Evaluating generalizability and parameter consistency in learning models

By Eldad Yechiam

Technion – Israel Institute of Technology

Jerome R. Busemeyer

Indiana University, Bloomington IN

Mail should be addressed to the first author:

Eldad Yechiam, PhD

Email: yeldad@tx.technion.ac.il

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.

Abstract:

The present study proposes a new evaluation method for learning models used for predicting decisions based on experience. The method is based on the generalization of models' predictions at the individual level. First, it evaluates the ability to make a priori predictions for decisions on a new task using parameters previously estimated from a different task performed by an individual decision maker. In addition, it evaluates the consistency of parameters estimated in different tasks performed by the same person. We use this method to examine two different rules for updating past experience with new payoff feedback: The Delta rule, where only the chosen option is updated; and a Decay-Reinforcement rule, where the chosen option is updated and non-chosen options are discounted by some amount. The results reveal an accuracy advantage to the Decay-Reinforcement model when fit to the same data. However, this model is outperformed by the Delta model in the test of generality and parameter consistency at the individual level. The current method thus improves the ability to select models based on their correspondence to consistent characteristics within the individual decision-maker.

1. Introduction

Recently, there has been a rising interest in learning models that are applied to choices from repeated play games. Studies of choice behavior in individual (see e.g., Busemeyer and Myung, 1992; Erev and Barron, 2005; Sarin and Vahid, 1999) and multi-player games (see e.g., Camerer and Ho, 1999; Cheung and Friedman, 1997; Erev and Rapoport, 1998, Erev and Roth, 1998; Fudenberg and Levine, 1995; Sarin and Vahid, 2001; Stahl, 1996) have shown that learning in repeated choice problems can be summarized by simple mathematical models. Two important issues have surfaced from this research.

First, methods for testing learning models have become a major concern. It is difficult to derive qualitative properties that distinguish the models unambiguously, and so researchers must resort to comparisons of the accuracy of model predictions . However, model comparisons must also take into consideration model complexity (Myung, 2000). A particular model might fit a data set better simply because more parameters were used to fit the same data. For this reason generalization tests provide an important model comparison method – model parameters are estimated from one learning condition, and then these same parameters are used to make a priori predictions for a new learning condition.

In the past, these generalization tests have been conducted between different groups or populations of decision makers (see e.g., Rieskamp et al., 2003). An important problem arises when using group data to perform generalization tests. Parameters differ across individuals, and therefore the predictions for groups must reflect the effects of parameter heterogeneity. Much stronger tests are possible by conducting generalization

tests across tasks within the very same person. This allows one to estimate parameters for a single individual on one learning task, and then examine how well these same parameters predict performance for the same individual on another learning task. This method has rarely ever been used to compare learning models.

Second, assessing the validity of model parameters is another major concern. The parameters are usually estimated by maximizing a fit statistic, which then raises the following question. Do these parameters actually measure stable characteristics of an individual, or do they simply reflect model mimicry for a particular task? If the former is true, then the parameters are stable across tasks and measure something meaningful about a person; but if the latter is true then the parameters lack stability and have no meaningful interpretation at the individual level. Having the same person perform several different learning tasks allows us to assess the validity of model parameters. We can estimate the parameters separately using data collected from the same person on each task, and examine the extent to which the estimated parameters remain invariant across tasks within a single person. This property of a model is hereafter called "*individual parameter consistency*".

Parameter consistency has only been partially addressed before, owing to the lack of proper methodology. Previous studies examined the consistency of parameters across conditions when the parameters were estimated from data averaged across individuals within a group (thus assuming homogeneity of parameters across individuals). Yet as far as the authors know, no previous study has assessed the consistency of parameters at the individual level.

Assessing the accuracy of each individual's parameters is particularly important for applications of learning models to the study of cognitive processes (see e.g., Busemeyer and Stout, 2002; Wallsten et al., 2005; Yechiam et al., 2005). In these applications, the model parameters provide measurements of latent cognitive and decision processes, such as utility coefficients or learning rates. However, the examination of individual parameter consistency is an important step for model testing and comparison in general, for two reasons. First, it can be used to improve the understanding of which components of the model should be fixed and which should be changed under different tasks and contexts (i.e., the parameters that are consistent across different tasks within the individual performer can be kept fixed, while the parameters that are not consistent should be re-estimated). A second benefit is that studying the consistency of parameters within the individual can identify models whose accurate prediction is due to flexibility and success in mimicking data rather than providing accurate representation of "stable" internal processes.

The present investigation uses the proposed method to compare two classes of learning models and evaluate the parameter consistency of each model. These two classes of models were previously evaluated in a study of the Iowa Gambling task, a complex task used for examining individual differences in choice behavior (Yechiam and Busemeyer, 2005). In the current studies we examine three tasks that comprise some of the components of the more complex Iowa Gambling task. Study 1 evaluates the two learning models in three tasks where the outcomes are, for the most part, in the gain domain. Study 2 evaluates the same learning models for three tasks in which the

outcomes are mostly in the loss domain. First we present the two classes of learning models.

2. A comparison of learning models

An examination of the learning models used in previous studies 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 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). In the present study we compare two learning models that posit different assumptions about the process of updating the expectancies.

2.1. Utility.

The evaluation of gains and losses experienced after making a choice is represented by a prospect theory type of utility function (Kahneman and Tversky, 1979). The utility is denoted $u(t)$, and is calculated as a weighted average of gains and losses produced by the chosen alternative in trial t.

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

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 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. In the present study, it was assumed that $L = 1- W$ (see Yechiam, Veinott, et al., 2007). Thus, a single parameter W denoted the relative attention weight to gains over losses. Furthermore, for the small amounts of money used in the present experiment (less than $1 per outcome), $\gamma = 1$ was considered to be a sufficiently good approximation to the utility function.

2.2. Updating of expectancies

Two general classes of models have been proposed to account for the way new information is accumulated in repeated choices (Yechiam and Busemeyer, 2005). Under one class of models, the expectancy for an alternative changes only if the alternative is selected. This class of models has been labeled "interference" models, because the memory representation is only modified by relevant events, and not simply as a function of time (e.g., Newell, 1992; Oberauer and Kliegl, 2001). In a second class of models, the expectancy of an alternative can diminish on each choice trial even if no new information concerning that particular alternative is presented. This class of models has been labeled "decay" models, because decay of memory can take place purely as a function of time even without the occurrence of interfering events (e.g., Atkinson and Shiffrin, 1968; Broadbent, 1958).

A Delta learning rule was used as an example of an interference class model. This model was found to have the best fit among interference class models in two previous studies that evaluated models at the individual level (see e.g., Busemeyer and Stout, 2002; Yechiam and Busemeyer, 2005). A Decay-Reinforcement model (Erev and Roth,

1998) was studied as an example of the decay class. This model was found to have the best fit among all models in our previous investigation (Yechiam and Busmeyer, 2005).

Delta model. Connectionist theories of learning usually employ a learning rule called the Delta learning rule (see Gluck and Bower, 1988; Rumelhart and McClelland, 1986; Sutton and Barto, 1998). It has been applied to learning in decision tasks by Busemeyer and Myung (1992) and by Sarin and Vahid (1999). According to this learning rule, the expectancy $E_j(t)$ for each alternative j on each trial t is updated as follows:

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

The variable $\delta_j(t)$ is a dummy variable which equals 1 if alternative j is chosen on trial t, and 0 otherwise. If alternative j is not chosen on trial t, then the new expectancy $E_j(t)$ simply remains the same as it was on the previous trial, $E_j(t\text{-}1)$. On the other hand, if alternative j is selected on trial t, then its expectancy changes in the direction of the prediction error given by $[u(t) - E_j(t)]$.

The parameter ϕ is the learning rate parameter. It dictates how much of the expectancy is changed by the prediction error. If $0 < \phi < 1$, then the effect of a payoff on the expectancy for an alternative decreases exponentially as a function of the number of times a particular alternative was chosen. Thus, recently experienced payoffs have larger effects on the current expectancy as compared to payoffs that were experienced in the more distant past.

Decay-Reinforcement Rule. More recently, Erev and Roth (1998) added a decay parameter to the reinforcement-learning model, which can be represented by the following equation:

$$E_j(t) = \phi \cdot E_j(t\text{-}1) + \delta_j(t) \cdot u(t) \tag{3}$$

Note that for this model, the past expectancy is always discounted, regardless of whether or not an alternative is chosen and new payoff information is experienced. This is implemented by multiplying the past expectancies of all alternatives $E_j(t\text{-}1)$ by the recency parameter ϕ in each trial, where $0 < \phi < 1$. The decay model can be more flexible because the expectancies of selected and unselected alternatives can both be changed. Previously, this model was found to have more accurate predictions than the Delta model (Yechiam and Busemeyer, 2005). Yet it is important to evaluate whether this improved accuracy is a result of more flexibility for the Decay-Reinforcement rule as compared to the Delta learning model. High model flexibility can lead to overfitting the choices made by an individual in a specific task.

