The effect of foregone payoffs on choices from experience: An individual level modeling analysis

Eldad Yechiam – Technion – Israel Institute of Technology, Haifa, Israel

Tim Rakow – University of Essex, Colchester, UK

Address for correspondence:

Behavioral Science Area, Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa 32000, Israel. Phone: (972) 4-829-4420, Fax: (972) 4-829-5688.

This research was funded in part by the Max Wertheimer Minerva Center for Cognitive Studies which supported the visit of the second author to the Technion, as well as by the Israel Science Foundation (Grant No. 244/06).

Abstract

Previous modeling studies that examined the relative weight given to obtained and foregone payoffs (i.e., payoffs from the non-chosen options) in risky choices have yielded mixed results. Studies of prediction using the one step ahead method demonstrated that people give *less* weight to foregone than to obtained payoffs, while studies using simulations of multiple steps ahead suggested that on average people give *equal* weight to foregone and obtained outcomes. The current study employs new generalization techniques at the individual level to examine if the findings for the prediction of one step ahead method are due to over-fitting. Using the dataset of Rakow and Miler (2009), the results show that on the traditional fit index foregone payoffs are calculated as receiving lower weight. However, in generalization criteria at the individual level, the assumption of equal weighting of foregone and obtained payoffs performs best. This indicates that the superior fit of the model that assumes discounting of foregone payoffs is due to over-fitting to specific tasks.

Keyword: decision making; learning; counter-factual; cognitive modeling; reinforcement learning

Foregone payoffs are the outcomes from the alternatives that are not chosen. Exposure to foregone payoffs is common in social environments, where one often sees what would have occurred had another course of action been taken through observing the outcomes for others who selected an option that you declined to take. Understanding the role of foregone payoffs could therefore provide a key for modeling a variety of real-world behavioral phenomena, such as people's risk taking in naturalistic situations (e.g., Grosskopf, Erev, & Yechiam, 2007), the effect of regret (Zeelenberg & Pieters, 2004), and the effect of exposure to others' choices (Noble, Todd & Tuci, 2000; Yechiam, Druyan, & Ert, 2009). In repeated decisions, foregone payoffs have been shown to decrease risk aversion, especially for risky options where the most common outcomes are favorable (Yechiam & Busemeyer, 2005; 2006; Yechiam et al., 2009) and for multiple choice alternatives (Grosskopf et al., 2006; Ert & Erev, 2007); and they are considered to be one of the main mechanisms used to encourage participation in casino gambling (Grosskopf et al., 2006) and state lotteries (Zeelenberg, 1999). Foregone payoffs have also been shown to be involved in social learning in humans (see e.g., Yechiam et al., 2009) and animals (Hishimura, 1998; Noble, Todd & Tuci, 2000). In social comparisons, the saying "the grass is greener on the other side of the fence" suggests that foregone outcomes may sometimes be given greater attention. Therefore, at least some of the time, foregone payoffs may receive more weight than obtained payoffs. In the experimental literature, there is however, controversy whether individuals give the same or different weights to obtained and foregone outcomes.

Some studies have concluded that foregone payoffs are discounted compared to obtained payoffs (Busemeyer & Myung, 1992; Camerer & Ho, 1999; Yechiam & Busemeyer, 2005). One possible explanation for this discounting process is that attention capacity is limited, and the decision maker focuses on the obtained payoff as the primary source of information, discounting other sources. Another possible reason is that direct experience may be more influential than experience realized as a counterfactual (“what if”) (Yechiam & Busemeyer, 2006). In contrast, a second line of research implies that foregone payoffs carry the same weight as obtained payoffs (Fudenberg & Levine, 1998; Grosskopf et al., 2006; Erev & Barron, 2005). This is consistent with the argument, often made in Economics, that decision makers seek to use all available information in making decisions (Fudenberg & Levine, 1998).

Both of these lines of inquiry have focused on the same type of tasks: repeated decision tasks known as decisions from experience. However, there is a major difference in their analysis method. The first line of studies, which imply that foregone payoffs are discounted (e.g., Camerer & Ho, 1999), used the prediction of one-trial-ahead choices. In this method a model is given the previous outcomes of the participants’ decisions as input, and predicts their next choice ahead on each trial. The second line of studies, which point to an equal effect of foregone and obtained payoffs (e.g., Erev & Barron, 2005), used simulations of *n* trials ahead. In this method the model is not given any information on the decision maker’s past choices but rather predicts their entire set of repeated choices (i.e., the proportion of choices for each option in aggregate).

One possible reason for the inconsistent findings concerning the effect of foregone payoffs is therefore the reliance on different methods. In an important critique of the reinforcement learning literature, Erev and Haruvy (2005) demonstrated that parameter estimation using the prediction and simulation methods can lead to parameter values that differ significantly from each other (between the two methods). They argued that the one trial ahead prediction method is disadvantageous because its reliance on actual past choices leads to over-fitting and diminishes its ability to predict behavior in different tasks (see also Mitropoulos, 2004). Partly in response to this critique, methods have been proposed to evaluate the success of models of one step ahead predictions based on generalizability criteria rather than the fit of prediction alone (Ahn, Busemeyer, Wagenmakers, & Stout, 2009; Yechiam & Busemeyer, 2008; Yechiam & Ert, 2007). These methods are based on a parameter-free generalization of the model prediction for different tasks performed by the same individual, and on the consistency of the parameters for each individual decision maker across tasks. Their essential advantage is that they allow the evaluation of models both at the intra- and inter-individual level (as would be elaborated below). The current paper applies this multi-criteria modeling evaluation approach for investigating the comparative weighting of foregone and obtained payoffs.

