectation.

Two methods were used to compare the empirical validity of the models. The first was based on "one-step-ahead" predictions, and the second on simulation of the entire game. Under the first method, the model predicted the player's next choice ahead using the past history of payoffs actually experienced by a player. In this case, the model parameters were estimated separately for each player to maximize the likelihood of the observed choices. Using the second method, model simulations were generated to predict the proportion of choices, averaged across players, for the entire length of the game. In this case, parameters estimated from one experiment were used to generate a priori predictions for a second experiment. The advantage of the first method was that it allowed tests of the model at the individual level, but it also had two disadvantages: (1) It had to rely on actual past choices of an individual, and (2) it relied on fitting parameters to the data. The second method circumvented both of these disadvantages, but it could not be used to test models at the individual level. By using both methods, we hoped to achieve a convergence of evidence.

This research was supported in part by Grant DA R01 014119 from the National Institute on Drug Abuse and by Shared University Research Grants from IBM to Indiana University. The authors thank Richard Shiffrin for his comments. Correspondence relating to this article may be sent to E. Yechiam, Behavioral Science Area, Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel (e-mail: yeldad@tx.technion.ac.il).

Table 1
General Design for Examining Basic Assumptions
of Competing Models

	Maximizing Choice Rules	Probabilistic Choice Rules
Interference Learning Models		
Decay Learning Models		

A four-alternative choice task was chosen for model comparison rather than a simple two-alternative (binary) task because it allowed us to explore the main difference between interference and decay models. In interference models, only experience related to the selected alternative is updated following a choice. Thus, whereas in a binary task an update in the experience of the chosen alternative implies a mirroring update in the unchosen alternative, this is not so in a multiple-alternative task. The particular task we selected was a highly studied problem called the Iowa gambling task (Bechara, A. R. Damasio, H. Damasio, & Anderson, 1994). This gambling task has been extensively studied in its association with decision-making deficits of individuals with brain damage (for a review, see Busemeyer, Stout, & Finn, in press; see also Bechara & H. Damasio, 2002; Clark & Robbins, 2002).

The present article begins by presenting the Iowa gambling task and summarizing the basic empirical results. Second, the models to be compared are described. Rather than contrasting specific models, we will compare basic assumptions made by these models about the updating of expectancies and examine different choice rules. Third, different methods for evaluating learning models are described—one method is based on predicting the choices of individuals in a one-step manner, the other on simulating the choice probabilities of the group average for the entire sequence of choice trials. Finally, the models are empirically evaluated on the basis of data (Yechiam, Stout, Busemeyer, Rock, & Finn, 2005) collected under different payoff conditions of the Iowa gambling task.

EXPERIMENT AND BASIC FINDINGS

In the Iowa gambling task, participants are presented with four card decks, labeled A, B, C, and D. They are told to accumulate as much (real) money as possible by picking cards from the decks. Initially, they know nothing about the payoffs produced by each deck, and this has to be learned from trial-by-trial feedback. Decks differ with respect to the payoff for each card selection and

Table 2
Choice Alternatives in the Iowa Gambling Task

Deck	Wins	Losses	Net
A	\$1 every card	.5 to lose \$2.50	Loss
B	\$1 every card	.1 to lose \$12.50	Loss
C	\$0.50 every card	.5 to lose \$0.50	Gain
D	\$0.50 every card	.1 to lose \$2.50	Gain

the frequency and severity of penalties, as indicated in Table 2.

Note that Decks C and D may appear to be disadvantageous when considering the gain domain. However, in each trial the winnings are also paired with losses on many cards in such a way that Decks A and B are disadvantageous overall, leading to an average expected loss of 25 cents per trial, whereas Decks C and D are advantageous overall, leading to an average expected gain of 25 cents per trial.

Yechiam et al. (2005) examined the choices of young adults, mostly (85%) college students, in the Iowa gambling task.¹ Two primary conditions were manipulated in their study: In the first, the standard “partial-information” version of the task was used; in this version, only the payoff for the chosen gamble was shown on each trial. The second condition featured a modified “full-information” version of the task; in this version, the payoffs from all four gambles were shown on each trial (although earnings were based solely on the chosen option).

Yechiam et al. (2005) also performed a secondary manipulation involving the size of the payoffs. In the low-payoff condition, disadvantageous decks (A and B) had the exact wins/losses indicated in Table 2. In the high-payoff condition, these payoffs (wins and losses) were multiplied by a constant factor of 1.5. This manipulation had little effect, and results from the two versions were pooled in the study. In the present analysis, the two payoff magnitude conditions were utilized to provide a cross-validation test of the models (as described in the Simulation Method section below).

Method

Participants. The original study included 162 young men and women. The participants were 22 years old on average and had 14 years of education. Because of a technical problem, the trial-by-trial data of 7 individuals were lost, permitting the modeling of the behavior of 76 participants in the partial-information condition and 79 participants in the full-information condition.

Apparatus. The experiment used a computer-simulated version of the gambling task developed by Bechara et al. (1994). Images of four card decks labeled A, B, C, and D were displayed horizontally and “face down” on a monitor controlled by a desktop computer. The participants were instructed to make a series of selections from the decks using the mouse and to try to win as much money as possible. They received a \$20.00 credit at the start of the task and were informed that their winnings would be paid at the end of the session as long as they continued until the game was completed (150 trials). Gains and losses were shown on two tally bars at the top of the display, the top one revealing the cumulative net win/loss and the bar below it indicating the win/loss for the most recent selection (see Figure 1).

Results

The results in the partial-information condition (see Figure 2) show a slow but significant effect of learning, with fewer choices from disadvantageous decks (A and B) and more from advantageous decks (C and D) as a function of time (first versus last block of 25 trials).² Note that these results deviate from Bechara et al.’s (1994) original results, in which control participants

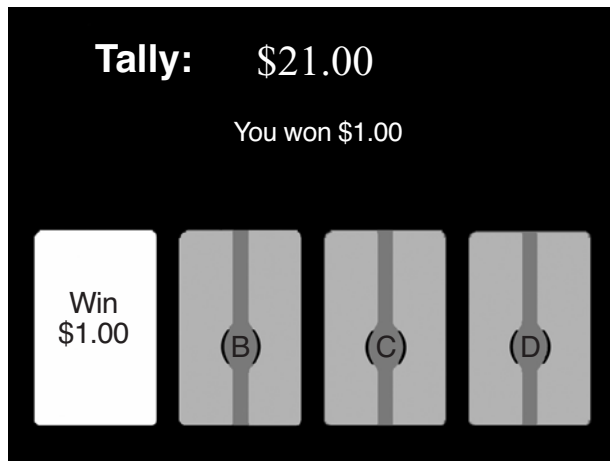


Figure 1. A screen shot from the simulated Iowa gambling task.

(who had lesions in the left somatosensory area and were about 20 years older than our participants on average) learned to choose Decks C and D consistently over Decks A and B. However, note that the present task did not include constraints on the order of losses as Bechara et al.'s did, and thus it might have been more difficult in our study to learn the contingencies. Also, the characteristics of the populations were different.

In contrast to most follow-ups of Bechara et al. (1994), Yechiam et al. (2005) focused on choices of specific decks (rather than grouping Decks C and D together as “advantageous”). The results revealed that for the partial-information condition, Decks B and D were chosen, on average, more often than Decks A and C [.62 vs. .38; $t(75) = 6.12, p < .01$]. The low-frequency decks (B and D) with respect to negative payoffs ($p = .1$ for a negative payoff) were chosen more often on average, despite having expected values equal to those of the high-frequency decks (A and C, respectively). This result was observed for both the advantageous and disadvantageous decks (B chosen more than A, D chosen more than C); it replicates previous findings in simpler binary tasks (Bar-

ron & Erev, 2003; Erev & Barron, in press; Hertwig, Barron, Weber, & Erev, 2004) and signal detection tasks (Barkan, Zohar, & Erev, 1998), showing that decision makers underweight small-probability events with experience.

A comparison of the partial- and full-information conditions (see Figure 2) reveals the following important differences: In the full-information condition, on average, more choices were made from decks with a low frequency of negative payoffs [B and D, .70 with full information vs. .62 with partial information; $t(153) = 3.07, p < .01$] and fewer choices from advantageous Deck C [.15 vs. .24; $t(153) = 3.95, p < .01$]. Furthermore, for Deck C, choices decreased over time (although the decrease was largely limited to the first 50 trials). This finding is interesting, since it indicates that more information led to poorer performance in the task (and a \$2 decrease in earnings from \$23 to \$21). It also replicates previous results indicating that forgone payoffs can, under certain conditions, increase risk seeking (see, e.g., Yechiam & Busemeyer, 2005). Thus, the results replicate some of the robust regularities in the performing of experience-based choice tasks. The next step was to systematically compare mathematical learning models that can capture these regularities.