We also implemented a second version of the Decay-Reinforcement rule in which the parameter ϕ is constrained so as not to affect the long-run expectation $E_j(t)$ as follows:

$$E_j(t) = \phi \cdot E_j(t\text{-}1) + \delta_j(t) \cdot u(t) \cdot (1\text{-} \phi) \tag{4}$$

This model, labeled Constrained Decay-Reinforcement, was constructed to avoid an embedded association between the recency parameter ϕ and the accumulated

expectancies, as the size of the expectancies is not a strict function of ϕ [1]. It therefore

more clearly sets the function of the recency parameter in relation to the expectancies.

2.3. Choice rule

The choice on each trial was determined by the expectancies for each alternative,

using to a ratio-of-strength choice rule (see Yechiam and Busemeyer, 2005). The ratio

rule assumes that the choice made on each trial is a probabilistic function of the relative

expectancies of the alternatives (Luce, 1959). It can be formalized by the following ratio-

of strengths rule:

$$\Pr[G_j(t+1)] = \frac{e^{\theta(t) \cdot E_j(t)}}{\sum_k e^{\theta(t) \cdot E_k(t)}} \tag{5}$$

The parameter $\theta(t)$ controls the sensitivity of the choice probabilities to the expectancies.

On one hand, setting $\theta(t) = 0$ produces random guessing; on the other hand, as $\theta(t) \to \infty$

we recover a strict maximizing rule. 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 $\theta(t)$.

It is assumed that the sensitivity to the expectancies, denoted by $\theta(t)$, may change

as a function of experience. This is parameterized by a power function for the sensitivity

change over trials:

---

[1] In the original Decay-Reinforcement model as $t \to \infty$, $E \to u/(1-\alpha)$. In the Constrained Decay-Reinforcement model as $t \to \infty$, $E \to u$.

$$\theta(t) = (t/10)^c \tag{6}$$

Here c is a free parameter. The parameter c determines how the consistency of the choice

and expectancies change across training. When the value of c is negative, choices become

more inconsistent with training (perhaps because of boredom or fatigue); if the value of c

is positive, then choices become more consistent with training (perhaps because of

increased experience with the task)[2].

2.4. Model evaluation

All of the model evaluations are based on a measure computed from the accuracy

of 'one step ahead' predictions generated by each model for each individual performer.

Specifically, define $Y(t)$ as a $t \times 1$ vector, representing the sequence of choices made by

an individual up to and including trial t; define $X(t)$ as the corresponding sequence of

payoffs produced by these choices; and define $\Pr[G_j(t)]$ as the probability that alternative

j will be selected on trial t by the model. Each model is given $X(t)$ and uses this

information to generate $\Pr[G_j(t+1) \mid X(t)]$ for choice trials $t = 1$ to 100 and alternatives j

$= 1$ to 2. The accuracy of these predictions is measured using the log likelihood criterion

for each individual:

---

[2] This model assumes that the slope of the choice sensitivity is the consistent trait of individual subjects. It sets up a point after an initial exploration period (the tenth trial) and compares the exploration before and after that point. We also contrasted it with a model that assumes independent magnitude and slope, as follows: $\theta(t) = \theta_0 \cdot (t/10)^c$, where θ_0 is an additional free parameter denoting the magnitude of the choice sensitivity ($0 \le \theta_0 \le 700$). However, the latter model did not improve the fit or the individual parameter consistency and generality. Therefore, for conciseness, it is not detailed here.

$$LL_{model} = lnL \text{ (model | data)} = \sum_t \sum_j ln(Pr[G_j(t+1) \mid X(t)]) \cdot \delta_j(t+1) \qquad (7)$$

Recall that $\delta_j(t) = 1$ if alternative j is chosen on trial t, and zero otherwise.

Each learning model has only three parameters $\{W, \phi, c\}$ estimated from each person's choices on the first 100 trials. When fitting parameters, we optimize the log likelihood for each participant and model using a robust combination of grid-search and simplex (Nelder and Mead, 1965) search methods. Each point on the grid serves as a starting position for the simplex search algorithm, which is then used to find the parameters that maximize the log likelihood for an individual. This generates a set of solutions, one for each starting point on the grid. Occasionally these solutions differ due to local maxima, and so we select the grid point that produces the maximum over all starting positions for the final solution.

The parameter search is forced to satisfy the following constraints. The value of W is limited between 0, denoting complete attention to losses, and 1, denoting complete attention to gains. The value of ϕ is limited between 0 and 1. Values lower than 0 for ϕ allow for gambler's fallacy and values higher than 1 allow for a primacy effect. These values produce predictions that are qualitatively different from that of recency. For parsimony, we focus on the range of the learning parameter that produces recency thus preventing the other response strategies to past expectancies[3]. The value of the sensitivity parameter c is set between -5 and 5, permitting the full range between deterministic and

---

[3] One way of measuring primacy and gambler's fallacy is to conduct a separate evaluation and constrain ϕ to be below zero or above 1, accordingly. This examination was considered to be beyond the scope of the current study (but see Yechiam and Busemeyer, 2006).

random choices (approximately 90% of the participants conform to this range in the Delta model and 95% in the Decay- Reinforcement model).

For the fit index, we compare the learning models presented above to a baseline statistical model. The baseline model assumes that the choices are generated by a statistical Bernoulli process. That is, the choice probabilities for each alternative are assumed to be constant and statistically independent across trials:

$$\Pr[\ G_j(t{+}1)\ |\ Y(t)] = p_j \qquad\qquad (8)$$

The baseline model has only a single parameter (p_1, $p_2 = 1 - p_1$) which corresponds to the proportion of choices from Alternative H, pooled across all of the trials (the proportion for L is determined from the proportion for H). The different models are evaluated by comparing a log likelihood score for the baseline and the learning model:

$$G^2 = 2 \cdot [\ LL_{model} - LL_{baseline}] \qquad\qquad (9)$$

Because the baseline model has only one parameter, while the learning models have three parameters, when we compare model fits, we adjust for the difference in number of parameters. This is achieved by using the Bayesian Information Criterion (BIC; Schwartz, 1978) statistic to compare models. The BIC is a model-comparison index based on Bayesian principles which penalizes models with additional parameters:

$$BIC = G^2 - k \cdot \ln(N) \tag{10}$$

where k equals the difference in the number of parameters and N equals the number of observations. For our comparisons, we have $k = 2$ (two additional parameters in the learning models compared to the baseline model) and $N = 100$. Thus, $2 \cdot \ln(100) \approx 9.2$. This constitutes 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.

In addition to the fit index, generalization tests were conducted for each individual performer. In the test of *generalization at the individual level*, the parameters in one task are used to form predictions for another task. The model's predictions are compared to a random prediction using the G^2 index. Clearly, the statistical baseline model is of no use in this generalization test, because its predictions only reflect the measured choice proportions in a given task.

Finally, we evaluated *individual parameter consistency* by examining the associations between parameter values extracted in different tasks performed by the same individual. If the parameters estimated in one choice task are consistent within individual performers, then their ranking across individuals in different choice tasks is expected to be similar. For example, if the parameter W, denoting the weighting to gains, has high individual parameter consistency, then the same individuals exhibiting high attention to gains in one task should have high attention to gain in other tasks, resulting in a high positive correlation between the parameters estimated in the different tasks. A correlation of zero implies no individual parameter consistency, and a negative correlation indicates a tendency to be consistent in the opposite direction in a given parameter across tasks.

3. Experimental setup

3.1. General layout

In two experimental studies participants performed a set of three distinct and independent choice tasks. In all three tasks the probabilities and payoffs were initially unknown and were learned by repeatedly choosing alternatives and obtaining immediate payoff feedback. Participants were not made aware that that the distributions were fixed; and this too was to be learned from experience (as in Barron and Erev, 2003). The alternatives were presented on the screen as two buttons, and choices were made by selecting one button on each trial. Each button was associated with a fixed payoff distribution.

The payoffs for the two studies are summarized in Table 1 and fully described in the methods section. Task 1 (Payoff-Sensitivity task) included two alternatives with different averages and the same variance. Task 2 (Small-Probability task) included an alternative with a small probability outcome and an alternative with a constant outcome. Task 3 (High-Variance task) included an alternative with high average and variance, and an alternative with low average and variance. The tasks were therefore considerably different in their payoff structure.

These three tasks were selected because they comprise the components of more complex tasks used for studying individual differences in decision making, such as the Iowa Gambling task (Bechara et al., 1994) and others (e.g., Lejuez et al., 2003; Newman et al., 1987). First, the tasks include choice alternatives that produce both gains and losses. This property is considered to be important because individual differences in the relative weight of rewards and penalties are central to most theories of cognitive style and

personality (see e.g., Gray, 1994; Higgins, 1997; Hjelle and Ziegler, 1981; Fowles, 1988). Secondly, the Iowa Gambling task in particular involves choosing between payoffs of different magnitude, dealing with differences in variance, and responding to small probabilities. Accordingly, the present tasks were considered as a good starting point for discovering whether the current method of model evaluation would reveal important characteristics of the studied models that are also relavant to the more complex tasks currently in use.