The dataset we used is reported in Rakow and Miler (2009)[1]. Participants performed a series of decision tasks consisting of 100 trials. On each trial, participants chose one of two options that returned points – either a 'win' (a positive amount) or a 'loss' (a negative amount) – according to pre-programmed probabilities. Participants were unaware of these payoff probabilities, but observed the outcome for both the selected and the unselected option after each choice (i.e., outcome feedback for obtained and foregone payoffs). A key feature distinguishing Rakow and Miler's (2009) data from similar investigations of decisions from experience (as reviewed in Erev & Barron, 2005) is that payoff probabilities for one of the options were non-stationary. For instance, in Experiment 1 (which comprised four separate tasks), after an initial period of several trials where payoff probabilities were static, the 'win' probability either increased or decreased (depending on the task) at a rate of .01 per trial for the next 40 trials – then remained static for the remaining trials. Payoffs were such that the better (i.e. higher expected value) option switched during this period of change: either because the better option deteriorated, or because the inferior option improved. Participants were sensitive to the change in the environment – the modal choice switched according to the relative payoffs. However, as in many probability learning or 'binary

[1] We focus on Study 1 in Rakow and Miler (2009), which had a relatively large number of trials, providing sufficient data-points per participant for modeling.

prediction' tasks, participants tended to alter their choices even in periods where the payoff probabilities were static (e.g., Goodnow, 1955; Peterson & Uhehla, 1965).

This task models decision making in 'slow-changing' dynamic environments where past outcomes provide only a partial indication of the probabilities associated with different outcomes, and recent outcomes are a better guide to the future than more distal ones. For instance, many innovations are associated with a 'learning curve' whereby the chance of success increases gradually as agents get to grips with new technology, skills or practices. Initial performance with the innovation may be only minimally indicative of current levels of success (Denrell & March, 2001). The data from Rakow and Miler (2009) are particularly conducive to considering the relative weight given to obtained and foregone payoffs, because options could either increase or decrease in value. If foregone payoffs are given less weight, participants should respond more slowly to changes in unselected options (usually increasing in value) than they do to changes in selected options (generally decreasing in value). Because participants performed several different tasks, the data are also conducive to testing models via the alternative evaluation criteria noted above; especially, in addition to model fit, the generalizability of model predictions and the consistency of estimated parameter values across tasks can be assessed (Yechiam & Busemeyer, 2008).

Learning in non-stationary environments has been explored in the early learning and reinforcement literature. For instance, responses 'extinguish' when payoff reinforcement is withdrawn, though extinction is more gradual when reinforcement prior to withdrawal is partial (e.g., Humphreys, 1939). The data from Rakow and Miler (2009) can be thought of as an examination of competing partial reinforcement schedules, where one schedule is allowed to change gradually (increasing or decreasing the reinforcement contingency). In the current study we examined the effect of foregone payoffs in these tasks, and, specifically, compared the two main assumptions concerning the effect of foregone payoffs noted above, which disagree on whether foregone payoffs are discounted compared to the obtained payoffs.

Experiment: Four Repeated Choice Tasks with Non-stationary Payoffs

Rakow and Miler's (2009) experiment comprised four decision problems, which were administered in a random order for each participant. The problems were presented via computer as two-button tasks consisting of 100 trials in which the participant had to pick one of two buttons in each trial. Both options were risky choices from which participants could win or lose points. The value of the "winning" or "losing" amount of points for a given option remains constant throughout the task – however, the probability of a win or loss could change. In every task, one option – the non-stationary (NS) option – changed over a period of 40 trials, whilst the other option – the stationary (S) option – remained unchanged.

The within-participant design enables the use of advanced model evaluation criteria of Individual Parameter Consistency – the internal validity of the estimated parameters in different tasks, and Generalizability at the Individual Level – the robustness of predictions from one task to other tasks performed by the same individual (Yechiam & Busemeyer, 2008). Note that Rakow and Miler (2009) used a simple Bush-Mosteller (1955) linear learning rule to predict the average learning curve of the participants in each condition (i.e., group-level modeling). This model predicts choice proportion based on the difference between the obtained and foregone payoff on each trial. It thus assumes equal weighting of foregone payoffs. The current paper compares different assumptions concerning foregone payoffs. Furthermore, it evaluates the models at the individual level, which allows understanding potentially heterogeneous individual processes and not only the population model, and avoids a variety of aggregation problems (Estes, 1956; Siegler, 1987).

Method

Participants

Forty unpaid volunteers (19 male) with a mean (SD) age of 21.4 (4.7) years were tested individually. Almost all participants were students at the University of Essex.

Task and Apparatus

The participants completed four order-controlled tasks, each consisting of 100 trials. In each task, they were presented with two options on screen, which were described as "money machines". Participants could win or lose points by choosing an option (by clicking on it with the mouse), and were required to select one of the two machines on each trial. Participants saw the outcome of their choices and were not given any initial information about the payoff outcome – making the task an example of "experience based decisions" (Erev & Barron, 2005). Participants were also provided with foregone payoffs on each trial (i.e., they saw the outcome for the unselected option) as per Yechiam and Busemeyer (2006).