EXPERIENCE-BASED DECISION-MAKING MODELS

An examination of the learning models used by previous researchers reveals that most models employ three groups of assumptions: First, a utility function is used to represent the evaluation of the payoff experienced immediately after each choice. Second, a learning rule is used to form an expectancy (or propensity) for each choice alternative that summarizes the experience of all of the past utilities produced by each choice alternative. Third, a choice rule is based on the comparison of the expectancies formed for each choice option. Various learning theories posit one set of assumptions about the process for updating expectancies and another set about the choice rule. We first present the formulation of the models for the partial-information

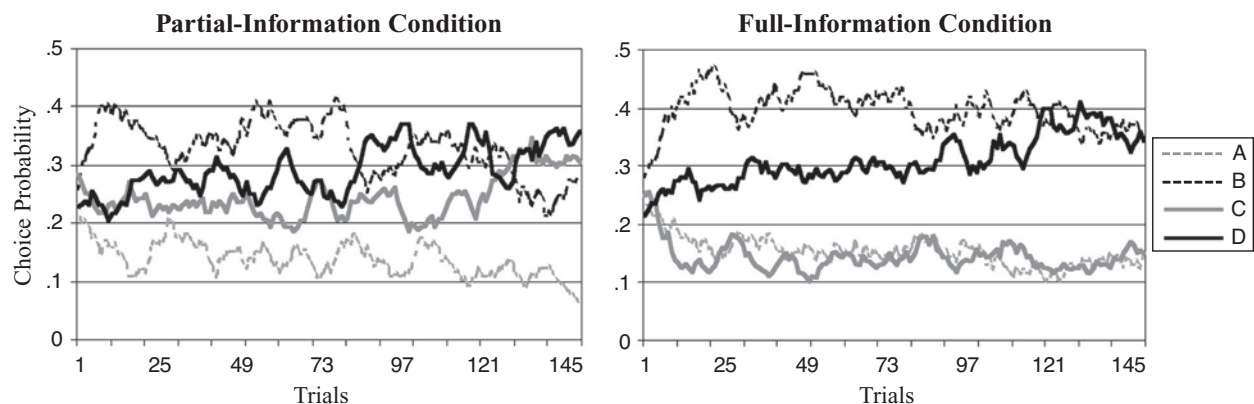


Figure 2. Proportions of choices from the four decks as a function of time for the partial- and full-information conditions. For purposes of comparison with the simulation, the high-payoff results are presented.

condition, in which the decision maker only experiences the payoff for the chosen alternative. Later, we will extend the formulation to the full-information condition, to include forgone-payoff information.

Utility

The evaluation of the gains and losses experienced after making a choice is represented by a prospect-theory type of utility function (Kahneman & Tversky, 1979). The utility is denoted $u(t)$ and is calculated as a weighted average of gains and losses in value for the chosen option (or deck) in trial t :

$$u(t) = W \cdot \text{win}(t)^\gamma - L \cdot \text{loss}(t)^\gamma. \quad (1)$$

The term $\text{win}(t)$ is the amount of money won on trial t ; the term $\text{loss}(t)$ is the amount of money lost on trial t ; W and L are parameters that indicate the weights to gains and losses, respectively; and γ is a parameter that determines the curvature of the utility function. Note that for the small amounts of money used in the present experiment (less than \$20), $\gamma = 1$ was found to be sufficient (estimation of γ produced only minor improvements).

Updating of Expectancies

A general linear model is used to represent all of the previously proposed learning models. (See Bush & Mosteller, 1955, and Estes & Burke, 1953, for the original work on such models.) The expectancy E_j for deck j is updated as a function of its value in the previous trial (which reflects past experience), as well as on the basis of new payoffs, as follows:

$$E_j(t) = \alpha_{jt} \cdot E_j(t-1) + \beta_{jt} \cdot u(t), \quad (2)$$

where α_{jt} and β_{jt} denote the weight given to the old expectancy $E_j(t-1)$ versus the value of the new information in the present trial $u(t)$. The expectancy (accumulated experience from past outcomes) of an alternative is determined by the previous expectancy and the weighted new outcome following a choice. For all models, the initial expectancies were set to $E_j(0) = 0$, reflecting the lack of any prior knowledge about the payoffs from each deck. Different assumptions about the learning weights result in different specific learning models.

Two general classes of models have been proposed to account for how new information is accumulated in a learning task. Under one class of models, the weight of the old expectancy from an alternative changes only if new information is added about that alternative. Thus, the old expectancy from an alternative is discounted only if new information about outcomes from that alternative is presented. This class of models can be labeled *interference models*, because memory is only changed by relevant events and not simply as a function of time (see, e.g., Newell, 1992; Oberauer & Kliegl, 2001). Examples include the delta learning model (see, e.g., Busemeyer & Myung, 1992; Sarin & Vahid, 1999) and Bayesian learning (see, e.g., J. R. Anderson & Matessa, 1992).

In a second class of models, the weight of the old expectancy of an alternative can decrease on each choice trial even if no new information about a particular alternative is presented. Thus, expectations about one option can change as a result of selections of other alternatives. This class can be labeled *decay models*, because decay of memory occurs purely as a function of time, even without the occurrence of interfering events (see, e.g., Atkinson & Shiffrin, 1968; Broadbent, 1958). Examples include the reinforcement learning model used by Erev and Roth (1998) and the EWA model (Camerer & Ho, 1999a, 1999b). The different learning models are reviewed next.

We will initially present the specific assumptions of the models for the partial-information condition. Recall that in this condition, the decision maker is only given feedback about the payoff for the chosen alternative on each trial.

Models That Assume Interference

Delta rule model. Connectionist theories of learning usually employ what is called the delta learning rule (see Gluck & Bower, 1988; Rumelhart, McClelland, & the PDP Research Group, 1987; Sutton & Barto, 1998). This type of learning rule was applied to learning in the repeated play of individual decision tasks by Busemeyer and Myung (1992). It has also been applied to repeated-play games by Sarin and Vahid (1999). According to this model, the two learning weights are defined as follows:

$$\alpha_{jt} = 1 - \delta_j(t) \cdot \phi, \quad (3)$$

and

$$\beta_{jt} = \delta_j(t) \cdot \phi, \quad (4)$$

where $\delta_j(t) = 1$ if payoff information about option j is presented on trial t , or it is 0 otherwise. Formally, Equation 2 can be developed as follows:³

$$\begin{aligned} E_j(t) &= [1 - \delta_j(t) \cdot \phi] \cdot E_j(t-1) + \delta_j(t) \cdot \phi \cdot u(t) \\ &= E_j(t-1) - \delta_j(t) \cdot \phi \cdot E_j(t-1) + \delta_j(t) \cdot \phi \cdot u(t) \\ &= E_j(t-1) + \phi \cdot \delta_j(t) \cdot [u(t) - E_j(t-1)]. \end{aligned} \quad (5)$$

In particular, for the partial-information condition, the delta rule assumes that no changes are made to expectancies unless a deck is chosen, in which case the expectancy is changed in the direction of the prediction error given by $[u(t) - E_j(t)]$. The parameter ϕ is the learning rate parameter, which dictates how much of the expectancy is changed by the prediction error. According to this model, if $0 < \phi < 1$, then the effect of a payoff on the expectancy for a deck decreases exponentially as a function of the number of times a particular deck has been chosen. Thus, recently experienced payoffs have larger effects on the current expectancy than do payoffs that were experienced in the distant past.

Bayesian learning. A model that is popular among economists is the Bayesian learning model (see, e.g., Brown, 1951; Cournot, 1960). For the present application, we assume a beta prior distribution, which leads to the following learning weights (see the Appendix for derivation):

$$\alpha_{jt} = [1 - \delta_j(t)] + \delta_j(t) \cdot \frac{\phi + N_j(t-1)}{\phi + N_j(t)}, \quad (6)$$

$$\beta_{jt} = \delta_j(t) \cdot \frac{1}{\phi + N_j(t)}. \quad (7)$$

For the partial-information condition, $N_j(t)$ is the number of times that deck j is chosen up to and including trial t . Like the delta rule, the Bayesian model also assumes that no changes in expectancies occur unless payoff information is presented for that deck. However, unlike the delta rule, the Bayesian model assumes that the effect of a payoff on the expectancy for a deck remains the same, independent of the number of times a particular deck was chosen. Thus, the Bayesian model implies no serial-position effects at all.⁴

Reinforcement learning with decreasing learning rate. Early reinforcement learning theories (see Luce, 1959) posited an additive effect of reinforcement on propensities. We implemented this idea by using the following set of learning weights:

$$\alpha_{jt} = 1; \quad (8)$$

$$\beta_{jt} = \left[\frac{\delta_j(t)}{1 + N_j(t)^\phi} \right]. \quad (9)$$

Once again, $N_j(t)$ denotes the number of choices from deck j up to and including trial t , and as with the models described above, no change is made unless a deck is chosen and a payoff is received. The parameter ϕ allows some flexibility in the effect of reinforcement with experience. For example, if $\phi = 0$, reinforcement does not depend on amount of experience; if $\phi = 1$, the effect decreases with experience.⁵

Models That Assume Decay

Decay reinforcement model. More recently, Erev and Roth (1998) added a decay or discount parameter to the reinforcement learning model, which can be represented by the following learning rates:

$$\alpha_{jt} = \phi; \quad (10)$$

$$\beta_{jt} = \delta_j(t). \quad (11)$$

This implies that Equation 2 can be expressed simply as:

$$E_j(t) = \phi \cdot E_j(t-1) + \delta_j(t) \cdot u(t). \quad (12)$$

Note that for this model, the past expectancy is always discounted, regardless of whether or not any new payoff information is experienced. In particular, for the partial-information condition, the expectancies for unchosen options are decayed toward zero. In other words, if an option is *not* chosen, so that no information is provided about it, then a payoff of zero is used to update the expectancy.