Participants' choices in these three tasks were used to systematically compare the two competing learning models with different rules for updating the expectancy of the alternatives. The first goal of the two studies was to compare the decay and interference class models by means of the new method that combines generalization at the individual level and individual parameter consistency. A secondary goal was to evaluate the three choice tasks. Namely, it is theoretically possible, and even likely, that some tasks would be better for extracting more robust parameters. This should be reflected in better generalizability from one task to another at the individual level, as well as by greater individual parameter consistency.

3.2. Method

One hundred and eighty students from Indiana University, Bloomington campus (90 males and 90 females), participated in the two studies (90 students in each study). The students' average age was 20, ranging from 18 to 30. In Study 1 participants were paid a sum of $15 to $35 ($20 on average) for their participation, and in Study 2 they

were paid a sum of $5 to $23 ($14 on average). The exact amounts depended on the

participants' performance in the experimental task.

The experiment took place in the Experimental Spatial Lab at Indiana University.

Participants filled informed consent forms prior to participating in the study. Participants

were asked to read the instructions, which were also read out loud. They were encouraged

to ask questions. The instructions read as follows:

"Your payoff in this experiment will be ___ ($8 in Study 1 and $30 in study 2)

plus your gains/losses during the experiment. Gains/losses will be accumulated during

200 trials. In each trial you will have to click a button. The payoff for your selection will

appear on the button that you selected. You receive 1 cent for every 10 points earned.

"You will immediately see a form with two buttons like the one in the picture

below (a picture of the form was shown at this point; see Figure 1). You can press either

of the two buttons in the form. The payoff for choosing a button appears on the respective

button. The accumulating payoff appears below. At the end of 200 trials you will see a

message to call the experimenter. When you get the message, please raise your hand.

After you finish the task there will be two other similar tasks but the payoffs for pressing

buttons might be different from the payoffs in the first task".

Altogether, the set of three tasks took about 40 minutes to complete. Subsequent

to performing the tasks participants filled a short questionnaire for another study

(Yechiam and Budescu, 2006). After filling out the questionnaire participants were paid,

thanked, and dismissed.

The three experimental tasks involved selecting one of two buttons, labeled A and

B, in each of the 200 trials (see Figure 1). The number of trials was unknown to the

players. The color of the buttons was different for each of the three tasks to emphasize their distinctiveness. Payoffs were presented in points (rounded to the next integer) and converted into money at the end of the experiment. In Study 1 payoffs were contingent upon the button chosen as follows:

Task 1 - Payoff-Sensitivity (PS). In this task, the payoff for Alternative H (denoting High expected value) is drawn from a discrete approximation to a normal distribution with an average of 20 and a standard deviation of 20. The payoff for alternative L (denoting Low expected value) is drawn from a discrete approximation to a normal distribution with an average of 10 and the same standard deviation. In addition, payoffs are rounded to 0, ±10, ±20, or ±30 around the mean. Accordingly, outcomes from H are superior about 70% of the time.

Task 2 – Small Probability (SP). In this task, alternative H produces 10 with certainty. Alternative L produces 20 with a probability of 0.9, and a negative value with a probability of 0.1, where the negative value is uniformly distributed within the range 85-95. This produces an expected value equal to 9.17 for alternative L. The payoffs from L thus have high variance (SD = 32.7; compared to 0 in H), but are positively skewed, so that L produces high positive outcomes 90% of the time.

Task 3 –High Variance (HV). In this task, the payoff for H is drawn from a normal distribution with an average of 100 and a standard deviation of 354. The payoff for L is drawn from a normal distribution with an average of 25 and a standard deviation of 17.7. The distribution of L is further truncated above 0. The variance of the payoffs from H is therefore very high, producing superior payoffs only 58% of the time. This task was originally used by Thaler, Tversky, Kahneman, and Schwartz (1997). Their results

indicated that, on average, decision makers preferred alternative L over H, indicating a strong variance aversion, even at the expense of losing about two thirds of the potential payoffs (see replication by Barron and Erev, 2003).

In study 2, the payoffs were identical, but the signs of all payoffs were reversed (see Table 1). The location of the H and L options was randomized for each participant and for each task. There was a delay of 1 second after pressing a button in which the two buttons could not be pressed. Two types of feedback immediately followed each choice: (1) The basic payoff for the choice, which appeared on the selected button for a duration of one second and then re-appeared on the caption below the buttons, (2) an accumulating basic payoff counter, which was displayed constantly.

Participants were randomly assigned to one of six groups. Each group performed the three tasks in a different order. An equal number of participants were assigned to each of the groups; also there was an equal proportion of males and females in each group (except for two groups in which there was either one more male or one more female).

4. Experimental results of Study 1 (Gain Domain)

4.1. Behavioral Results

Figure 2 summarizes the average choice proportion from the High expected value alternative in each task. The results in the Payoff-Sensitivity task show a slow but significant effect of experience. Choices from H increased significantly from the first to the last block of 10 trials ($t (89) = 6.35$, $p < .01$). In the other two tasks, performers did not learn to select the high expected-value alternative. In the Small-Probability task H was chosen 49% of the time, and there was no significant effect of experience. In the

High-Variance task, consistent with previous findings, participants learned to avoid the

High-Variance alternative (H) despite its much higher expected value. In this task,

choices from H decreased significantly from the first to the last block of 10 trials (t (89) =

2.74, p < .01). Thus, the prescriptive predictions of expected value clearly do not drive

the participants' behavior in all of the current tasks.

We first examined the consistency of the participants' choices across tasks at the

individual level using Spearman rho. There was no association between choice of H in

the Payoff-Sensitivity task and choice of H in each of the other two tasks (Small-

Probability: r = 0.10, NS; High Variance: r = 0.03, NS). We also calculated the

association between choices from the high variance alternatives in the Small-Probability

task (L) and the High-Variance task (H). The Spearman correlation was small but

significant for the average choice proportion in 200 trials (r = 0.37, p < .01) as well as in

the first 100 trial block (r = 0.21, p < .05). This indicates that there is some consistency in

the choice of the high variance (risky) alternative. Learning models can shed light on the

component processes that modulate this association.

4.2. Model evaluation

Our comparison of learning models focused on the first 100 trials, because in the

second half of the task many participants reached a plateau, repeatedly selecting from the

same alternative.

Model fit. The average BICs of the learning models, indicating their accuracy

compared to the baseline model, are summarized in Table 2. The average BIC of the

Decay-Reinforcement was higher than the BIC of the Delta model (t (269) = 6.09, p <

.01), with the Constrained Decay-Reinforcement model falling in between but still significantly lower than the Decay-Reinforcement model (t (269) = 4.03, p < .01). An examination at the individual level confirms these results. Under the Delta model only 42% of the participants had BICs above the baseline level, compared to 55% under the Decay-Reinforcement model, and 48% under and the Constrained Decay-Reinforcement model.

Focusing on the Decay-Reinforcement model, an examination of the fit in specific tasks showed an advantage of this model over the Delta model in the Payoff-Sensitivity task (mean BIC of 13.7 compared to 8.4; t(89) = 6.28, p < .01) and in the Small-Probability task (mean BIC of 11.4 compared to 2.7; t(89) = 7.23, p < .01). In the High-Variance task, the BICs of the models were similar, with both models falling below the baseline model (mean BICs of -1.9 and -2.8, respectively). Thus, overall the Decay-Reinforcement model had better fit for the data used for parameter estimation.

A comparison of different tasks indicated that the Payoff-Sensitivity task produced the best fit (with a BIC of 11.3, averaged across all three models), and that the High-Variance task produced the poorest fit (BIC of –2.8). The fit of the Small-Probability task fell in between (BIC of 7.2). The difference in fit between tasks was significant for all three models (Delta: (F(2, 267) = 6.97, p < .01, MS_E = 342.1), Decay-Reinforcement: F(2, 267) = 17.53, p < .01, MS_E = 408.1, Constrained Decay-Reinforcement: F(2, 267) = 13.02, p < .01, MS_E = 437.7).

Generalization at the Individual Level. The parameters estimated in each of the three tasks were used to generate predictions for the other two tasks. This was conducted separately for each individual. As indicated above, in this comparison the fit index is G^2

because the baseline model is a random draw. Another way to present the results is by the percent of individuals for whom $G^2 > 0$. A percent above 50 implies above-chance success in predicting the next choice ahead in the generalization test.

The results (summarized in Table 3) show an unexpected interaction between experimental task and the generalization success of the models. The generalization from the Payoff-Sensitivity and the Small-Probability tasks was not successful under all three models, being below chance level. In the High-Variance task, the predictions of all three learning models (for the other two tasks) were better than the random baseline model. However, the predictions of the Delta model were superior, with 74% of the participants having better fit than the baseline model compared to 59% in the Decay-reinforcement model, and 53% in the Constrained Decay-reinforcement model. The differences between the Delta model and the other two models were significant in a Z-test for proportion differences (Decay-reinforcement: Z = 3.01, p < .01; Constrained Decay-reinforcement model: Z = 5.10, p < .01)[4].