In Tasks 1-3, the amount that could be won or lost was the same for each option – so the options were differentiated only by the probabilities of wins/losses. This means that one option stochastically dominated the other: the probability of a given prize was always greater with one option than with the other (though the dominated option changed in the course of the task). Descriptions of the task payoffs appear in Table 1. In Tasks 1 and 2 the payoffs from the initially superior option began to decline after 20 trials (Task 1), or 40 trials (Task 2). In Task 3 the payoffs from the initially inferior option began to improve after 20 trials. Task 4 was a variant of Task 3 where both the outcome values and the probabilities differed between the two options. This enabled contrasting situations where assessment of win/loss probabilities is sufficient to determine which option is better with a situation (Task 4) in which participants must weigh up both the amount to be won or lost and their respective

An additional feature of the experiment was that half of the participants received a summary of past outcomes (the ‘experience plus history condition’, which is contrasted with an ‘experience only’ condition). This had the effect of improving the initial identification of the better option, but sometimes slowed the rate of adaptation to change in the environment (for full details, see Rakow & Miler, 2009). In the current study we pooled the model estimates across these two conditions although we graphically present the results for both conditions (in Table 1).

Competing Models

We compared three different reinforcement learning models, varying in their assumption about the effect of foregone payoffs. Most reinforcement learning models employ three groups of assumptions: first, a utility function is used to represent the evaluation of outcomes experienced immediately after each choice; second, a learning rule is used to form an expectancy (or propensity) for each alternative, which summarizes the experience of all past utilities produced by each alternative; third, a choice rule selects the alternative based on the comparison of the expectancies (see Yechiam and Busemeyer, 2005). We focus, however, on a fourth component that addresses the effect of foregone payoffs.

1. Utility: Given the small nominal outcomes provided in the current studies, a linear utility model was assumed with weighted averaging of gains and losses (as in Yechiam and Busemeyer, 2005, 2008). The utility for trial t is denoted $u(t)$, and is calculated as follows:

$u(t) = W \cdot win(t) - (1- W) \cdot loss(t)$ (1)

The term *win*(*t*) is the amount of money won on trial *t*; the term *loss*(*t*) is the amount of money lost on trial *t*; *W* is a parameter that indicates the relative weight to gains versus losses, and its value was limited between 0 and 1.[2]

2. Updating of expectancies. A delta learning rule was used for updating the expectancies based on past outcomes (see Busemeyer & Myung, 1992; Sarin & Vahid, 1999). According to this learning rule, the expectancy $E_j(t)$ for each alternative *j* on each trial *t* is updated as follows:

$E_j(t) = E_j(t\text{-}1) + \phi \cdot [u_j(t) - E_j(t\text{-}1)] \cdot \delta_j(t)$ (2)

The variable $\delta_j(t)$ is a dummy variable which equals 1 if alternative *j* is chosen on trial *t,* and 0 otherwise. This implies that if alternative *j* is selected on trial *t,* then its expectancy changes in the direction of the prediction error given by $[u(t) - E_j(t\text{-}1)]$. The parameter ϕ is the learning rate parameter. It dictates how much of the expectancy is changed by the prediction error. Its value was set to $0 \leq \phi \leq 1$. In this range, recently experienced payoffs have larger effects on the current expectancy as compared to payoffs experienced in the more distant past. The delta learning rule has been shown to have better generalizability at the individual level compared to several alternative models (see e.g., Worthy, Maddox, & Markman, 2009; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008), and there are brain areas in the frontal cortex that exhibit activation patterns consistent with its implied

[2] The assumption of relative weight to gains and losses was checked by examining a model with equal weighting of gains and losses and no free parameter. This produced lower fits as well as poorer generalizability. For conciseness we do not provide the details of this analysis.

mechanism (see e.g., Schultz, Dayan, & Montague, 1997). These findings suggest that the rule can be used as a robust estimation for the updating of preferences based on new information.

3. Choice rule: The choice on each trial was determined by the expectancies for each alternative, using to a ratio-of-strength choice rule (Luce, 1959) formalized by the following equation:

$$\Pr[G_j(t+1)] = \frac{e^{\theta \cdot E_j(t)}}{\sum_k e^{\theta \cdot E_k(t)}} \tag{3}$$

The parameter θ controls the sensitivity of the choice probabilities to the expectancies. This rule allows a large variety of choice responses. Setting $\theta(t) = 0$ produces random guessing, while as $\theta(t) \rightarrow \infty$ a strict maximizing rule is produced (i.e., the decision maker *always* chooses the option with the higher expectancy). The probability of choosing the alternative producing the largest expectancy increases according to an S-shaped logistic function with a slope (near zero) that increases with θ. The value of the parameter θ was allowed to range between 0 and 250, enabling a wide range of exploration tendencies (see e.g., Yechiam & Ert, 2007).

4. Foregone payoffs. It is this fourth component that was manipulated to create three alternative models. Two major assumptions concerning the effect of foregone payoffs were contrasted. Under one assumption (e.g., Fudenberg & Levine, 1998; Grosskopf et al., 2006; Erev & Barron, 2005) foregone payoffs receive similar weight to obtained payoffs. In this model the Delta rule (see Formula 2) updates simultaneously the expectancies of all alternatives based on the payoffs in trial *t*, as follows:

$$E_j(t) = E_j(t\text{-}1) + \phi\cdot[u_j(t) - E_j(t\text{-}1)] \quad (4)$$

This model will be referred to as the F=O model (denoting the fact that the weight of foregone payoffs equals that of the obtained payoff).