Experience-weighted attraction (EWA) model. The EWA model (Camerer & Ho, 1999a, 1999b) pro-

vides more general assumptions about learning than do the previous models. It makes the following assumptions about α_{jt} and β_{jt} :

$$C_t = \rho \cdot C_{t-1} + 1; \quad (13)$$

$$\alpha_{jt} = \phi \cdot \frac{C_{t-1}}{C_t}; \quad (14)$$

$$\beta_{jt} = \frac{1}{C_t}. \quad (15)$$

The new term C_t represents an estimate of the count for the number of trials experienced by the decision maker. If we set $\rho = 0$, there is no count, and this model reduces to the updating rule used by Erev and Roth (1998). In order to examine a different model class, we need to impose different constraints on the count for the number of trials. Here we consider two other special cases of the more general EWA model. In the first case, we set $\rho = \phi$, which assumes a common discounting rate for estimating the count and the expectancies. The second case, $\rho = 1$, assumes no discounting of the count and allows for an independent discount of expectancy. In both cases, the expectancy formula is the same:

$$E_j(t) = \frac{C_{t-1} \cdot \phi \cdot E_j(t-1) + u(t) \cdot \delta_j(t)}{C_t}. \quad (16)$$

As with the decay reinforcement model, the weight on past experience can decrease on each trial even if no new payoff information is provided. In particular, for the partial-information condition, expectancies for unchosen options are discounted toward zero. Note that all three versions of the EWA model ($\rho = 0$, ϕ , or 1) have the same number of parameters.

Choice Rules

The choice on each trial is assumed to be determined by the expectancies for each deck according to a choice rule. In the present study, three choice rules are compared.

Max rule with guessing. In the expectation maximization model (see Harless & Camerer, 1994), decisions are based on a comparison of the expected utilities estimated for the different decks. A strict maximizing rule would imply that on each trial, the deck producing the maximum expectancy would be chosen. A more lenient and general rule allows for occasional random guessing. We define $\Pr[G_j(t)]$ as the probability that deck j will be selected on trial t by the model and g_t as a free parameter denoting the probability of guessing. Formally:

$$\chi_j = 1 \text{ if } E_j = \text{Max}\{E_A, E_B, E_C, E_D\};$$

$$\text{otherwise zero.} \quad (17)$$

$$\Pr[G_j(t)] = (1 - g_t) \cdot \chi_j(t) + g_t/4. \quad (18)$$

We compared two common assumptions about the guessing parameter. One alternative is constant guessing, in which $g_t = g$. An alternative suggestion is that guessing

decreases with time (see, e.g., Barron & Erev, 2003). This can be formally presented as:

$$g_t = (1 + t^c)^{-1}, \tag{19}$$

where c is a free parameter that denotes the decrease in guessing as a function of time.⁶

Ratio choice rule. The ratio rule assumes that the choice made on each trial is a probabilistic function of the relative expectancies of the alternatives. This principle implies that decision makers do not optimize utilities but rather use a matching rule in which the probability of choosing a deck is proportional to the relative strength of its expectancy (Luce, 1959). This can be formalized by the following ratio-of-strengths rule:

$$\Pr[G_j(t)] = \frac{e^{\theta(t) \cdot E_j(t)}}{\sum_k e^{\theta(t) \cdot E_k(t)}}, \tag{20}$$

where the parameter $\theta(t)$ controls the sensitivity of the choice probabilities to the expectancies. On the one hand, setting $\theta(t) = 0$ produces random guessing; on the other, as $\theta(t) \rightarrow \infty$ we recover a strict maximizing rule. The probability of choosing the deck producing the largest expectancy increases according to an S-shaped logistic function with a slope (near zero) that increases with $\theta(t)$.

It is assumed that the sensitivity to expectancies, which is denoted by $\theta(t)$, may change as a function of experience. Potentially, the sensitivity can increase in magnitude, reflecting more confidence in the expectancies as experience is accumulated. It can also decrease, reflecting fatigue or boredom. This is parameterized by a power function for the sensitivity change over trials:

$$\theta(t) = (t/10)^c, \tag{21}$$

where c is a free parameter. If $c = 0$, sensitivity does not change over time; positive values of c indicate increasing sensitivity over time; and negative values indicate decreasing sensitivity.

Full-Information Models

Recall that in the full-information condition, the decision maker receives feedback about the payoffs from all options (although payment is based solely on the chosen option). In most economic theories, it is understood that decision makers use all available information in making decisions. However, this is more controversial in experimental economics, particularly among those studying

adaptive choice behavior. We considered three prominent assumptions concerning the effect of full information: The first assumes that weight is given only to feedback from the chosen option (the one that actually pays out), thus ignoring forgone payoffs (see, e.g., Roth & Erev, 1995; Sarin & Vahid, 2001). A second assumes that equal weight is given to feedback from all four decks, indicating an identical effect of forgone and actual payoffs (Fudenberg & Levine, 1998). A third assumes that more weight is given to the chosen option, but some weight is also given to unchosen options (Busemeyer & Myung, 1992; Camerer & Ho, 1999b). In this model, an additional parameter is added to account for the weight given to forgone payoffs. Including this weight parameter in the general learning model produces the following updating equation:

$$E_j(t) = \alpha_{jt} \cdot E_j(t - 1) + \beta_{jt} \cdot \{ \delta_j(t) + \gamma \cdot [1 - \delta_j(t)] \} \cdot u_j(t). \tag{22}$$

In particular, the delta (Equation 5), decay reinforcement (Equation 12), and EWA (Equation 16) models are modified as shown at the bottom of this page.

The parameter γ denotes the weight assigned to payoff feedback from unchosen options, and $u_j(t)$ is the utility of the payoff for option j : $u_j(t)$ is the actual payoff if deck j was chosen on trial t , or it is the forgone payoff if deck j was not chosen on trial t . A γ value of 0 implies no weight, in which case the models all reduce to the same forms used above for the partial-information condition; a γ value of 1 implies full weight, in which case there is no longer any difference between the delta and decay reinforcement models; and values between 0 and 1 indicate decreased weight to forgone in comparison with experienced payoffs.

MODEL EVALUATION

Prediction Method

The first method of model evaluation examined the accuracy of the one-step-ahead predictions generated by each model for each individual. To be more specific, $Y_i(t)$ is defined as a $t \times 1$ vector, representing the sequence of choices made by individual i up to and including trial t , and $X_i(t)$ as the corresponding sequence of payoffs produced by these choices. Each model was given $X_i(t)$ and used this information to generate $\Pr[G_j(t + 1) | X_i(t)]$ for choice trials $t = 1$ to 149 and for

Delta	$E_j(t) = E_j(t - 1) + \phi \cdot \{ \delta_j(t) + \gamma \cdot [1 - \delta_j(t)] \} \cdot [u_j(t) - E_j(t - 1)];$
Reinforcement	$E_j(t) = \phi \cdot E_j(t - 1) + \{ \delta_j(t) + \gamma \cdot [1 - \delta_j(t)] \} \cdot u_j(t);$
EWA	$E_j(t) = \frac{C_{t-1} \cdot \phi \cdot E_j(t-1) + \{ \delta_j(t) + \gamma \cdot [1 - \delta_j(t)] \} \cdot u_j(t)}{C_t}.$

choice options $j = 1$ to 4. The accuracy of these predictions was measured using the log likelihood criterion:

$$\ln(L | \text{model}) = \sum_t \sum_j \ln \{ \Pr[G_j(t+1) | X_i(t)] \cdot \delta_j(t+1), \quad (23)$$

where $\delta_j(t) = 1$ if deck j was chosen on trial t or 0 otherwise.

For the partial-information condition, each model has four parameters $\{W, L, \phi, c\}$ estimated from each person's choices on 150 trials. For the full-information condition, one additional parameter, γ , is added. To optimize the log likelihood for each participant and model, we searched for parameters using a robust combination of grid-search and simplex search (Nelder & Mead, 1965) methods. Each point on the grid served as a starting position for the simplex search algorithm, which was then used to find the parameters that maximized the log likelihood for an individual. This generated a set of solutions, one for each starting point on the grid. Occasionally, these solutions would differ due to local maxima, so we selected the grid point that produced the maximum overall starting positions for the final solution. The parameter search was done to satisfy the following constraints: The values of the W and L parameters were limited between 0 (indicating no attention) and 3. The value of ϕ was limited between -2 and 2 (values lower than 0 for ϕ allow for the gambler's fallacy).⁷ The value of the guessing parameter g in the case of a constant guessing rate was set between 0 and 1, allowing the full range between a deterministic and a random choice. The value of the sensitivity parameter c was set between -5 and 5 , allowing for similar flexibility. Finally, for the full-information condition, the value of the γ parameter was set between 0 and 1, denoting the range between no weight and full weight to forgone payoffs.