The success of the Delta model in the generalization test cannot be attributed to a smaller sampling variance (or standard error) for the parameter estimates. To rule out this possibility we conducted a parametric bootstrap cross-fitting analysis (Efron, 1979; Wagenmakers et al., 2003). The procedure and the results are detailed in Appendix 1. The outcomes of this analysis indicated no advantage to the Delta model over the Decay-Reinforcement model. Thus, the success of the Delta model in the generalization test is not due to an interaction between the specific tasks used here and the sampling variance of the model parameters.

---

[4] Note that the G^2 averages are below zero because of some extreme low values.

The success of the Delta model (and of the High-Variance task) also cannot be attributed to high between-performer variance in model parameters. Fisher's F-tests showed no significant differences in the variance of the parameters in different models or in different tasks (in fact, the variances, presented in Table 2, were very similar). The High-Variance task also did not have more variability in preferences: The between-performer standard deviation in selections from alternative H (or L) was similar in all tasks (High-Variance: 0.25, Payoff-Sensitivity: 0.25, Small-Probability: 0.23).

Individual Parameter Consistency. In addition to studying the generalizability of the model predictions, we studied the consistency of each of the estimated parameters in the three tasks. For this purpose, one-sided Spearman correlations were used to examine the association between individual parameter values. Table 4 summarizes the results for the three models. As expected from the generalization tests, the correlations were generally higher for the Delta model than for the two versions of the Decay-reinforcement model. However, the only significant correlations were for parameters estimated in the High-Variance task.

Specifically, both the Delta and Decay-Reinforcement models showed a significant correlation between the weight to gains parameters extracted in the Small-Probability and High-Variance tasks (although it was higher for the Delta model). However, the Delta model also showed a significant correlation between the recency parameter in the High-Variance task and the same parameter extracted in the Small-Probability task ($r = 0.20$, $p < .05$) and in the Payoff-Sensitivity task ($r = 0.18$, $p < .05$). In the Constrained Decay-Reinforcement only the latter correlation was significant ($r =$

0.24, p < .05). A factor analysis of the parameters (see Appendix 2) confirms that the results are not due to covariance between different parameters.

4.3. Summary

Differences in the ranking of models emerged between the conventional evaluation method based on fit and the two generalization tests at the individual level. The two Decay-Reinforcement models produced high fit, especially in the Payoff-Sensitivity and Small-Probability tasks. However, the Delta model produced better generalization at the individual level, especially in the High-Variance task, as well as improved individual parameter consistency.

Interestingly, the High-Variance task produced the poorest fit, yet it led to the best generalizability to other tasks at the individual level. It is not clear why the High-Variance task (replication of Thaler et al.'s 1997 resource allocation task) produced the best generalizability. One possibility that was *refuted* is that this task has more between-performer variability in selections or parameter values.

It is our view that the task was successful simply because it is relatively difficult for decision makers to calculate the advantage of each choice alternative based on trial to trial outcomes. This was confirmed in discussions with participants following the experiment, and is further supported by the finding that, on average, the estimated choice sensitivity in this task was low but positive (mean c of 0.47 compared to 1.61 and 1.98 in the Payoff-Sensitivity and Small-Probability task, respectively). This implies that on average participants did learn from experience, but that their learning process was

relatively slow. A slow learning process based on trial to trial reinforcement may calibrate the model better. This task is re-examined in the next study.

5. Experimental results of Study 2 (Loss Domain)

In study 2 the tasks were changed so that the sign of all payoffs was reversed, and outcomes were mostly negative (see table 1). In this way, we assessed whether the results of Study 1 are restricted to tasks where most outcomes are in the gain domain[5].

5.1. Behavioral Results

Figure 3 summarizes the average choice proportion from the High expected value alternatives in the three tasks. As in Study 1, in the Payoff-Sensitivity task participants learned to select the High expected value alternative. Choices from H increased significantly from the first to the last block of 10 trials (t (89) = 4.98 , p < .01). However, in the Small-Probability task participants displayed learning towards the Low expected-value alternative that produced −10 with certainty (t (89) = 5.60 , p < .01). Finally, in the High-Variance task participants learned to stay away from the High-Variance alternative (t (89) = 1.92 , p = .06). Interestingly, the high variance option was equally enticing in the loss domain (where it produced an average of −100) and gain domain (where it had an average of +100), capturing about 35% of the choices in both studies.

As in study 1, there were no significant correlations between choices from the High expected value alternatives. However, the association between choices from the two

---

[5] For instance, if the advantage of the High-Variance task is due to the frequency of losses (upon choosing H; see Table 1) then it is not likely to be replicated in the inverted payoff version, in which the other tasks produce more frequent losses.

high-variance alternatives (H in the Small-Probability task and L in the High-Variance task) was significant for the entire task (r = 0.29, p < .01) and for the first 100 trials (r = 0.36, p < .01), indicating that participants were consistent in their choice of the high-variance alternative.

5.2. Model evaluation

We focused on the original version of the Decay-Reinforcement model because it yielded better fits and generalizations in the first study, as compared to the Constrained Decay model. In addition, we applied an adjustment proposed by Erev and his colleagues (Bereby-Meyer and Erev, 1998; Erev et al., 1999) for dealing with negative expectancies. Such negative expectancies are likely to develop when the outcomes are mainly losses. The adjustment, called Low Reference Point (LRP) solution, forces the expectancies to be positive. This is implemented by deducting the worst possible outcome (in a given task) from all payoffs. Accordingly, models with the LRP adjustment do not have a weight to gains parameter, as all payoffs occur in the gain domain. We implemented a version of the LRP that was considered to be best suited for the analysis at the individual level: In this version the payoff that was deducted was the worst payoff encountered by the individual up to the current trial[6].

Model fit. The average BICs of the models in the three tasks are summarized in Table 5. As can be seen, without the LRP adjustment, the fit of the Delta model was above the fit of the Decay-Reinforcement model (t(269) = 9.15, p < .01). The LRP

---

[6] The LRP is useful because without this adjustment the Decay-Reinforcement model cannot converge on a single alternative when the expectancies are negative (because if $\phi < 1$ then the unchosen alternative always improves faster than the chosen alternative).

adjustment significantly improved the fit of both learning models (Delta: $t(269) = 4.17$, p $< .01$, Decay-Reinforcement: $t(269) = 6.06$, $p < .01$). Under the LRP adjustment there was a slight advantage to the Decay-Reinforcement model over the Delta model ($t(269) = 0.63$, NS).

A comparison of specific tasks shows that as in Study 1, the Delta model produced lower fits than the Decay-Reinforcement model in the Payoff-Sensitivity task and in the Small-Probability task (both with and without the LRP adjustment). The fit of all four models was highest in the Payoff-Sensitivity task (average BIC of 6.4 across all models, compared to -14.4 in the Small-Probability task and -6.3 in the High-Variance task). The differences in BICs between tasks were significant for all models. However, for conciseness, the tests are not presented here.

Generalization at the Individual Level. The results of the generalization test appear in Table 6. The table indicates that as in Study 1 there was an interaction between task and learning model. Consistent with the previous results, the task that produced the best generalization was the High-Variance task. The predictions from the High-Variance task (to the other two tasks) using the Delta model were better than random for 61% of the participants, compared to only 41% under the Decay-Reinforcement model ($Z = 3.69$, $p < .01$). Thus, without the LRP adjustment, the Delta model produced better generalization at the individual level.

The LRP adjustment improved the generalization from the High-Variance task using both models. Yet the Delta model retained its advantage over the decay model (81% compared to 65%; $Z = 3.44$, $p < .01$). Moreover, the High-Variance task was again the only task that produced adequate fits in the generalization test.

Individual Parameter Consistency. One-sided Spearman correlations were used to examine the association between the parameter values estimated in the different tasks. Table 7 presents the results for the four models. Only the Delta model produced significant positive correlations. These correlations were primarily between the weight to gains and recency parameters estimated in two tasks: The Payoff-Sensitivity task and the Small-Probability task. However, as in our investigation in the gain domain, there were positive correlations between the parameters estimated in the Small-Probability task and High-Variance task as well, with p-values approaching significance (weight to gains: r = 0.15; p < .1; recency: r = 0.16; p < .1; see also factor analysis on Appendix 2).

None of the other models had any significant correlations. Interestingly, the LRP adjustment for the Delta model, which improved the generalizability of the model at the individual level, produced poor individual parameter consistency.

5.3. Summary

As in Study 1, the results show that the fit index and tests of generalizability at the individual level produced systematic differences in the ranking of models and tasks. The Decay-Reinforcement model (using the LRP adjustment) produced the best fit for the estimated data. However, the Delta model was more accurate in both tests of generalization at the individual level. Likewise, the High-Variance task produced poor fit, yet was proven to be the most useful task for generalization[7]. Finally, some differences

---

[7] We have previously argued that this task has a slow learning process, and therefore more trial-to-trial adaptation. This slow learning process also appeared in the current study. Under the Delta model the average of the choice consistency parameter was 0.73 (compared to 1.39 in the Payoff Sensitivity task and 2.01 in the Small Probability task; see Table 6).

emerged between the two methods of generalization at the individual level concerning the ranking of models with and without the LRP adjustment. This discrepancy is re-examined in the general discussion section below.