An alternative view argues that foregone payoffs are discounted compared to obtained payoffs (Busemeyer & Myung, 1992; Camerer & Ho, 1999; Yechiam & Busemeyer, 2005). This model implies the following update to the Delta rule:

$$E_j(t)= E_j(t\text{-}1) + \phi\cdot\delta_j(t) \cdot[u_j(t) - E_j(t\text{-}1)] + \delta_j(t)] + \gamma\cdot\phi\cdot(1\text{-}\delta_j(t))\cdot[u_j(t) - E_j(t\text{-}1)] \quad (5)$$

The parameter γ denotes the weight assigned to payoff feedback from non-chosen options. A γ value of 0 implies no weight to foregone payoffs, and the model then reduces to the simple Delta rule; a value of 1 implies equal weighting for foregone and realized payoffs; values ranging between $0 < \gamma < 1$ imply some discounting of foregone payoffs compared to obtained ones. Combining all these three eventualities yields a parameter range of $0 \le \gamma \le 1$, which is the range that was used to allow the weak assumption that some individuals discount foregone payoffs. This model will be therefore referred to as the F≤O model.

We also examined a third approach, which suggests that while some individuals discount foregone payoffs others overweight them as compared to obtained payoff (Yechiam & Busmeyer, 2005). Under this assumption the value of γ is limited between $0 \le \gamma \le 5$, implying that foregone payoffs can be overweighted at a maximum ratio of up to five times compared to obtained payoffs. This final model will be referred to as the F≠O model.[3]

[3] This model also incorporate the possibility of the weight of obtained and foregone outcomes could be equal. However, it rests on the assumption that for most people this is not the case. Otherwise, it can be replaced by the more parsimonious F=O model.

Model selection

The models were initially evaluated for their ability to predict the 'one step ahead' choices on each trial. Specifically, the model parameters were estimated (on each trial *t*) based on the fit of the prediction for trial $t+1$ to the actual choice, using log likelihood estimation (Busemeyer & Wang, 2002).

The parameter optimization process was a robust combination of grid-search and simplex (Nelder & Mead, 1965) search 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 minimized the log likelihood for an individual. This generated a set of solutions, one for each starting point on the grid. Occasionally these solutions differed due to local maxima, and the grid point that produced the maximum over all starting positions was selected as the final solution. The parameter search was forced to satisfy the constraints on the parameter ranges noted above.

Additionally, the final fit of the learning models was compared to a baseline statistical model which assumes that the choices are generated by a statistical Bernoulli process (following Busemeyer & Stout, 2002). According to this model, for each participant the predicted choice probabilities for each trial are simply the average proportions of alternative selections throughout the task. The comparison to the benchmark model does not change the differences between the reinforcement learning models that are compared (because all log likelihoods are deducted by a constant) but rather provides a benchmark for evaluating the fit of all models. Positive values of the corrected log likelihood fit statistic indicate that a learning model performs better than the baseline model. The comparison is made using the G^2 statistic as follows (Busemeyer & Wang, 2000):

$$G^2 = 2\cdot[\,LL_{model} - LL_{baseline}] \qquad (6)$$

Finally, the fit of each model was corrected based on the Bayesian Information Criterion (BIC; Schwartz, 1978):

$$\text{BIC} = G^2 - k \cdot \ln(N) \qquad (7)$$

where k equals the difference between models in the number of parameters and N equals the number of observations, 100). For example, in our examination the F=O model had three parameters whereas the F≤O model has four. Thus, $1 \cdot \ln(100) \approx 4.6$. This constitutes the deduction from the G^2 of the four-parameter model.[4]

In addition to the fit index used for parameter estimation, generalization tests were conducted for each individual performer. In the test of Generalization at the Individual Level (GIL), the parameters in one task were used to form predictions for each of the other three tasks, and these predictions were compared with a random prediction.[5] This method provides an intra-individual analysis as it assesses the models ability to predict behavioral changes and consistencies within the individual. Additionally, we evaluated the Individual Parameter Consistency (IPC) of each model parameter by examining the associations between parameter values extracted in different tasks performed by the same person. This produces a score denoting the internal consistency of parameters estimated in different tasks, enabling the inter-individual assessment of a stronger assumption: that some of the estimated parameters reflect consistent mechanisms within the individual.

---

[4] Making this correction in the current dataset using the Aikake Information Criterion (AIC) does not change the results of any of the significance tests. The current findings are thus robust to the differences in the assumptions of alternative information criteria.

[5] Clearly, the statistical baseline model is of no use in this generalization test. One could just flip the alternatives of the first (source) task in a second (target) task to lead to a total failure of this model in a generalization test.

Results

Behavioral results

The main behavioral results are summarized in Table 1. The results showed that adaptation to the change in payoffs was more successful when an initially inferior (and relatively unchosen) option increased in EV than when an initially superior (and relatively chosen) option decreased in EV (e.g., in Task 3 compared to Task 1) (statistical analyses appear in Rakow & Miler, 2009). This can be interpreted as denoting greater weight to foregone compared to obtained payoffs. Alternatively, the same pattern can be due to the differences in the probability of the better outcomes in these tasks. Exploration and dynamic adaptation might have been facilitated by the relatively low outcomes as implied by Luce's (1959) rule (described above). The modeling analysis sheds more light on this issue.