In addition to the learning models presented above, we employed a baseline statistical model. The baseline model assumed that the choices were generated by a statistical Bernoulli process. That is, the choice probabilities for each deck were assumed to be constant and statistically independent across trials:

$$\Pr[G_j(t) | X_i(t), Y_i(t)] = p_j. \quad (24)$$

The baseline model had three parameters $\{p_1, p_2, p_3, p_4 = 1 - p_1 - p_2 - p_3\}$, which corresponded to the proportions of choices of Decks A, B, and C pooled across all 150 trials (the last proportion, for Deck D, was determined from the previous three). Therefore, a learning model could do better than the baseline model only if it explained learning effects or trial-to-trial dependencies. The different cognitive models were evaluated by comparing log likelihood scores for the baseline and each of the learning models:

$$G^2 = 2 \cdot [\ln(L | \text{model}) - \ln(L | \text{baseline})]. \quad (25)$$

Because the baseline model had only three parameters, whereas the learning models had at least four, we adjusted for the difference in number of parameters. This

was accomplished by using the Bayesian information criterion (BIC; Schwartz, 1978) statistic to compare models. The BIC is a correction in a model fit that penalizes models with additional parameters:

$$\text{BIC} = G^2 - k \cdot \ln(N), \quad (26)$$

where k is the difference in number of parameters and N is the number of observations. For our comparisons, we had $k = 1$ (one additional parameter in the learning models in comparison with the baseline model) and $N = 150$. Thus, $1 \cdot \ln(150) \approx 5$. This number could be thought of as the deduction from the G^2 of the learning models. Positive values of the BIC statistic indicate that a learning model performs better than the baseline model.

Simulation Method

The second method of model evaluation did not use any information about the actual choices made by any of the participants. In other words, this method was based on evaluation of the predictions of each model (Busemeyer & Wang, 2000), which was done as follows.

Recall that our experiment included two different payoff conditions. We used parameters estimated from the first payoff condition to generate new predictions for the second payoff condition, and then these predictions were evaluated using the independent data from the second payoff condition. Thus, the parameters used to generate the predictions for the second payoff condition were based on the parameter estimates obtained in the first condition (see Rieskamp, Busemeyer, & Laine, 2003). This procedure is relatively parsimonious, since it is based on no further parameter estimation. However, because the two payoff conditions were relatively similar, the procedure cannot be considered a true test of generalization, but rather a cross-validation of the robustness of the set of parameters to changes in the evaluation method and in the payoff magnitude.

The parameters of the model were estimated in the low-payoff condition for one-step-ahead predictions (as detailed in the previous section). Then, the same parameters were used to generate the full simulation path in the high-payoff condition. A total of 1,000 simulations were generated to produce a distribution of choice sequences from a given model in the high-payoff condition, and these results were averaged to produce the probability of choosing each deck on each trial. We then examined the mean square deviation (*MSD*) of the model's predicted probability as compared with the observed proportion of choices on each trial, averaged across participants.

EVALUATION USING THE PREDICTION METHOD

Partial-Information Condition

For examining the (standard) partial-information condition, eight alternatives for the updating of propensities were compared. Five of these alternatives were essen-

tially interference models: the delta rule, the original and modified Bayesian rule, and a reinforcement rule assuming a decreasing learning rate (as a function of either time or number of choices from a deck). The other alternatives were decay models, including the decay reinforcement model and the two instantiations of the EWA model. Three choice rules were simultaneously evaluated: a deterministic maximization rule, with or without a decrease in guessing, and a probabilistic ratio-of-strengths rule. These models were evaluated in an 8 (learning) \times 3 (choice) table of models, which allowed decomposing the unique contribution of each component of a model. A BIC score was obtained for each participant and each model.

Table 3 presents the median BIC scores of models compared using the prediction method, pooled across participants. The results show that it is possible to distinguish the competing models empirically using the BIC index. Most importantly, the models that had the highest BIC scores were all in the decay category, which included the decay reinforcement and the EWA models. The decay reinforcement model produced the highest BIC scores across all assumptions about the choice rule.

An examination of different choice rules shows that, overall, models including the ratio choice rule produced higher BIC scores than did models including either of the two max rules. This result was consistent for all learning models. Thus, the learning model with the highest BIC score (22.17) was the reinforcement learning model that featured decay and ratio choice, although the two instantiations of the EWA model had high fit scores as well.

In addition to examining group medians, we assessed between-participants heterogeneity. Table 4 summarizes the percentages of participants whose average choice proportions were approximated best (had the highest BICs) under the different models. As in the previous comparison, the three decay models outperformed the five interference models. The single model that had the highest BIC for the largest proportion of individual participants was the decay reinforcement model with a ratio choice rule (31.6%). Across all choice rules, the two in-

stances of the EWA model put together had the highest BIC for 29.3% of the participants. These results also show that, interestingly, some of the models that produce a low average BIC for all participants still approximate best the behavior of a small proportion of the players. Notably, the delta learning model under the ratio rule captured best the behavior of about 7% of the participants, as did also the modified Bayesian learning model. The clear bottom line, however, is that a decay model with the ratio-of-strengths choice rule produces better approximations of the next choice ahead.

To examine the effect of individual differences in decay and interference updating, we conducted post hoc Spearman correlation tests between the fit of the model (expressed by the BIC score) and choice proportions from specific decks. The only significant result was a negative correlation between the fit of the delta model and choice from disadvantageous Deck B, the deck with a 10% chance for large losses [$r(76) = -.28, p < .05$]. Thus, it appears that the delta model was less successful in predicting the behavior of those who had many choices from Deck B. The decay models were impartial to this behavior, because the preference for Deck B is easily captured by a decay formula with an extreme recency parameter.⁸

Full-Information Condition

For the full-information condition, three learning models were selected for further testing. The delta learning model was selected as the one with the highest BIC from the interference models. The decay reinforcement and the EWA (with $\rho = \phi$) models were selected for their high BIC scores in the decay group. These learning models were examined under two choice rules: the probabilistic ratio rule and the maximization with constant guessing rule. In addition, three basic assumptions regarding full information were compared: (1) Decision makers may simply ignore the forgone payoff information. In this case, the difference between decay and interference models is exactly the same as in the partial-information condition. Under this assumption, expectancies are updated with $\gamma = 0$ —that is, weight is given only to feedback from the chosen deck. (2) Decision makers may

Table 3
Partial-Information Condition: Bayesian Information Criterion (BIC)
Medians and Standard Deviations of the Models

	Max, Constant Guess		Max, Decrease Guess		Ratio Choice	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Delta learning (I)	-1.72	26.0	-7.97	46.1	3.01	21.0
Bayesian learning (I)	-9.50	28.8	-11.50	30.1	-3.55	20.6
Modified Bayesian learning (I)	-3.67	47.6	-4.29	47.4	0.94	40.3
Reinforcement learning—Time dependent (I)	-8.60	21.0	-12.07	20.8	1.11	15.9
Reinforcement learning—Choice dependent (I)	-8.30	21.1	-9.33	20.9	1.21	17.8
Decay reinforcement (D)	5.32	44.9	5.15	45.6	22.17	37.9
EWA, $\rho = \phi$ (D)	4.90	46.0	5.10	47.0	19.48	39.9
EWA, $\rho = 1$ (D)	4.00	46.2	2.87	47.6	18.72	34.4

Note—I, interference models; D, decay models.

Table 4
Partial-Information Condition: Percentages of Individual Participants Whose Average Behavior Was Best Approximated (According to BIC) by the Different Models

	Max, Constant Guess (%)	Max, Decrease Guess (%)	Ratio Choice (%)
Delta learning (I)	1.3	0	6.6
Bayesian learning (I)	0	0	2.6
Modified Bayesian learning (I)	1.3	1.3	6.6
Reinforcement learning—Time dependent (I)	0	0	3.9
Reinforcement learning—Choice dependent (I)	0	0	6.6
Decay reinforcement (D)	4.8	6.6	31.6
EWA, $\rho = \phi$ (D)	3.5	4.6	10.5
EWA, $\rho = 1$ (D)	2.2	7.2	1.3

Note—I, interference models; D, decay models.

give equal weight to actual and forgone payoffs, so that $\gamma = 1$ and equal weight is given to payoffs observed for all four decks. In this case, there is no mathematical difference between the delta and decay reinforcement models.⁹ (3) Decision makers may give more weight to actual payoffs, but still give smaller weight to forgone payoffs. In this last case, γ is a free parameter, and the difference between interference and decay models depends on the weight assigned by the model to unchosen alternatives.

Table 5 presents the BIC scores of the models. The results indicate that as in the partial-information condition, the BIC scores differentiate between competing models. First, as in the partial-information results, the BIC scores under the ratio choice rule were relatively higher than under the maximization rule (the latter scores were all negative). Second, under the ratio rule, the most flexible of the assumptions, assigning more weight to the chosen deck produced the highest BIC scores for all three learning models. Finally, again as in the partial-information condition, the decay reinforcement model produced the highest BIC score (7.76).