6. Can the model parameters improve the ability to predict risky choices?

The learning model parameters are usually interpreted as reflecting different facets of "risky" choices, such as overweighting gains compared to losses (Busemeyer and Stout, 2002; Wallsten et al., 2005). However, could it be the case that the parameters estimated in a given task would improve predictions of risky choices in a *different* task beyond the risky choices in the initial task? Theoretically, if the parameters measure consistent predispositions of an individual in the response to payoffs, and the tendency to select risky choices is due to such predispositions, then the parameters should be able to predict the choice of risky alternatives in a different task.

To examine the utility of the parameters for predicting risky choices in a different task, we used the High-Variance task as the source task and the Small-Probability task as the target task. We conducted two regression analyses, one for each study. The predictors were the proportions of risky choices in the first 100 trials in the High-Variance task and the rank order score of each parameter estimated in this task using the Delta model. The criterion variable was the tendency to take risk (or choose the high variance option) in the first 100 trials in the Small-Probability task. The regression was conducted with the Stepwise Selection procedure.

The results were as follows. In Study 1, the model parameters accounted for 9.2% of the observed variance, compared to only 4.5% using the choices from the High-

Variance task. Risk taking on the Small-Probability task was associated with high weight to gains (Partial r = 0.22; β= 0.21; t (87) = 2.05, p < .05) as well as to some extent with low recency (Partial r = 0.18; β= 0.18; t (87) = 1.71, p < .1) [8], estimated on the High-Variance task.

In study 2, the model comprising the weight to gains and recency parameters also achieved comparable predictions to the risky choice proportion, accounting for 12.0% of the observed variance compared to 12.7% based on the choice proportion. An examination of specific parameters reveals again that the weight to gains was the most prominent predictor (r = 0.29; β= 0.30; t (87) = 3.00, p < .01), followed by the recency parameter (r = 0.17; β= 0.20; t (87) = 1.95, p = .05).

In summary, the regression analyses confirm the role of the weighting to gains compared to losses and the recency parameters as predictors of risky choice in the current tasks. Moreover, in Study 1 the combined prediction of these two parameters from one task (the High-Variance task) to a different task (the Small-Probability task) was more accurate than the prediction using the risky choice proportion. This indicates the potential value of the current modeling approach for predictive purposes.

7. General Discussion

The results of the present two studies indicate that as postulated, superior accuracy does not guarantee high generablizability and strong parameter consistency at

---

[8] An in depth examination of the association between the recency parameter and the tendency to choose option L in the Small-Probability task reveals a U-shape association. Evidently the tendency to under-weight small probabilities can be associated with either very high recency (i.e., a hotstove effect; Denrell and March, 2001) or very low recency combined with discounting of losses compared to gains. For similar findings, see Yechiam and Busemeyer (2006).

the individual decision maker level. In particular, the Delta model emerged as a model with a lower accuracy level, but with better generablizability and consistency across tasks.

Specifically, in Study 1 participants showed similar rankings in two of the parameters of the Delta model across different tasks: The weight to gains parameter and the recency parameter. This was partially replicated in Study 2 (see also Appendix 2). The findings therefore suggest that these two parameters are stable across different tasks and may represent internal traits of individuals. Additionally, in an examination of the ability of different parameters to predict risky choices in a different task, high weighting to gains was the strongest predictor of risk taking in both studies.

The weight to gains (compared to losses) parameter may tap a motivational disposition similar to chronic promotion and prevention focus (Higgins, 1997). Promotion focus denotes greater concerns with the presence or absence of gains, and prevention focus denotes greater concerns about the presence or absence of losses. Consistent with Higgins's (1997) view that stable behavioral strategies are associated with this motivational disposition, high weighting to gains (on the Iowa Gambling task) was observed in drivers with multiple traffic offences (Lev et al., in press; Yechiam et al., 2007) and in incarcerated criminals sentenced for theft and drug crimes (Yechiam et al., 2007).

Surprisingly, the current evaluation method, developed for analyzing learning models, was also useful for evaluating an important related issue: Selecting choice tasks that produce stable parameters. The High-Variance task yielded better generality at the individual level in both studies. We have suggested that this was due to the prolonged

period of learning and exploration in this task. However, further research is required in order to pinpoint the exact conditions that make a task produce stable parameters. The present study is considered to be a preliminary step in evaluating these models and tasks. Its main goal was to demonstrate that the tests of generalizability and consistency at the individual level rank models differently than the conventional fit index; and that this ranking is stable in different conditions (Study 1 and 2). To fully evaluate different models and tasks, one should define the relevant task population and sample tasks from this population (see Roth and Erev, 1995).

Note also that the current studies focused on a specific learning situation where the distributions of payoffs are unknown and stationary, and where the aspiration level of players is ill defined. In our studies the distribution of payoffs could therefore only be discovered through sampling and observation. Generalizations to other settings require further research. Yet the principles of the current method involving multiple tasks, generalization at the individual level, and individual parameter consistency are by no means limited to this setting. We end this paper with a discussion of those principles.

What are the advantages of measuring generalization at the individual level *simultaneously* with individual parameter consistency? The demand for generalizability at the individual level and individual parameter consistency is derived from the theoretical assumption that the parameters represent internalized latent constructs of the decision maker. Accordingly, these tests are relevant when the goal of the scientific investigation is to construct a model that would produce adequate prediction in a variety of tasks. This is important in applications that seek to predict choice behavior in new situations and in examinations of individual differences. The test of generalization at the

individual level allows the examination of the overall model, integrating the specific contribution of each component process; whereas the test of individual parameter consistency enables the evaluation of the generalizability of specific component processes. Taken together, these two tests provide a diagnostic tool for explaining variations in generalizability and parameter consistency within individual decision makers (see Table 8).

For example, consider a person playing separate prisoner dilemma games (A and B), each with a different opponent. Let us assume that the choices of the player in the two games are heavily dictated by her cooperation and reciprocation (e.g., use of "tit for tat") levels. Let us also assume that these two parameters are robust within the player (i.e., they are applied similarly in the different games), and that the model successfully estimates these parameters. In this case, given the parameters estimated in game A, we can expect that we will be able to predict the player's choices in game B as well as obtain consistent levels in the two parameters (e.g., a high reciprocation / high cooperation player will be characterized as such in both games).

In some cases, however, the two tests are expected to give different results, and this is indicatory of specific modeling concerns. One such case is where individual parameter consistency is high for some parameters but low for others, and where the overall generalizability is impaired (compared to the example above). This can occur when some of the parameters are estimated inaccurately or when they are highly task specific. For instance, in the prisoner dilemma example, if the model succeeds in estimating the cooperation parameter but fails in estimating the reciprocation parameter,

then the estimated cooperation parameter might still be consistent in the two games. However, the overall prediction from one game to the other would be impaired.

Yet another possible case is where individual parameter consistency is low but generalizability at the individual level remains high. In the above example, a model might accurately predict the player's choices in game A using certain levels of the cooperation and reciprocation parameters. Moreover, it might be that this combination of parameters will also produce adequate predictions on game B. Yet there could still be poor individual parameter consistency in one or both of the parameters! Assuming that the true parameters are consistent, this interesting state of affairs implies that while the overall prediction of the model is general, the accurate prediction has been obtained by a set of parameters that *take on a different role when applied to each task separately*. This could happen if the predictions of the model are accurate for a wide range of parameter values, or in other words, the predictions are insensitive to the exact parameter values. Actually this is a good property for a model with respect to prediction, but a bad property with respect to estimating individual differences.

To demonstrate this case, consider the results in Study 2, in which the LRP adjustment improved the generalizability but impaired the individual parameter consistency. The present analysis suggests that this is due to the fact that LRP model parameters take on dissimilar roles in different tasks. One candidate parameter for that is the recency parameter. The LRP model lacks the weight to gains parameter, and so it could be that the recency parameter mimics this parameter in tasks involving

considerable losses (thereby changing its role in different tasks)[9]. While further investigation is necessary to discover the exact processes that take place in the LRP model, the present method is useful in diagnosing the general problem, as indicated by a specific discrepancy in the two tests of generality at the individual levels and individual parameter consistency.

8. Conclusions and future directions

The present paper describes a new evaluation method for learning models, which focuses on the generality of models at the individual level. This method opens a new avenue for discovering models that can make a priori predictions for how individual decision makers learn from experience. Furthermore, it enables us to select tasks on the basis of their success in eliciting parameters that are consistent across other tasks and conditions.

Note that while the present study focused on individual decision tasks, the current analysis is also relevant to repeated games. Learning models have been extensively used for predicting choices in games with feedback, yet no previous study has examined the generalizability of model parameters at the individual level. This is a challenging avenue for further research.

---

[9] Specifically, if a player is somewhat loss averse and if losses are pertinent in a certain task (as in the High-Variance task) then the recency parameter in that task may increase so as to allow quick learning to avoid losses. This would imply that the recency parameter for the LRP model would be (negatively) associated with weight to gains parameter. Indeed, this postulation is empirically supported. The Spearman correlation between the recency parameter of the LRP model and the weight to gains parameter of the original Delta model is -0.31 ($p < .01$), compared to a correlation of only -0.15 ($p < .05$; $Z = 1.96$, $p < .05$) between the recency and the weight to gains parameter in the original Delta model.