Model fit and estimated parameters

The highest correct BIC was obtained by the F≠O model, 47.8 (±2.71) and a similar score was obtained by the F≤O model, 46.3 (±2.74). The fit of the F=O model was lower, at 32.4 (±3.34). The fits of all three models were significantly above that obtained by the statistical baseline model assuming no learning (F≠O: $t(159) = 17.67$, $p < .01$; F≤O: $t(159) = 16.86$, $p < .01$; F=O: $t(159) = 9.68$, $p < .01$). The two models assuming different weight to foregone and obtained payoffs (F≠O, F≤O) had no difference in fit, and both had significantly higher fit than the F=O model (F≠O: $t(159) = 7.66$, $p < .01$; F≤O: $t(159) = 6.88$, $p < .01$).

We also examined the fit at the individual level. A comparison of the F≠O model and the F=O showed that for 88% of the participants the F≠O provided a better fit. Comparing the F=O model and the F≤O showed that for 80% of the participants the F≤O provided a better fit. This shows that for most individuals, the assumption of an equal weight to foregone and obtained payoff resulted in poorer fit.

Finally, we also examined the foregone payoff parameter of the most successful model, the F≠O model. This parameter, to reiterate, dictates the weight of foregone compared to obtained payoffs, with values smaller than one denoting discounting of foregone payoffs. The results showed that in 74% of all cases the value of the parameters was smaller than 1, suggesting that most participants tended to discount foregone payoffs, compared to obtained ones. This, of course, is under the assumption that this model is the most accurate one, which is supported by the fit indices. We then moved to studying the success of the models in the generalization tests.

<u>Generalizability at the Individual Level (GIL)</u>

The GIL tests examined the ability to use the parameters estimated for each individual in a given task to predict his/her choices in the other tasks. The results of the tests appear in Table 2. As indicated in the table, the model with the best GIL fit was the F=O, even though it had the poorest fit in predicting the next step ahead on each individual task. Its fit, compared to the random baseline model, was 9.26, compared to -10.61 for the F≠O model and 0.97 for the F≤O model. In fact, the F=O model was the only model that showed a significant improvement beyond random prediction (F=O: $t(159) = 2.11$, $p = .04$). Its fit in the generalization test was significantly above that of the F≠O model ($t(160) = 3.25$, $p < .01$) though not above that of the F≤O model ($t(160) = 1.34$, $p = .18$).

A non-parametric analysis yielded a similar rank of models. Seventy-two percent of predictions were above the random model for the F=O model, 62% were above baseline for the F≠O model, and 67% for the F≤O model. Similarly, an examination at the individual level showed that for 65% of the participants, the fit of the F=O model was better than that of the F≠O model and for 55% it was better than the fit of the F≤O model. Notice that in these analyses the differences between models were smaller. What this pattern suggests is that there are large individual differences in people's weighting of foregone payoffs, but that the models' assumptions of overweighting or discounting foregone payoffs leads to larger prediction errors than the assumption of no consistent weighting to these payoffs across the four tasks.

Finally, the analysis also allows comparison of the ability of the four tasks to produce parameters that successfully predict individuals' performance. In this respect, it appears that the most successful task under all three models was Task 4, which is characterized by changes not only in the likelihood of payoffs but also in their magnitude. Under the F=O model, for instance, this task led to predictions that were better than those of the random model in 83% of all cases within the other three tasks.

Individual Parameter Consistency (IPC)

We examined parameter consistency by assessing the extent to which individual parameter values obtained from one task predict parameter values on the other three tasks. Cronbach's α was used as a measure of consistency for each set of four parameter values (see Table 3). Examination of Table 3 suggests that the only reasonable consistency appeared for the learning rate parameter (ϕ) under the F=O model (Alpha = 0.76) (according to Nunnally's 1978 criteria). The consistency of the learning rate parameter was somewhat harmed by the addition of the γ parameter in the F≤O and F≠O models. In contrast, the F≤O and F≠O models exhibited moderate IPC for the weight to gains/losses parameter (W), which was somewhat decreased by the omission of γ in the F=O parameter. IPC was poor for γ (the weight to foregone payoffs) in both models that include this parameter, and (with the slight exception of the F≤O model) the choice consistency parameter (θ) showed poor IPC in all models. In summary, the addition of a parameter for the weight given to foregone payoffs (γ) in the F≤O and F≠O models did little, if anything, to enhance overall parameter consistency.

The effect of task and condition, and the psychological plausibility of the models

The individual parameter values displayed a coherent pattern of relationships with two behavioral measures of task performance: the number of times the higher EV option was chosen per task (a measure of successful responding), and the number of alternations between options per task (the number of times that different options were chosen on successive trials). Using mean values (averaged within participants across the four tasks), the learning rate parameter (ϕ) was positively correlated with the number of alternations (r ≈ .6 to .7 depending upon the model). Moreover, for the F≤O and F≠O models, the consistency parameter had a strong negative relationship with the number of alternations (r ≈ –.8) – lower consistency being indicative of a greater tendency to 'flip-flop' between alternatives.