Note that using both choice rules, the difference between the delta model's BIC and the two other models' is largest under the assumption of giving weight only to chosen decks (with the ratio rule, a difference of 9.8 from the best BIC; with the max rule, a difference of 10.5). Under the assumption of equal weight to both payoffs, in which case all decks are updated regardless of choice, there is mathematically no difference between the delta and the decay reinforcement models, and both are slightly

better than the EWA model with $\rho = \phi$. Under the assumption of differential weighting among decks, the difference between the delta model and the best BIC increases again (with ratio, 9.9; with max, 4.5). Thus, it appears that the poorer performance of the delta model for predictions of one step ahead is largely due to the facts that (1) models that assign low weight to unchosen decks are more successful and (2) the delta model assumes that expectancies of unchosen decks are updated at a different rate than those of chosen decks.

EVALUATION USING THE SIMULATION METHOD

Partial-Information Condition

The simulation focused on the combination of the best three learning rules from the interference and decay classes and the two choice rules examined above. To recall, the simulation was an examination of the models using the parameters estimated in the previous section with the prediction method. Table 6 summarizes the parameters of each model. These parameters were computed by averaging across the individual estimates obtained using the prediction method for participants in the low-payoff condition.

The parameters were used to generate the simulations for the high-payoff condition. Note that a priori predictions were generated for the high-payoff condition because the parameters were based on the estimates obtained from the low-payoff condition. Table 7 (top) presents the *MSDs*

Table 5
Full-Information Condition: Bayesian Information Criterion (BIC) Medians and Standard Deviations of the Models Under the Assumption of Partial Weighting of Forgone Payoffs

Choice Rule	Learning	Weight Only to Chosen Deck		Equal Weight to All Decks		More Weight to Chosen Deck	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Ratio choice	Delta model	-6.60	19.9	-4.95	50.8	-2.10	32.1
	Decay reinforcement	3.19	23.0	-4.95	50.9	7.76	23.5
	EWA model, $\rho = \phi$	0.95	23.9	-17.4	54.1	3.10	24.2
Max, constant guessing	Delta model	-18.27	25.1	-15.60	34.5	-13.66	26.4
	Decay reinforcement	-7.75	27.3	-15.60	34.5	-9.14	26.6
	EWA model, $\rho = \phi$	-13.99	39.6	-20.96	52.7	-9.52	26.5

Table 6
Partial-Information Condition: Averages and Standard Deviations
of the Estimated Parameters in the Low-Payoff Version

Choice Rule	Learning	ϕ		W		L		c/g	
		M	SD	M	SD	M	SD	M	SD
Ratio choice	Delta model	0.31	0.44	0.25	0.51	0.18	0.34	0.33	1.97
	Decay reinforcement	0.37	0.47	0.73	0.90	0.69	1.04	0.68	1.42
	EWA model, $\rho = \phi$	0.49	0.38	1.60	1.20	0.99	1.10	0.75	1.11
Max, constant guessing	Delta model	0.66	0.65	0.55	0.91	0.51	0.82	0.40	0.18
	Decay reinforcement	0.54	0.77	0.64	1.06	0.35	0.74	0.43	0.23
	EWA model, $\rho = \phi$	0.01	1.32	0.67	1.06	0.67	1.00	0.34	0.15

Note— ϕ , learning rate; W , win weight; L , loss weight; c , sensitivity; g , guessing rate.

of the different models for the high-payoff condition. Note that the *MSDs* are based on percentage scores rather than proportions.

The results indicate that the simulation method shows smaller differences between models than does the prediction method. In terms of updating propensities, the best model for both payoff conditions was the EWA model with $\rho = \phi$, but the delta model with the maximization choice rule had slightly better fit than did the decay reinforcement model. As for choice rules, it appears that the large advantage of the ratio choice rule over the maximization rule that appeared for the prediction method diminished in the simulation. This result is likely due to the fact that, averaged over many individuals, sensitivity in the simulation remained constant in time. Thus, the assumption of changing sensitivity over time embedded in the ratio choice rule was unnecessary.

To understand these results, we plotted in Figure 3 the path estimated by the decay reinforcement and the delta models with either the ratio rule or the maximization rule. The trial-by-trial predictions were smoothed by using a moving average filter of seven trials. The results show that the delta model with the ratio rule predicts fewer choices made from Deck B, the deck with the infrequent but high losses. In fact, under any combination of parameters (that we examined) the delta plus ratio model cannot reproduce the strong preference for Deck B. Under the delta plus ratio rule, B cannot be preferred over both D, which produces smaller gains than B, and A, which produces smaller losses. Guessing and decay both lead to a recovery of the expectancies of B, because most of the time this deck produces results that have a high expected value.

Under the maximization rule, the fit of the delta model improves, but it predicts no learning. The maximization rule likewise predicts no learning with the decay reinforcement model. For example, for the decay reinforcement model with the ratio choice rule, the average predicted change in a participant's choices (between the first block of 25 trials and the last) is 8.9%. Under the maximization choice rule, the change is much smaller (only 2% on average). Thus, the maximization choice rule produces a prediction that is less changeable in time, which fits the relatively flat average choice path.

Full-Information Condition

As in the partial-information condition, the parameters obtained with the prediction method were used to generate simulations for the high-payoff condition with full information. Table 8 summarizes the estimated parameters (using the prediction method) for the low-payoff condition. Table 9 (top) presents the *MSDs* of the different models for the high-payoff condition. Only the simulations for the unequal-weight model, which produced the highest BICs, are shown.

The results show that in the full-information condition, both decay models (decay reinforcement and EWA) were advantageous over the delta model. In addition, they show (as in the partial-information condition) that the clear difference between choice rules that was observed using the prediction method disappeared. The maximization choice rule led to better fits than did the ratio rule for the delta model, and the two choice rules produced roughly similar fits for the EWA model.

Figure 4 presents the paths estimated in the full-information condition for the decay reinforcement and

Table 7
Partial-Information Condition: *MSDs* of the Models Examined Using Simulations With
Parameters Extracted From the Low-Payoff Version by the Prediction Method (Top)
or by Estimation for the Entire Game Path (Bottom)

Origin of Parameters	Model	Ratio Choice	Max, Constant Guess
Prediction method (low-payoff condition)	Delta model	1.50	0.62
	Decay reinforcement	0.86	0.68
	EWA model, $\rho = \phi$	0.59	0.92
Simulation method (low-payoff condition)	Delta model	1.10	0.48
	Decay reinforcement	0.48	0.52
	EWA model, $\rho = \phi$	0.50	0.50

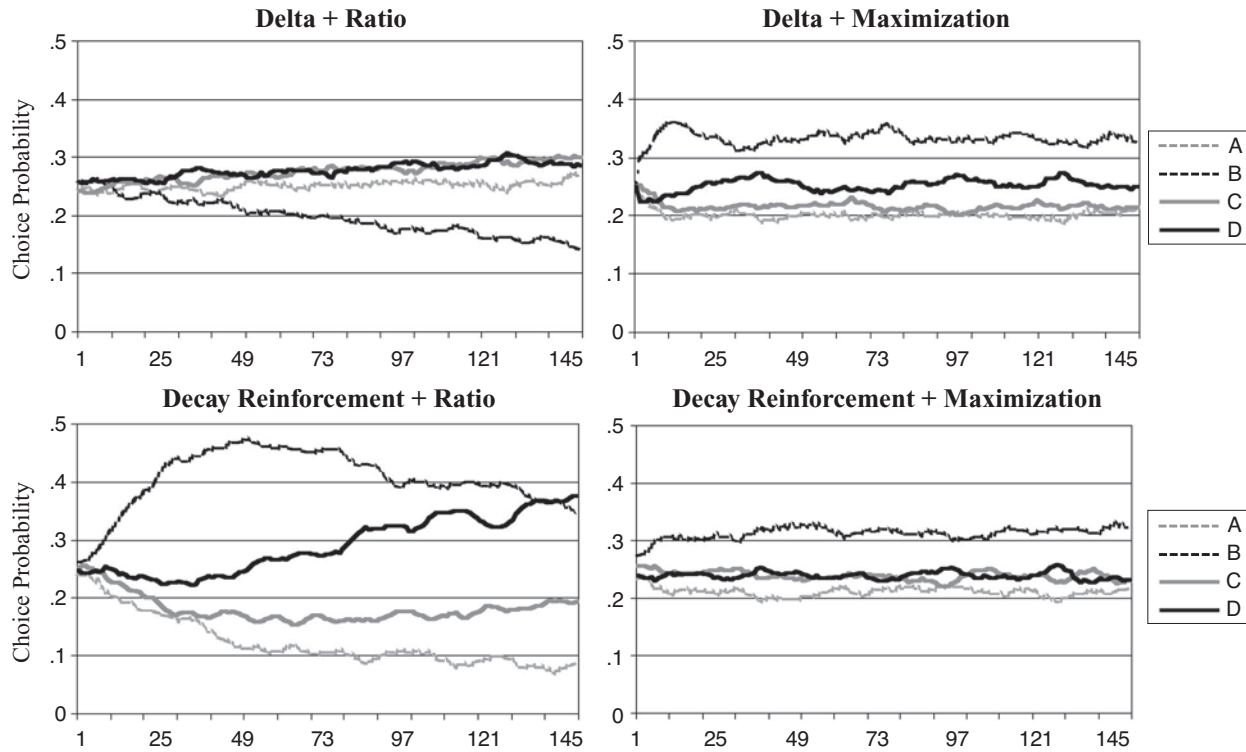


Figure 3. Partial-information condition: Simulated choices in the high-payoff version using four alternative models.