Appendix 1: Parametric bootstrap cross-fitting analysis

To rule out potential interactions between the specific tasks used here and parameter variance we conducted a parametric bootstrap cross-fitting analysis (Efron, 1979; Wagenmakers et al., 2003) as follows. Ten individuals were randomly selected and data was simulated using their estimated parameters on each of the three experimental tasks. We employed two models to generate this data: the Delta model and the Decay-reinforcement model. Fifty modeling agents were run in each task condition with these models' parameters (a total of 6000 agents)[10]. After the data was generated, we estimated the parameters again using the prediction of one-step-ahead method (as in our main analysis). Each model was run on the data it generated as well as on the data produced by the other model. We then examined the sampling distribution of the parameters. The results appear in the table below.

An examination of the sampling variance shows that the Decay-Reinforcement model generally had smaller parameter variances. Consider for example, the High-Variance task, which produced the best generalization. When the data was generated by the Delta model, the variances of the Decay parameters were consistently smaller than those of the Delta model. When the data was generated by the Decay model, the first Decay model parameter had significantly smaller variance than the first Delta model parameter, the variances were equal across models for the second parameter, and the variance was non significantly larger for the Decay as compared to the Delta for the third parameter. Thus, the success of the Delta model in the generalization test is not because of a smaller sampling variance in the current tasks. In addition, we also examined the Mean Square Deviation of the re-estimated parameters from the actual parameters used to generate the data. The results (appearing in parentheses in the table) show that neither model dominates the other. The Delta model had smaller MSDs for the Delta model data and the Decay-

---

[10] To the extent possible we used the exact same payoffs of the actual players. When the payoffs experienced by the player "ran out" we used a simulation based on the payoff distributions, as described in Table 1.

Reinforcement model had smaller MSDs for the data produced by this model. One exception is the Small-Probability task in which the Decay model had somewhat smaller MSDs even on the data produced by the Delta model. Clearly, the advantage of the Delta model in the generalization test does not stem from the sampling properties of true (estimated) parameters in the current tasks.

App. 1 Table. Parametric bootstrap cross-fitting analysis: Standard deviation and MSD (in parenthesis) of the re-estimated parameters in the three experimental tasks in Study 1 (averaged across 10 individuals and 50 simulations per individual, for a total of 500 simulations per row).

Task	Data	Model	Weight to gains (W)	Recency (ϕ)	Sensitivity (c)
Payoff Sensitivity	Delta	Delta	0.39 (0.29)	0.22 (0.14)	1.56 (3.87)
		Decay-Reinforcement	0.32 (0.42)	0.26 (0.48)	1.77 (6.30)
	Decay-Reinforcement	Delta	0.35 (0.31)	0.23 (0.36)	1.75 (4.37)
		Decay-Reinforcement	0.33 (0.20)	0.24 (0.09)	1.60 (4.53)
Small Probability	Delta	Delta	0.29 (0.45)	0.26 (0.58)	2.60 (16.50)
		Decay-Reinforcement	0.23 (0.25)	0.24 (0.31)	1.23* (10.79)
	Decay-Reinforcement	Delta	0.43 (0.35)	0.37 (0.50)	2.63 (6.81)
		Decay-Reinforcement	0.37 (0.33)	0.35 (0.21)	2.05 (5.70)
High Variance	Delta	Delta	0.36 (0.29)	0.28 (0.24)	2.00 (9.92)
		Decay-Reinforcement	0.12* (0.20)	0.28 (0.43)	1.76 (18.92)
	Decay-Reinforcement	Delta	0.34 (0.46)	0.39 (0.54)	2.90 (8.47)
		Decay-Reinforcement	0.07* (0.08)	0.39 (0.51)	3.17 (17.69)

* = p < .05 (paired t-tests for the average variance of the two models).

Appendix 2: Factor analysis of the model parameters

To control for covariance between different parameters, we conducted a Factor Analysis with Varimax Rotation and Kaiser Normalization for the parameters extracted in all tasks using the Delta model (total of 9 variables).

In study 1 the model explained 57.7% of the total variance. Four factors exceeded the cut-point of Eigenvalue > 1 (see Table; for conciseness only the first three factors are presented). Factors 1 and 3 were consistently associated with the same parameter across different tasks. Factor 1 (17.0% of the variance) was associated with the recency parameters estimated in all three tasks. Factor 3 (14.4% of the variance) was associated with weight to gains parameters estimated in the Small-Probability task and High-Variance task.

In study 2 the model explained 78.9% of the variance. Five factors exceeded the cutoff point. The first factor (24.8% of the variance) was associated with the weight to gains parameter in all three tasks. The second factor (16.8% of the variance) was associated with the recency parameter in two of the tasks. However, the associations in Factor 2 were much smaller than in the corresponding factor in Study 1 ($r < 0.4$ in all tasks). Thus, while a major factor was again related to the weight to gains parameter, there was no clear factor associated with recency

App. 2 Table. Factor loading for the parameters extracted using the Delta model in the three tasks. Loadings larger than 0.20 appear in bold font and consistent associations with the same parameter appear in gray background.

Task	Parameter	Study 1			Study 2		
		1 (17.0%)	2 (15.2%)	3 (14.4%)	1 (24.8%)	2 (16.8%)	3 (13.6%)
Payoff Sensitivity	Weight to gains (W)	**-0.33**	**0.48**	**-0.21**	**0.31**	**0.75**	**0.20**
	Recency (ϕ)	**0.69**	-0.01	-0.18	-0.18	**0.38**	0.02
	Sensitivity (c)	-0.17	0.07	-0.05	-0.04	**-0.86**	-0.09
Small Probability	Weight to gains (W)	-0.03	-0.11	**0.61**	**0.86**	0.04	0.02
	Recency (ϕ)	**0.29**	**0.61**	**0.41**	**0.80**	**0.20**	0.05
	Sensitivity (c)	0.03	**-0.83**	0.08	-0.19	**0.22**	-0.07
High Variance	Weight to gains (W)	-0.01	0.13	**0.78**	**0.23**	-0.10	-0.01
	Recency (ϕ)	**0.70**	0.02	-0.04	0.09	0.16	**0.84**
	Sensitivity (c)	**-0.46**	-0.02	**-0.26**	0.02	-0.08	**-0.86**

References

Atkinson, R. C., Shiffrin, R. M., 1968. Human memory: A proposed system and its control processes. In Spence, K. W., Spence, J. T. (Eds.), The Psychology of Learning and Motivation Vol. 2, Academic Press, New York, pp. 89-195.

Barron, G., Erev, I., 2003. Small feedback-based decisions and their limited correspondence to description-based decisions. J. Behav. Decis. Making 16, 215-233.

Bechara, A., Damasio, A. R., Damasio, H., Anderson, S., 1994. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 50, 7-15.

Bereby-Meyer, Y., Erev, I., 1998. On learning to become a successful loser: A comparison of alternative abstractions of learning processes in the loss domain. J. Math. Psychol. 42, 266-286.

Broadbent, D. E., 1958. Perception and Communication. Pergamon, London.

Busemeyer, J. R., Myung, I. J., 1992. An adaptive approach to human decision-making: Learning theory, decision theory, and human performance. J. Exp. Psychol. Gen. 121, 177-194.

Busemeyer, J. R., Stout, J. C., 2002. A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. Psychol. Assessment 14, 253-262.

Camerer, C., Ho, T. –H, 1999. 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 Econ. Behav. 19, 46-76.

Denrell, J., March, J. G., 2001. Adaptation as information restriction: The hot stove effect. Organ. Sci. 12, 523-538.

Erev, I., Barron, G., 2005. On adaptation, maximization, and reinforcement learning among cognitive strategies. Psychol. Rev. 112, 912-931.

Erev, I., Bereby-Meyer, Y., Roth, A. E., 1999. The effect of adding a constant to all payoffs: Experimental investigation, and implications for reinforcement learning models. J. Econ. Behav. Organ. 39, 111-128.

Erev, I., Rapoport, A., 1998. Magic, reinforcement learning and coordination in a market entry game. Games Econ. Behav. 23, 146-175.

Erev, I., Roth, A. E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. Am. Econ. Rev. 88, 848-881.

Fowles, D.C., 1988. Psychophysiology and psychopathology: A motivational approach. Psychophysiology 25, 373-391.

Fudenberg, D., Levine, D. K., 1995. Consistency and cautious fictitious play. J. Econ. Dyn. Control 19, 1065-1089.

Gluck, M. A., Bower, G. H., 1988. From conditioning to category learning: An adaptive network model. J. Exp. Psychol. Gen. 128, 309-331.

Gray, J. A., 1994. Personality dimensions and emotion systems. In Ekman, P., Davidson, R.J. (Eds.), The Nature of Emotion: Fundamental Questions. Oxford University Press, New York, pp. 329–331.

Hjelle, L. A., Ziegler, D. J., 1981. Personality Theories. McGraw-Hill, Auckland.

Higgins, E. T., 1997. Beyond pleasure and pain. Am. Psychol. 52, 1280-1300.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 263-291.