In addition, there was a quadratic relationship between ϕ and the frequency of choosing the better alternative, whereby the best performing participants had ϕ-values of .4 to .5. This quadratic component was significant for each model, and accounted for 14 to 28% of additional variance in performance (depending on the model). This accords with the principle that in a dynamic environment one must 'spread' attention between recent observations (in order to identify changes in the environment) and more distal ones (to learn the underlying payoffs). Too much reliance on past incomes makes one insensitive to changes in the environment (Rakow & Miler, 2009); too much reliance on recent outcomes makes one too prone to probability matching to the detriment of performance.

The effect of the 'history' manipulation was apparent in the values of the learning rate (recency) parameter (ϕ). For all models, ϕ was higher for the experience only condition – significantly so for the mean value of ϕ and, individually, for at least three of the four tasks for each model. Thus, according to the models, expectancies change more quickly in the absence of a summary of past outcomes. This is consistent with the behavioral data, which implied greater ‘stickiness’ in choices when past outcomes were summarized and presented to the participant.

Discussion

Erev and Haruvy’s (2005) critique challenges the modeling of a series of data using the “one step ahead” method (Yechiam & Busemeyer, 2008). This method of analysis has become popular for studying individual differences in tasks where there is ample data for each individual, enabling conclusions to be drawn at the individual level (e.g., Ahn et al., 2008; Busemeyer & Stout, 2002; Walsten, Pleskac, & Lejuez, 2005). Yet, Erev and Haruvy’s critique does not offer a reasonable alternative to modeling repeated performance data at the individual level. The simulation-based method they recommend is considered to have too few data-points to generate useful solutions at the individual level (Feltovich, 2000). The current paper suggests the value of a third approach – using one step ahead modeling but referring to parameter-free predictions across tasks as a means of evaluating the generalizability of the obtained parameters.

We focus on the issue of the weighting of foregone payoffs, which has generated competing conclusions from one-step head (e.g., Busemeyer & Myung, 1992; Camerer & Ho, 1999; Yechiam & Busemeyer, 2005) and simulation based modeling (e.g., Fudenberg & Levine, 1998; Grosskopf et al., 2006; Erev & Barron, 2005). Our main results show that, consistent with previous findings using the traditional one step ahead fit index, a model assuming lower weight to foregone (relative to obtained) payoffs is the most successful in fitting individuals’ data. However, the analysis using generalizability criteria at the individual level reveals completely different rankings: in this analysis, the additional parameter denoting the discounted weight of foregone payoffs *decreases* the ability to use this model for predicting choices in a different task performed by the same individual. Thus, the relative increase in fit due to the additional assumption that foregone payoffs are discounted appears to be the result of over-fitting of the model to specific tasks. Moreover, assuming differential weighting of foregone and obtained payoffs does not improve the ability to estimate parameters that are consistent within individual decision makers.

This modeling exercise suggests three main conclusions. First, it appears that the dependence on the fit index for a single task in the method of one step ahead modeling can produce inaccuracies that are due to over-fitting of parameters to the task. While an analysis of models that are relatively simple (such as a model assuming no differences in the weight foregone payoffs) can to a certain extent decrease the inaccuracies, a much better solution is to use multiple tasks. These should then be analyzed either using generalization criteria as in the current study, or by estimating the parameters simultaneously for the different tasks (e.g., using Bayesian frameworks).

A second conclusion regards the common use of information indices, such as the BIC. The goal of these indices is to compensate for the addition of parameters by penalizing the fit of the model. Here, we have seen that the BIC, which is considered a relatively orthodox information index, still did not sufficiently penalize the more complex model, which failed in the generalization test. The reason for this could be the assumption, made in information indices, of a linear effect of parameters. Clearly, though, when a parameter is used to modify inputs (such as the weight of foregone payoffs or positive versus negative outcomes), interactions between parameter values and specific input pattern can produce non-linear effects. Thus, for parameters that modulate input information the use of information criteria should be rethought.

Our last conclusion touches the specific feature investigated – the use of foregone payoffs. The findings are illuminating with respect to the ability of the human operators to process information even if he/she is not directly impacted by it. It appears that at least in the two choice setting we examined, decision makers integrate this information as well as they use information which impacts them directly.

References

Ahn, W. Y., Busemeyer, J. R., Wagenmakers, E. J., & Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cognitive Science*, *32*, 1376-1402.

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.

Busmeyer, J.R., & Stout, J.C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment*, *14*, 253-262.

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., & Mostellar, F. (1955). *Stochastic models for learning*. New York: Blackwell.

Camerer, C., & Ho, T. -H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, *67*, 827-874.

Denrell, J., & March, J. G. (2001). Adaptation as information restriction: The hot stove effect. *Organization Science*, *12*, 523-538.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies *Psychological Review*, *112*, 912-931.

Erev, I., & Haruvy, E. (2005). Generality, repetition, and the role of descriptive learning models. *Journal of Mathematical Psychology*, *49*, 357-371.

Ert, E., & Erev, I. (2007). Replicated alternatives and the role of confusion, chasing, and regret in decisions from experience. *Journal of Behavioral Decision Making*, *20*, 305-322.

Estes, W.K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, *53*, 134-140.

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.

Goodnow, J. J. (1955). Determinants of choice-distribution in two-choice situations. *American Journal of Psychology*, *68*, 106-116.

Grosskopf, B., Erev, I., & Yechiam, E. (2006). Foregone with the wind: Indirect payoff information and its implications for choice. *International Journal of Game Theory*, *34*, 285-302.