delta models with either the ratio or the maximization rule. The figure indicates that one likely reason for the advantage of decay over interference models is the ability of the decay models to predict the increase in choices from disadvantageous Deck B in the full-information condition as compared with the partial-information condition. It appears that in the full-path simulation, the delta model cannot predict the extent to which this deck (with its low-frequency but high negative payoff) is preferred. The *MSD* for this particular deck for the delta model with maximization is 0.11, in comparison with 0.07 in the corresponding decay reinforcement mode (a 37% difference). For the other decks, the *MSD* of the delta model is better (A, 0.04; C, 0.09; D is not independent) and more similar to the decay reinforcement model (A, 0.06; C, 0.08).

In summary, although an evaluation of the learning models in the partial-information condition using the simulation method yielded less conclusive results, in the full-information condition decay models were clearly advantageous. The differences in the degree of fit between interference and decay models can largely be attributed to the fact that decay models can more easily capture the preference for the deck that featured large losses occurring 10% of the time.

Verification Using an Alternative Simulation Method

One potential limitation of the results using the simulation method is that they rely on the generalization of parameters estimated for one-step-ahead predictions to simulations of the entire average choice path. An alter-

Table 8
Full-Information Condition: Averages and Standard Deviations of the Estimated Parameters in the Low-Payoff Version Under the Assumption of Partial Weighting of Forgone Payoffs

Choice Rule	Learning	ϕ		W		L		c/g		γ	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Ratio choice	Delta model	0.73	0.66	0.36	0.70	0.30	0.66	0.19	1.72	0.42	0.42
	Decay reinforcement	0.50	0.50	0.67	0.91	0.76	1.16	0.31	1.32	0.25	0.38
	EWA model, $\rho = \phi$	0.53	0.43	1.64	1.26	1.07	1.16	0.70	1.08	0.22	0.37
Max, constant guessing	Delta model	0.61	0.75	0.39	0.63	0.65	0.84	0.38	0.16	0.36	0.31
	Decay reinforcement	0.65	0.56	0.94	1.04	0.59	0.87	0.40	0.18	0.37	0.36
	EWA model, $\rho = \phi$	0.58	0.37	0.92	1.04	0.45	0.76	0.40	0.17	0.45	0.35

Note— ϕ , learning rate; W , win weight; L , loss weight; c , sensitivity; g , guessing rate; γ , weight for unchosen options.

Table 9
Full-Information Condition: MSDs of the Models Examined Using Simulations With Parameters Extracted From the Low-Payoff Version by the Prediction Method (Top) or by Estimation for the Entire Game Path (Bottom)

Origin of Parameters	Model	Ratio Choice	Max, Constant Guess
Prediction method (low-payoff condition)	Delta model	1.20	0.79
	Decay reinforcement	0.58	0.68
	EWA model, $\rho = \phi$	0.69	0.71
Simulation method (low-payoff condition)	Delta model	1.90	0.49
	Decay reinforcement	2.00	0.43
	EWA model, $\rho = \phi$	0.59	0.41

Note—Only results for unequal weights to all decks are presented, since these produced the highest Bayesian information criterion scores.

native view (Haruvy & Erev, 2002) argues that such generalization is not always possible.

Supporting the latter view is the present finding that using the prediction method under all models, the attention to gains parameter W was on average higher than the attention to losses parameter L , indicating that gains loom greater in the mind than do losses (see Tables 6 and 8). This finding appears to be inconsistent with Barron and Erev's (2003) simulation results, and with other robust findings obtained using the simulation method, showing that people are more sensitive to losses than to gains. An alternative explanation is that, in the gambling task, losses occur frequently and are therefore less salient.

To guard against the possibility that the parameters assessed by the prediction method do not generalize to the simulation method (see Haruvy & Erev, 2002), we reestimated the parameters by fitting the simulations to the entire average choice path in the low-payoff condition. An examination of the two best-fitting models in this simulation (delta plus maximization and decay reinforcement with the ratio rule) shows that for both, the losses parameter was higher than the gains parameter in the partial- and full-information conditions (for conciseness, the full results are not detailed). That is, in this simulation, losses loomed greater than gains. However, the simulation with estimated parameters replicated the small advantage of

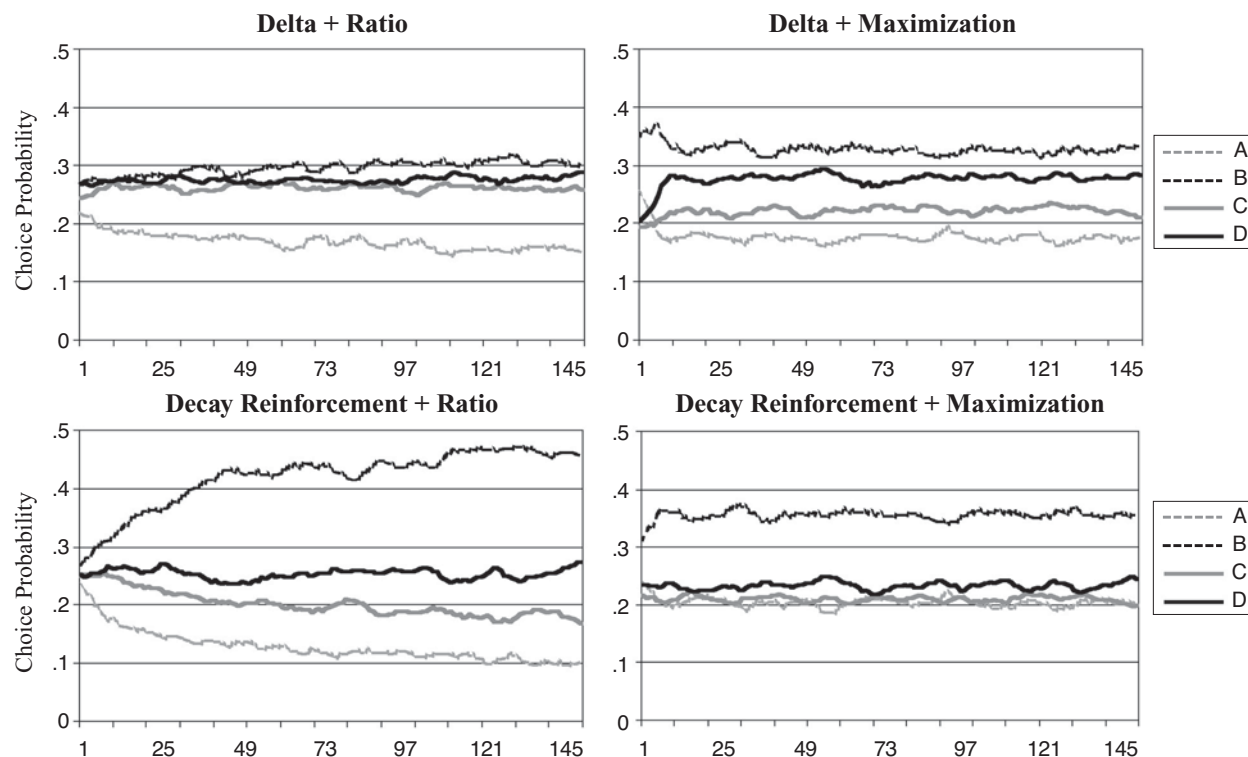


Figure 4. Full-information condition: Simulated choices in the high-payoff version using four alternative models.

decay models in the partial- and full-information conditions (for *MSDs*, see the bottom sections of Tables 7 and 9). Furthermore, the results with this method distinctly show that the maximization with guessing rule was as adequate as the ratio rule in the partial-information condition and that it improved fits over the ratio rule in the full-information condition.

GENERAL DISCUSSION

The present study highlights the importance of examining the different assumptions that underlie learning models about the components of the learning and choice processes. Regarding the learning process, an important difference emerged between decay and interference models. The results of the analysis using the prediction method showed that, in the standard Iowa gambling task, decay models were superior to interference models under robust assumptions pertaining to the choice rule. In the simulation, the advantage of decay models was less distinct for the partial-information condition, but it was clear nonetheless that decay models were not outperformed, indicating that their advantage in predictions of one step ahead is not an artifact due to post hoc model fitting of parameters using information about past choices.