Lev, D., Hershkovitz, E., Yechiam, E., In press. Decision making and personality in traffic offenders: A study of Israeli drivers. Accident Anal. Prev.

Luce, R. D., 1959. Individual Choice Behavior. Wiley, New York.

Myung, I. J., 2000. The importance of complexity in model selection. J. Math. Psychol. 44, 190-204.

Nelder, J. A., Mead, R., 1965. A simplex method for function minimization. Comput. J. 7, 308-313.

Newell, A., 1992. Unified theories of cognition and the role of soar. In Michon, J. A., Anureyk, A. (Eds.), Soar: A Cognitive Architecture in Perspective. Kluwer Academic Publishers, Dordrecht, pp. 25-75.

Newman, J. P., Widom, C. S., Nathan, S., 1985. Passive avoidance in syndromes of disinhibition, psychopathy and extraversion. J. Pers. Soc. Psychol. 48, 1316-1327.

Oberauer, K., Kliegl, R., 2001. Beyond resources - formal models for complexity effects and age differences in working memory. Eur. J. Cogn. Psychol. 13, 187-215.

Rieskamp, J., Busemeyer, J., Laine, T., 2003. How do people learn to allocate resources? Comparing two learning theories. J. Exp. Psychol. Learn. 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 Econ. Behav. 8, 164-212.

Rumelhart, D. E., McClelland, J. E., and the PDP Research Group, 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volumes 1 and 2. MIT Press, Cambridge, MA.

Sarin, R., Vahid, F., 1999. Payoff assessments without probabilities: A simple dynamic model of choice. Games Econ. Behav. 28, 294-309.

Sarin, R., Vahid, F., 2001. Predicting how people play games: A simple dynamic model of choice. Games Econ. Behav. 34, 104-122.

Schwartz, G., 1978. Estimating the dimension of a model. Ann. Stat. 5, 461-464.

Stahl, D., 1996. Boundedly rational rule learning in a guessing game. Games Econ. Behav. 16, 303-330.

Sutton, R. S., Barto, A. G., 1998. Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA.

Thaler, R. H., Tversky, A., Kahneman, D., Schwartz, A., 1997. The effect of myopia and loss aversion on risk taking: An experimental test. Q. J. Econ. 112, 647-661.

Wallsten, T. W., Pleskac, T., Lejuez, C. W., 2005. Modeling a sequential risk-taking task. Psychol. Rev. 112, 862-880.

Yechiam, E., Budescu, D. V., 2006. The sensitivity of probability assessments to time units and performer characteristics. Risk Anal. 3, 177-193.

Yechiam. E., Busemeyer, J. R., 2005. Comparison of basic assumptions embedded in learning models for experience based decision-making. Psychon. B. Rev. 12, 387-402.

Yechiam, E., Busemeyer, J. R., 2006. The effect of foregone payoffs on underweighting small probability events. J. Behav. Decis. Making 19, 1-16.

Yechiam, E., Busemeyer, J. R., Stout, J. C., Bechara, A., 2005. Using cognitive models to map relations between neuropsychological disorders and human decision making deficits. Psychol. Sci. 16, 973-978.

Yechiam, E., Kanz, J., Bechara, A., Stout, J. C., Busemeyer, J. R., Altmaier, E. M., Paulsen, J., 2007. Neurocognitive deficits related to poor decision-making in people behind bars: A peek inside their brain. Manuscript submitted for publication.

Yechiam, E., Veinott, E. S., Busemeyer, J. R., Stout, J. C., 2007. Cognitive models for evaluating basic decision processes in clinical populations. In Neufeld, R. (Ed.), Advances in Clinical Cognitive Science: Formal Modeling and Assessment of Processes and Symptoms. APA Publications, Washington, DC, pp 81-111.

Table 1. The payoff distributions of the three tasks in Study 1 and 2. H denotes High expected value, and L denotes Low expected value; N~(X,Y) denotes a Normal distribution with a mean of X and a standard deviation of Y.

Task	Payoff in Study 1	Payoff in Study 2
1) Payoff-Sensitivity (PS)	H: N~(20,20 truncated at –10, 50) L: N~(10,20 truncated at –20,40)	H: N~(-10,20 truncated at –20,40) L: N~(-20,20 truncated at –10, 50)
2) Small-Probability (SP)	H: 10 with certainty. L: 20 90% of the time, otherwise –90*	H: -20 90% of the time, otherwise 90* L: -10 with certainty
3) High-Variance (HV)	H: N~(100,354) L: N~(25, 17.7, truncated at 0)	H: N~(-25, 17.7, truncated at 0) L: N~(-100,354)

* This outcome includes a noise factor, so that it is constantly distributed between 85 and 95, rounded to the closest integer.

Table 2. Study 1: Means and standard deviations (in parenthesis) of the BIC scores and estimated parameters of the models in the three experimental tasks.

Task	Model	BIC	Weight to gains (W)	Recency (ϕ)	Sensitivity (c)
Payoff Sensitivity	Delta	8.41 (18.1)	0.59 (0.4)	0.62 (0.4)	1.11 (2.2)
	Decay-Reinforcement	13.72 (19.3)	0.57 (0.4)	0.47 (0.4)	0.95 (1.5)
	Constrained Decay-Reinforcement	11.64 (19.9)	0.60 (0.4)	0.42 (0.3)	1.61 (1.6)
Small Probability	Delta	2.72 (20.2)	0.60 (0.4)	0.43 (0.4)	1.78 (1.3)
	Decay-Reinforcement	11.37 (20.8)	0.62 (0.4)	0.64 (0.4)	1.01 (1.2)
	Constrained Decay-Reinforcement	7.36 (22.3)	0.68 (0.4)	0.43 (0.4)	1.98 (1.3)
High Variance	Delta	-1.86 (17.0)	0.27 (0.4)	0.31 (0.3)	0.47 (1.7)
	Decay-Reinforcement	-2.76 (20.4)	0.36 (0.4)	0.64 (0.4)	-0.63 (1.6)
	Constrained Decay-Reinforcement	-3.77 (20.6)	0.35 (0.4)	0.78 (0.3)	1.34 (2.1)

Table 3. Study 1: Mean G^2 scores and percent of individuals for which the generalization prediction is better than a random model (in parenthesis). Comparison of the three models in the three experimental tasks (PS = Payoff Sensitivity; SP = Small Probability; HV = High Variance). Shaded cells denote tests that do not involve generalization.

Delta model			
Target task	Tasks used for estimating the parameters		
	PS task	SP task	HV task
PS task	33.7 (100%)	-32.3 (46.7%)	-8.5 (82.2%)
SP task	-60.7 (34.4%)	37.3 (100%)	1.5 (66.7%)
HV task	-303.3 (18.9%)	-314.1 (16.7%)	44.5 (100%)
Average*	-181.9 (26.7%)	-173.2 (31.7%)	-3.5 (74.4%)

Decay-Reinforcement model			
Target task	Tasks used for estimating the parameters		
	PS task	SP task	HV task
PS task	39.1 (100%)	-32.2 (44.4%)	0.5 (64.4%)
SP task	-69.5 (52.2%)	46.0 (100%)	-17.0 (54.4%)
HV task	-198.1 (22.2%)	-224.9 (17.7%)	43.6 (87.8%)
Average*	-66.0 (37.2%)	-91.6 (31.1%)	-8.2 (59.4%)

Constrained Decay-Reinforcement model			
Target task	Tasks used for estimating the parameters		
	PS task	SP task	HV task
PS task	36.3 (97.8%)	-31.1 (48.9%)	-20.1 (56.7%)
SP task	-51.9 (48.9%)	41.8 (98.9%)	-32.2 (50.0%)
HV task	-142.9 (23.3%)	-151.9 (22.2%)	40.6 (93.3%)
Average*	-97.4 (36.1%)	-91.5 (35.6%)	-26.1 (53.3%)

* = Average for the generalization tests only

Table 4. Study 1: Spearman correlations between parameter values estimated in the different tasks (PS = Payoff Sensitivity; SP = Small Probability; HV = High Variance).

Delta model

	Weight to gains (W)	Recency (ϕ)	Sensitivity (c)
PS – SP tasks	-0.02	0.09	0.00
PS – HV tasks	-0.04	0.18*	0.11
SP – HV tasks	0.24*	0.20*	0.05

Decay-Reinforcement model

	Weight to gains (W)	Recency (ϕ)	Sensitivity (c)
PS – SP tasks	-0.03	-0.08	-0.07
PS – HV tasks	0.02	0.06	0.10
SP – HV tasks	0.20*	-0.07	0.13

Constrained Decay-Reinforcement model

	Weight to gains (W)	Recency (ϕ)	Sensitivity (c)
PS – SP tasks	0.13	0.09	0.11
PS – HV tasks	-0.01	0.24*	0.08
SP – HV tasks	0.07	-0.16	0.10

* = p < .05

Table 5. Study 2: Means and standard deviations (in parenthesis) of the BIC scores and

estimated parameters of the models in the three experimental tasks.