Hishimura, Y. (1998). Food choice in rats (Rattus norvegicus): The effect of exposure to a poisoned conspecific. *Japanese Psychological Research*, *50*, 172-177.

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

Luce, R. D. (1959). *Individual Choice Behavior*. New York: Wiley.

Mitropoulos, A. (2004). *Economic learning, experiments and the limits to information.* Cheltenham, UK: Edward Elgar Publishing

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308-313.

Noble, J., Todd, P.M., & Tuci, E. (2001). Explaining social learning of food preferences without aversions: An evolutionary simulation model of Norway rats. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *268*, 141-149.

Nunnally, J.C. (1978). *Psychometric Theory*. New York: McGraw-Hill.

Peterson, C. R., & Ulehla, Z. J. (1965). Sequential patterns and maximizing. *Journal of Experimental Psychology*, *69*, 1-4.

Rakow, T., & Miler, K. (2009). Doomed to repeat the successes of the past: History is best forgotten for repeated choices with non-stationary payoffs. *Memory and Cognition*, *37*, 985-1000.

Sarin, R., & Vahid, F. (1999). Payoff assessments without probabilities: A simple dynamic model of choice. *Games and Economic Behavior*, *28*, 294-309.

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *5*, 461-464.

Schultz, W., Dayan, P., & Montague, R. (1997). A neural substrate of prediction and reward. *Science*, *14*, 1593-1599.

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

Wallsten, T. W, Pleskac, T., & Lejuez, C. W. (2005). Modeling a sequential risk- taking task. *Psychological Review*, *112*, 862-880.

Worthy, D.A., Maddox, W.T., & Markman, A.B. (2008). Ratio and Difference Comparisons of Expected Reward in Decision Making Tasks. *Memory and Cognition*, *36*, 1460-1469

Yechiam. E., & Busemeyer, J.R. (2005). Comparison of basic assumptions embedded in learning models for experience based decision-making. *Psychonomic Bulletin and Review*, *12*, 387-402.

Yechiam, E., & Busemeyer, J.R. (2006). The effect of foregone payoffs on underweighting small probability events. *Journal of Behavioral Decision Making*, *19*, 1-16.

Yechiam, E., & Busemeyer, J.R. (2008). Evaluating generalizability and parameter consistency in learning models. *Games and Economic Behavior*, *63*, 370-394.

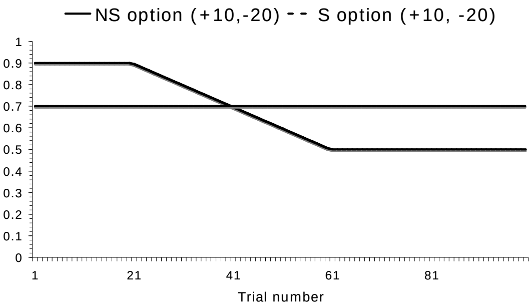
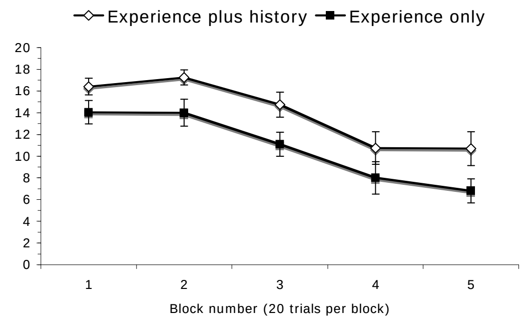
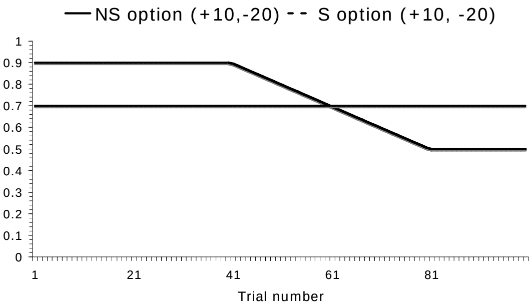
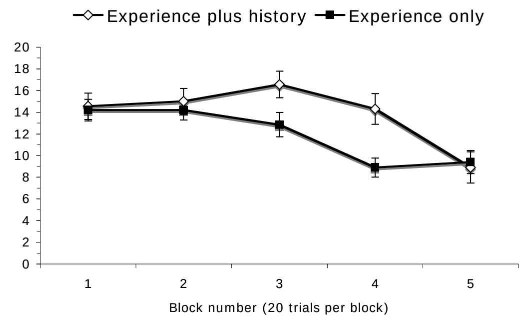
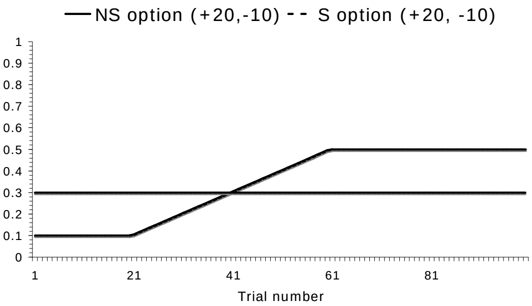
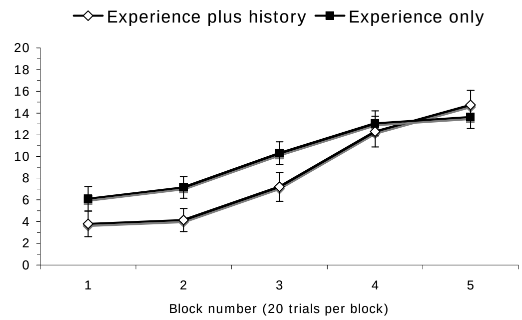
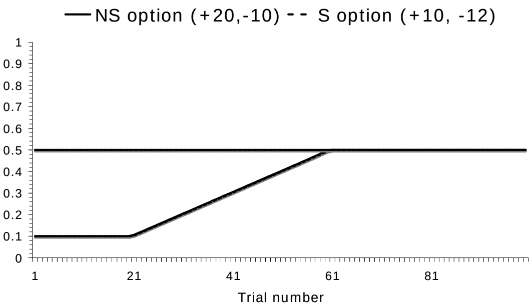
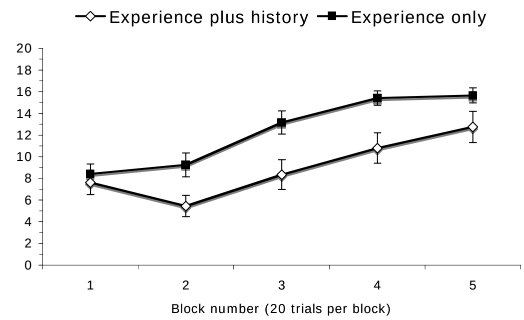
Yechiam, E., Druyan, M., & Ert, E. (2008). Observing others' behavior and risk taking in decisions from experience. *Judgment and Decision Making*, *3*, 493-500.