The advantage of decay models appeared more strongly in the full-information condition. In this condition, the advantage of decay over interference models was highest when the model ignored forgone payoffs. In this case, interference models update only the selected alternative, and they are therefore most distinct from decay models. Models assuming partial weighting of forgone payoffs take into account the experience of unchosen alternatives. Hence, under such partial weighting the similarity of the interference and decay models increases, because interference models update expectancies for unselected decks as well; the value of the interference models also appears to increase in this case. Yet, the fact that interference models do not fully update the expectancies of unchosen alternatives still appears to lead to poorer predictions. Most people's behavior is best described by a model that discounts past expectancies of alternatives, regardless of whether they were selected.

Previously, it has been suggested that in repeated-choice tasks decision makers are extremely sensitive to the value of the last three or four choices made (Hertwig et al., 2004). This is considered to be one of the factors that leads to underweighting small probabilities in repeated choices (Barron & Erev, 2003; Hertwig et al., 2004), as opposed to overweighting them in single choices (Kahneman & Tversky, 1979). In the present context, underweighting was explained on the basis of the tendency to discount large losses that occur infrequently. Indeed, there was a positive correlation between the amount of decay in the decay reinforcement model (expressed by the recency parameter) and choices from disadvantageous Deck B, which produced rare but large losses.

One plausible interpretation for the tendency to decay expectancies of unchosen decks is that decay in expectancy is a motivational phenomenon similar to the recency effect. The advantage of the decay model implies that players are sensitive to what has happened in the last couple of trials and discount outcomes from the more distant past. This tendency, coupled with a nondeterministic choice rule, seems to be important for adapting to a rapidly changing choice environment in which the most recent trials provide a good sample of the present state of the environment (Gonzalez, Lerch, & Lebiere, 2003). Extreme decay fares less well (in terms of performance level) in a static environment with rare negative payoffs. Such rare negative payoffs can be discounted more easily, which allows the expectancy of the alternatives that produced them to recover.

An alternative explanation is that the tendency of past expectancies to decay is a result of working memory limitations. That is, keeping track of unselected decks requires mental effort, especially in the full-information condition. This can reduce the attractiveness (i.e., the expectancies) of unselected decks as a function of the lag in their selection (see C. J. Anderson, 2003), as is formalized in the decay model. Future studies will be necessary to examine whether the decay in expectancy is a purely motivational phenomenon or is affected by memory constraints.

In our study, we examined the advantage of decay models using a choice task in which not updating unchosen alternatives was expected to have significant outcomes. In simpler binary choice tasks (which are more commonly examined in studies of individual decision-making tasks), the difference between the two classes of models is assumed to be relatively small, because both models change the ratio of the expectancies between the two alternatives. In the Iowa gambling task, where there are four alternatives, this difference appears to create significant disparities between models that update and those that do not update the expectancy of unchosen alternatives.

In contrast to the converging evidence pertaining to the updating of expectancies, the results of our two evaluation methods, prediction and simulation, showed less similarity regarding choice rules. Using the prediction-of-one-step-ahead method, the ratio-of-strengths rule was advantageous under all assumptions pertaining to the updating of expectancies. Using the simulation method, the differences in results between the choice rules were smaller, and in some cases (most notably, in the full-information condition) the maximization rule with constant guessing actually outperformed the ratio rule. We argue that this difference derives from the value of the ratio rule for capturing individual differences in sensitivity changes over time. The ability of the ratio rule to tap varying changes in sensitivity in different performers leads to improved fit in the prediction of individual performers' choices. However, when aggregated over dif-

ferent performers, the change in sensitivity is marginal, and modeling it does not improve predictions.

A Note on Methodology

The present study employed two methods for evaluating the accuracy of the assumptions of learning models. The first was to use the prediction of the next step ahead made by individual decision makers. Under this method, an advantage of decay models was due to their allowance for a steeper reduction in the weight of old expectancies as a function of time. The predictions of decay models may therefore mimic a regression model based on the player's previous choices. This implies that the advantage to decay models observed in the prediction method could be specific to the method of predicting the next choice ahead based on the player's past experiences.

For this reason, we extended our investigation in two ways. First of all, we examined an agent-based simulation of an entire game path that predicted the aggregated choice pattern. This simulation method did not have input from previous choices made by a player, but rather, only received prior choice of the model itself (i.e., the agent). Second, we examined the generality of the simulation for a data set in which the parameters were not estimated. This was done by using the parameters estimated in one payoff condition to simulate the behavior observed in another payoff condition.

Some researchers have suggested that the value of different models may be highly specific to the precise game, its parameters, and the evaluation method (Feltovich, 2000; Haruvy & Erev, 2002; Salmon, 2001). For example, it has been argued that model evaluation based on the examination of group averages can lead to different results than the examination of individuals' choices. Likewise, an evaluation based on the prediction of the next choice ahead may lead to different results than another evaluation based on the prediction of many choices ahead (see Rapoport, Daniel, & Seale, 1998; Stahl, 1996).

In the present study, differences between the rankings of basic assumptions under different evaluation techniques did appear in the ranking of choice rules. The ratio-of-strengths rule was clearly advantageous when using the prediction method, but it fared less well in a simulation of the entire average choice path. However, for another basic assumption, the updating of expectancies, there were more similarities than differences under the distinct evaluation methods.

Thus, the present study goes one step farther than previous studies have, by showing an implicit embedded association between basic assumptions and the evaluation method. It is the belief of the authors that we were enabled to observe this interaction by insisting on two methodological constraints: First, our predictions were driven at the individual level of analysis (see Hertwig & Todd, 2000). Second, we simultaneously compared different assumptions of learning models rather than whole

modeling approaches. For example, consider a pair of modeling approaches: One has better predictions regarding the updating of expectancies, whereas the other has better predictions regarding the choice rule. A global comparison of the predictions of the two approaches (the common method in experimental economics) would, in this case, miss the value of particular assumptions in each model that lead to better predictions. Under the present approach, the different components were evaluated simultaneously by examining all of the possible combinations of assumptions, and an examination of different data sets prevented potential overfitting due to multiple comparisons. The present approach therefore provides an attempt to examine learning models in a way that would ensure the accumulation of knowledge regarding the value of basic assumptions for predicting behavior.

REFERENCES

- ANDERSON, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin*, **129**, 139-167.
- ANDERSON, J. R., & MATESSA, M. (1992). Explorations of an incremental, Bayesian algorithm for categorization. *Machine Learning*, **9**, 275-308.
- ATKINSON, R. C., & SHIFFRIN, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.
- BARKAN, R., ZOHAR, D., & EREV, I. (1998). Accidents and decision making under uncertainty: A comparison of four models. *Organizational Behavior & Human Decision Processes*, **74**, 118-144.
- BARRON, G., & EREV, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, **16**, 215-233.
- BECHARA, A., DAMASIO, A. R., DAMASIO, H., & ANDERSON, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, **50**, 7-15.
- BECHARA, A., & DAMASIO, H. (2002). Decision-making and addiction (part I): Impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia*, **40**, 1675-1689.
- BROADBENT, D. E. (1958). *Perception and communication*. London: Pergamon.
- BROWN, G. W. (1951). Iterative solution of games by fictitious play. In T. C. Koopmans (Ed.), *Activity analysis of production and allocation* (pp. 374-376). New York: Wiley.
- BUSEMEYER, J. R., & MYUNG, I. J. (1992). An adaptive approach to human decision-making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General*, **121**, 177-194.
- BUSEMEYER, J. R., STOUT, J. C., & FINN, P. R. (in press). Using computational models to help explain decision making: Processes of substance abusers. In D. Barch (Ed.), *Cognitive and affective neuroscience of psychopathology*. New York: Oxford University Press.
- BUSEMEYER, J. R., & WANG, Y.-M. (2000). Model comparisons and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology*, **44**, 171-189.
- BUSH, R. R., & MOSTELLER, F. (1955). *Stochastic models for learning*. New York: Wiley.
- CAMERER, C., & HO, T.-H. (1999a). EWA learning in games: Preliminary estimates from weak-link games. In D. V. Budescu, I. Erev, & R. Zwick (Eds.), *Games and human behavior: Essays in honor of Amnon Rapoport* (pp. 31-52). Mahwah, NJ: Erlbaum.
- CAMERER, C., & HO, T.-H. (1999b). Experience-weighted attraction learning in normal form games. *Econometrica*, **67**, 827-874.