Task	Model	BIC	Weight to gains (W)	Recency (ϕ)	Sensitivity (c)
Payoff Sensitivity	Delta	2.75 (19.3)	0.46 (0.5)	0.69 (0.4)	1.39 (2.0)
	Delta – LRP	6.57 (15.7)	-	0.68 (0.4)	0.40 (1.8)
	Decay-Reinforcement	3.56 (15.9)	0.72 (0.4)	0.55 (0.4)	0.80 (2.2)
	Decay-Reinforcement – LRP	12.71 (19.5)	-	0.35 (0.3)	0.97 (1.3)
Small Probability	Delta	-7.30 (17.5)	0.30 (0.4)	0.62 (0.4)	2.01 (1.8)
	Delta – LRP	-3.33 (15.5)	-	0.64 (0.4)	-0.82 (2.0)
	Decay-Reinforcement	-52.53 (51.2)	0.83 (0.3)	0.46 (0.4)	-0.15 (3.9)
	Decay-Reinforcement – LRP	5.42 (17.22)	-	0.68 (0.4)	0.44 (1.4)
High Variance	Delta	2.40 (14.1)	0.51 (0.4)	0.40 (0.4)	0.73 (2.0)
	Delta – LRP	2.34 (12.22)	-	0.24 (0.3)	1.26 (2.0)
	Decay-Reinforcement	-14.69 (35.3)	0.58 (0.4)	0.51 (0.4)	-1.15 (1.9)
	Decay-Reinforcement – LRP	-15.39 (39.4)	-	0.26 (0.3)	2.51 (1.4)

Table 6. Study 2: Mean G^2 scores and percent of individuals for which the generalization prediction is better than a random model (in parenthesis). Comparison of the two models in the different tasks (PS = Payoff Sensitivity; SP = Small Probability; HV = High Variance). Shaded cells denote tests that do not involve generalization.

	Delta model			Delta - LRP		
Target task	Task used for estimating the parameters			Task used for estimating the parameters		
	PS task	SP task	HV task	PS task	SP task	HV task
PS task	26.5 (99.9%)	-84.7 (28%)	-4.17 (72%)	26.0 (98.9%)	-128.9 (24.4%)	1.6 (77.8%)
SP task	-47.8 (54%)	51.5 (99.9%)	-20. 0 (49%)	-9.2 (68.9%)	52.1 (98.9%)	-0.4 (83.3%)
HV task	-255.5 (13%)	-247.4 (13.5%)	22.6 (100%)	-214.6 (20.2%)	-234.8 (20.0%)	33.4 (100%)
Average*	-151.7 (33.9%)	-166.0 (21.0%)	-12.1 (60.6%)	-111.9 (44.6%)	-181.9 (22.2%)	0.6 (80.6%)
	Decay-Reinforcement model			Decay-Reinforcement - LRP		
Target task	Task used for estimating the parameters			Task used for estimating the parameters		
	PS task	SP task	HV task	PS task	SP task	HV task
PS task	27.5 (100%)	-25.5 (41.1%)	-12.2 (57.8%)	32.1 (94.4%)	-124.2 (24.4%)	-10.9 (56.6%)
SP task	-146.9 (11.1%)	7.3 (70.0%)	-32.4 (23.3%)	10.5 (78.9%)	60.9 (94.4%)	4.3 (73.3%)
HV task	-269.0 (4.5%)	-110.0 (14.6%)	4.1 (88.9%)	-222.2 (19.1%)	-294.5 (16.7%)	15.3 (53.3%)
Average*	-208.0 (7.8%)	-51.3 (42%)	-22.3 (40.6%)	-105.9 (49.0%)	-209.4 (21%)	-3.3 (65.0%)

* = Average for the generalization tests only

Table 7. Study 2: Spearman correlations between parameter values estimated in the different tasks (PS = Payoff Sensitivity; SP = Small Probability; HV = High Variance).

	Delta model			Delta - LRP	
	Wgt. gains (W)	Recency (ϕ)	Sensitivity (c)	Recency (ϕ)	Sensitivity (c)
PS – SP tasks	0.220*	0.179*	-0.031	0.083	-0.081
PS – HV tasks	0.119	-0.120	0.191*	-0.123	0.058
SP – HV tasks	0.163	0.151	0.106	0.122	-0.090
	Decay-Reinforcement model			Decay-Reinforcement - LRP	
	Wgt. gains (W)	Recency (ϕ)	Sensitivity (c)	Recency (ϕ)	Sensitivity (c)
PS – SP tasks	-0.196	-0.069	-0.053	0.075	-0.010
PS – HV tasks	0.133	0.131	0.145	0.117	-0.120
SP – HV tasks	-0.118	0.057	0.061	0.101	0.083

* = $p < .05$

Table 8. The general characteristics of the estimated model parameters based on the tests of generalizability at the individual level and individual parameter consistency.

		Generalizability at the individual level	
		Low	High
Individual parameter consistency	Low	Robust parameters *not* estimated by the model	Predictions are not dependent on exact parameters
	High	Certain parameters are task specific or inaccurate	Robust parameters estimated by the model

Figure 1: A screen capture of the experimental task in Studies 1 and 2.

Figure 2: Proportion of choices from the High expected value alternative (H) as a function of time in the three experimental tasks of study 1.
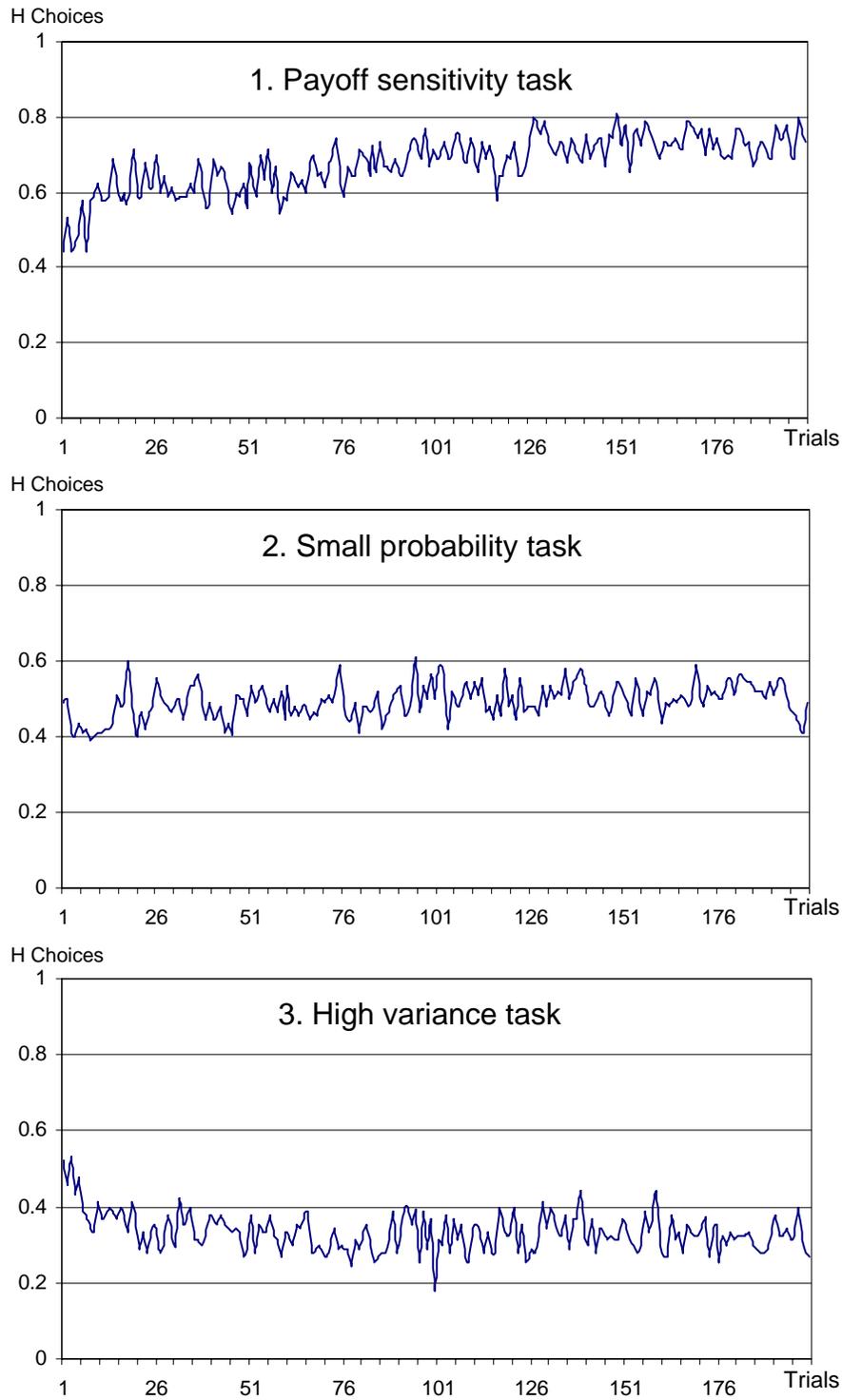
Figure 3: Proportion of choices from the High expected value alternative (H) as a function of time in the three experimental tasks of study 2.