Yechiam, E., & Ert, E. (2007). Evaluating the reliance on past choices in adaptive learning models. *Journal of Mathematical Psychology*, *51*, 75-84.

Zeelenberg. M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making*, *12*, 93-106.

Zeelenberg, M., & Pieters, R.G.M. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes*, *93*, 155-168.

Table 1: Task payoffs and behavioral data from Rakow and Miler (2009).

<table>
<tr><th>Task</th><th>Payoffs for the non-stationary (NS) and stationary (S) options: Probability of the better outcome by trial number</th><th>Number of times per 20-trial block that the NS option is picked for each condition.</th></tr>
<tr><td>1</td><td>
</td><td>
</td></tr>
<tr><td>2</td><td>
</td><td>
</td></tr>
<tr><td>3</td><td>
</td><td>
</td></tr>
<tr><td>4</td><td>
</td><td>
</td></tr>
</table>

Error bars are ± 1 Standard Error

Table 2: Generalization at the Individual Level (GIL) test results: Mean G^2 scores and percent of individuals for which the model's prediction was better than a random prediction (in parenthesis). Three models are compared: F=O, F≤O, and F≠O (where F and O denote the possible weights of Foregone compared to Obtained payoffs). Shaded cells denote tests that do not involve generalization.

Model	Target task	Task used for estimating the parameters			
		Task 1	Task 2	Task 3	Task 4
F=O	Task 1	51.7 (95.0%)	23.1 (72.5%)	13.9 (70.0%)	15.8 (82.5%)
	Task 2	4.8 (72.5%)	54.1 (97.5%)	14.8 (77.5%)	29.0 (87.5%)
	Task 3	16.5 (77.5%)	24.5 (77.5%)	56.8 (97.5%)	22.5 (80.0%)
	Task 4	-33.3 (50.0%)	-0.9 (60.0%)	-19. 4 (55.0%)	43.1 (100%)
	Average*	-4.0 (66.7%)	15.6 (70.0%)	3.1 (67.5%)	22.4 (83.3%)
		Task 1	Task 2	Task 3	Task 4
F≤O	Task 1	81.0 (100%)	35.1 (70.0%)	29.6 (75.0%)	14.9 (77.5%)
	Task 2	23.8 (77.5%)	83.1 (97.5%)	24.4 (75.0%)	20.4 (80.0%)
	Task 3	-43.8 (55.0%)	-12.9 (47.5%)	65.3 (100%)	-0.9 (75.0%)
	Task 4	-36.4 (62.5%)	-31.2 (52.5%)	-11. 2 (60.0%)	52.3 (100%)
	Average*	-18.8 (65.0%)	-3.0 (56.7%)	14.2 (70.0%)	11.5 (77.5%)
		Task 1	Task 2	Task 3	Task 4
F≠O	Task 1	80.3 (97.5%)	5.2 (70.0%)	2.0 (64.1%)	15.5 (75.0%)
	Task 2	21.1 (75.0%)	83.2 (97.5%)	-8.8 (61.5%)	6.4 (67.5%)
	Task 3	-45.3 (52.5%)	-20.0 (47.5%)	67.1 (97.5%)	0.2 (72.5%)
	Task 4	-30.6 (60.0%)	-46.5 (50.0%)	-26. 4 (45.0%)	51.9 (95.0%)
	Average*	-18.3 (62.5%)	-20.4 (55.8%)	-11.1 (56.9%)	7.3 (71.7%)

* = Average for the generalization tests only

Table 3: Indicators of individual parameter consistency (IPC) by parameter and model: Cronbach alpha for each condition.

Model	Parameter			
	Learning rate (ϕ)	Gain/Loss Wt (W)	Consistency (θ)	Wt to Foregone (γ)
F=O	0.76	-0.06	0.43	
F≤O	0.51	0.36	0.46	-0.11
F≠O	0.56	0.44	0.20	0.13