- CHEUNG, Y.-W., & FRIEDMAN, D. (1997). Individual learning in normal form games: Some laboratory results. *Games & Economic Behavior*, **19**, 46-76.
- CLARK, L., & ROBBINS, W. (2002). Decision-making deficit in drug addiction. *Trends in Cognitive Psychology*, **6**, 361-362.
- COURNOT, A. (1960). *Researches into the mathematical principles of the theory of wealth* (N. Bacon, Trans.). London: Haffner.
- EREV, I., & BARRON, G. (in press). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*.
- EREV, I., & RAPOPORT, A. (1998). Magic, reinforcement learning and coordination in a market entry game. *Games & Economic Behavior*, **23**, 146-175.
- EREV, I., & ROTH, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, **88**, 848-881.
- ESTES, W. K., & BURKE, C. J. (1953). A theory of stimulus variability in learning. *Psychological Review*, **6**, 276-286.
- FELTOVICH, N. (2000). Reinforcement-based vs. beliefs-based learning in experimental asymmetric-information games. *Econometrica*, **68**, 605-641.
- FUDENBERG, D., & LEVINE, D. K. (1995). Consistency and cautious fictitious play. *Journal of Economic Dynamics & Control*, **19**, 1065-1089.
- FUDENBERG, D., & LEVINE, D. K. (1998). *The theory of learning in games*. Cambridge, MA: MIT Press.
- GLUCK, M. A., & BOWER, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, **117**, 227-247.
- GONZALEZ, C., LERCH, J. F., & LEBIERE, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, **27**, 591-635.
- HARLESS, D. W., & CAMERER, C. F. (1994). The predictive utility of generalized expected utility theories. *Econometrica*, **62**, 1251-1289.
- HARUVY, E., & EREV, I. (2002). On the application and interpretation of learning models. In R. Zwick & A. Rapoport (Eds.), *Experimental business research* (pp. 285-300). Boston: Kluwer.
- HERTWIG, R., BARRON, G., WEBER, E. U., & EREV, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, **15**, 534-539.
- HERTWIG, R., & TODD, P. M. (2000). Biases to the left, fallacies to the right: Stuck in the middle with null hypothesis significance testing: Commentary on Krueger on social bias. *Psychology*, **11**(28).
- KAHNEMAN, D., & TVERSKY, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47**, 263-291.
- LUCE, R. D. (1959). *Individual choice behavior*. New York: Wiley.
- NELDER, J. A., & MEAD, R. (1965). A simplex method for function minimization. *Computer Journal*, **7**, 308-313.
- NEWELL, A. (1992). Unified theories of cognition and the role of Soar. In J. A. Michon & A. Akyurek (Eds.), *Soar: A cognitive architecture in perspective* (pp. 25-75). Dordrecht: Kluwer.
- OBERAUER, K., & KLIEGL, R. (2001). Beyond resources—Formal models for complexity effects and age differences in working memory. *European Journal of Cognitive Psychology*, **13**, 187-215.
- RAPOPORT, A., DANIEL, T. E., & SEALE, D. A. (1998). Reinforcement-based adaptive learning in asymmetric two-person bargaining with incomplete information. *Journal of Experimental Economics*, **1**, 221-253.
- RIESKAMP, J., BUSEMEYER, J. R., & LAINE, T. (2003). How do people learn to allocate resources? Comparing two learning theories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 1066-1081.
- ROTH, A. E., & EREV, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games & Economic Behavior*, **8**, 164-212.
- RUMELHART, D. E., MCCLELLAND, J. L., & THE PDP RESEARCH GROUP (1987). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vols. 1 & 2). Cambridge, MA: MIT Press.
- SALMON, T. (2001). An evaluation of econometric models of adaptive learning. *Econometrica*, **69**, 1597-1628.
- SARIN, R., & VAHID, F. (1999). Payoff assessments without probabilities: A simple dynamic model of choice. *Games & Economic Behavior*, **28**, 294-309.
- SARIN, R., & VAHID, F. (2001). Predicting how people play games: A simple dynamic model of choice. *Games & Economic Behavior*, **34**, 104-122.
- SCHWARTZ, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **5**, 461-464.
- STAHL, D. (1996). Boundedly rational rule learning in a guessing game. *Games & Economic Behavior*, **16**, 303-330.
- SUTTON, R. S., & BARTO, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- YECHIAM, E., & BUSEMEYER, J. R. (2005). *The effect of forgone payoffs on underweighting small probability events*. Manuscript submitted for publication.
- YECHIAM, E., STOUT, J. C., BUSEMEYER, J. R., ROCK, S. L., & FINN, P. R. (2005). Individual differences in the response to forgone payoffs: An examination of high functioning drug abusers. *Journal of Behavioral Decision Making*, **18**, 97-110.
- YECHIAM, E., VEINOTT, E. S., BUSEMEYER, J. R., & STOUT, J. C. (in press). Cognitive models for evaluating basic decision processes in clinical populations. In R. Neufeld (Ed.), *Advances in clinical cognitive science: Formal modeling and assessment of processes and symptoms*. Washington, DC: American Psychological Association.

NOTES

1. The study examined the relationship between drug abuse and performance in the Iowa gambling task.

2. Choices of Deck A decreased from an average proportion of .18 to .13 [$t(75) = 2.62, p < .05$]. Choices of B decreased from an average of .36 to .28 [$t(75) = 2.85, p < .01$]. Choices of advantageous Deck C increased from an average of .23 to .29 [$t(75) = 2.11, p < .05$]. Choices of D likewise increased from an average of .24 to .31 [$t(75) = 2.55, p < .05$].

3. Equation 5 was further developed as follows: $E_j(t) = E_j(t-1) + \phi \cdot \delta_j(t) \cdot [u(t) - E_j(t-1)] = E_j(t-1) + \phi \cdot \delta_j(t) \cdot [W \cdot \text{win}(t) - L \cdot \text{loss}(t) - E_j(t-1)] = E_j(t-1) + \delta_j(t) \cdot [W \cdot \phi \cdot \text{win}(t) - L \cdot \phi \cdot \text{loss}(t) - \phi \cdot E_j(t-1)] = E_j(t-1) + \delta_j(t) \cdot [b_1 \cdot \text{win}(t) - b_2 \cdot \text{loss}(t) - \phi \cdot E_j(t-1)]$. This last step was taken to make parameter ϕ independent of the loss and gain parameters, as it is in the decay models.

4. A modification of the Bayesian learning model can be made to allow for serial-position effects:

$$\alpha_i = [1 - \delta_j(t)] + \delta_j(t) \cdot \frac{N_j(t)^\phi}{1 + N_j(t)^\phi}, \quad \beta_i = \delta_j(t) \cdot \frac{1}{1 + N_j(t)^\phi}.$$

5. An alternative reinforcement learning model defines $\alpha_{ji} = 1$ and $\beta_{ji} = \delta_j(t)/(1 + t^\phi)$, where ϕ is a free parameter. In this model, the weight is given to the current payoff decreases as a function of time rather than of number of choices.

6. We also examined a decrease in guessing as a function of the number of choices of a deck, but since this model did not improve predictions, for the sake of conciseness we do not report these results.

7. Camerer and Ho (1999b) suggested that in the EWA model, values of ϕ that are greater than 1 require a modified learning rule. We therefore constrained the value of the ϕ parameter in both versions of that model to be between 0 and 1. Thus, for the sake of parsimony, our models capture only specific cases of the more general EWA model.

8. This conclusion is further supported by correlations between the recency parameter and choices from the different decks under the ratio rule. There was a positive correlation between the recency parameter of the decay reinforcement model and choices from disadvantageous Deck B, which produces rare but large losses [$r(56) = .28, p < .05$; only for $0 < \phi < 1$; see Yechiam, Veinott, Busemeyer, & Stout, in press]. In contrast, there was no such association for the delta model [$r(42) = .06, n.s.$].

9. In this case, the difference between the EWA model and the other two is a result of the inclusion of the C_i factor in the EWA model. Note also that, although the delta model is identical mathematically to the decay reinforcement model, small differences may emerge as a result of parameter constraints.

APPENDIX
A Derivation of the Bayesian Formula for Choice of Gambles

$$E_j(t) = p_{\text{lose},j}(t) \cdot u_{\text{lose}} + p_{\text{win},j}(t) \cdot u_{\text{win}}$$

Following a choice of gamble j , we update using beta priors:

$$p_{\text{lose},j}(t) = \frac{n_{\text{lose},0} + n_{\text{lose},j}(t)}{N_0 + N_j(t)} = \frac{n_{\text{lose},0} + n_{\text{lose},j}(t-1) + \delta_j(t)}{N_0 + N_j(t)}$$

$$p_{\text{win},j}(t) = \frac{n_{\text{win},0} + n_{\text{win},j}(t)}{N_0 + N_j(t)} = \frac{n_{\text{win},0} + n_{\text{win},j}(t-1) + [1 - \delta_j(t)]}{N_0 + N_j(t)}$$

$$u(t) = \delta_j(t) \cdot u_{\text{lose}} + [1 - \delta_j(t)] \cdot u_{\text{win}}.$$

Inserting these terms yields

$$E_j(t) = \frac{n_{\text{lose},0} + n_{\text{lose},j}(t-1)}{N_0 + N_j(t)} \cdot u_{\text{lose}} + \frac{n_{\text{win},0} + n_{\text{win},j}(t-1)}{N_0 + N_j(t)} \cdot u_{\text{win}} + \frac{u(t)}{N_0 + N_j(t)}$$

$$= \frac{N_0 + N_j(t-1)}{N_0 + N_j(t)} \cdot E_j(t-1) + \frac{1}{N_0 + N_j(t)} \cdot u(t)$$

$$= \alpha_t \cdot E_j(t-1) + \beta_t \cdot u(t).$$

(Manuscript received June 2, 2004;
revision accepted for publication November 10, 2004.)