In press for *Decision*

Absolute Performance of Reinforcement-Learning Models for the Iowa Gambling Task

Helen Steingroever[a], Ruud Wetzels[b], and Eric-Jan Wagenmakers[a]

[a] University of Amsterdam, Department of Psychology, Weesperplein 4,

1018 XA Amsterdam, The Netherlands

[b] The European Foundation for the Improvement of Living and Working Conditions,

Dublin, Ireland

Short title: Absolute Performance of RL Models for the IGT

Abstract

Decision-making deficits in clinical populations are often studied using the Iowa gambling task (IGT). Performance on the IGT can be decomposed in its constituent psychological processes by means of cognitive modeling analyses. However, conclusions about the hypothesized psychological processes are valid only if the model provides an adequate account of the data. In this article, we systematically assessed absolute model performance of the Expectancy Valence (EV) model, the Prospect Valence Learning (PVL) model, and a hybrid version of both models –the PVL-Delta model– using two different methods. These methods assess (1) whether a model provides an acceptable fit to an observed choice pattern, and (2) whether the parameters obtained from model fitting can be used to generate the observed choice pattern. Our results show that all models provided an acceptable fit to two stylized data sets; however, when the model parameters were used to generate choices, only the PVL-Delta model captured the qualitative patterns in the data. These findings were confirmed by fitting the models to five published IGT data sets. Our results highlight that a model's ability to fit a particular choice pattern does not guarantee that the model can also generate that same choice pattern. Future applications of RL models should carefully assess absolute model performance to avoid premature conclusions about the psychological processes that drive performance on the IGT.

Keywords: Decision making, Expectancy Valence Model, Prospect Valence Learning Model, Bayesian Hierarchical Analysis

The Iowa gambling task (IGT; Bechara, Damasio, Damasio, & Anderson, 1994) is arguably the most popular neuropsychological paradigm to assess decision-making deficits in clinical populations. Originally, the IGT was developed to assess decision-making deficits of patients with lesions to the ventromedial prefrontal cortex (vmPFC), but in the last two decades the task has been applied to a variety of clinical populations, such as patients with Asperger's disorder (e.g., Johnson, Yechiam, Murphy, Queller, & Stout, 2006), attention-deficit-hyperactivity disorder (e.g., Agay, Yechiam, Carmel, & Levkovitz, 2010; Toplak, Jain, & Tannock, 2005), bipolar disorder (e.g., Brambilla et al., 2012), obsessive-compulsive disorder (e.g., Cavedini, Riboldi, D'Annucci, et al., 2002), pathological gambling disorder (e.g., Cavedini, Riboldi, Keller, D'Annucci, & Bellodi, 2002), psychopathic tendencies (e.g., Blair, Colledge, & Mitchell, 2001), and schizophrenia (e.g., Martino, Bucay, Butman, & Allegri, 2007; Premkumar et al., 2008). In addition, the IGT has been applied to cocaine addicts (e.g., Stout, Busemeyer, Lin, Grant, & Bonson, 2004), chronic cannabis users (e.g., Fridberg et al., 2010), heavy drinkers (e.g., Gullo & Stieger, 2011), inmates (e.g., Yechiam et al., 2008), and traffic offenders (e.g., Lev, Hershkovitz, & Yechiam, 2008). Impaired performance on the IGT may be caused by several factors, such as only focusing on immediate rewards, avoidance of immediate losses, poor memory for past payoffs, or underweighting of rare events (e.g., Barron & Erev, 2003; Yechiam & Busemeyer, 2005).

In order to isolate and identify the psychological processes that drive performance on the IGT, behavioral analyses of IGT data need to be complemented with cognitive modeling analyses. To further this goal, several reinforcement-learning (RL) models have been proposed, and here we focus on the two most popular exemplars –the Expectancy Valence model (EV; Busemeyer & Stout, 2002) and the Prospect Valence Learning model (PVL; Ahn, Busemeyer, Wagenmakers, & Stout, 2008; Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011)– and the hybrid PVL-Delta model (Ahn et al., 2008; Fridberg et al., 2010; a detailed description of the three models can be found in section 1.2). The

Correspondence concerning this article should be addressed to: Helen Steingroever, Ph: +31 20525 8869, E-mail: helen.steingroever@gmail.com.

parameters of these models correspond to psychological processes such as motivation, learning/memory, and response consistency (Busemeyer, Stout, & Finn, 2003); hence, the purpose of fitting these models to empirical data is to allow applied researchers to draw conclusions about the latent psychological processes that drive performance on the IGT. Yechiam, Busemeyer, Stout, and Bechara (2005), for instance, fit the EV model to data of 10 groups of people suffering from various neuropsychological disorders (e.g., Asperger's syndrome, vmPFC lesions, chronic cannabis abuse), and mapped these groups according to the differences between their model parameters and those of their control group. The purpose of this analysis was to characterize the decision-making deficits of each clinical group in terms of underlying psychological processes, and to examine whether differences in neuropsychological disorders can be explained by differences in psychological processes underlying decision-making deficits (i.e., differences in the model parameters).

A prerequisite for drawing valid conclusions from RL model parameters is that the model provides an adequate account for the IGT data. However, systematic and detailed evaluations of model performance are virtually absent from the applied literature (see section 1.3 for more details on previous methods that applied studies used to assess model performance). This state of affairs makes it difficult to determine whether researchers can draw valid conclusions from parameters of RL models.

Here we outline two methods for the assessment of absolute model performance (i.e., the degree to which the choice behavior produced by a certain model matches the observed choice behavior). One method, the post hoc absolute fit, assesses a model's ability to fit an observed choice pattern when provided with information on the observed choices and payoffs. The other method, the simulation method, assesses a model's ability to generate the observed choice pattern with parameter values obtained from model fitting. The crucial difference between the two methods is that the first method is guided by information on the observed choices and payoffs, whereas the second method makes predictions using new, unobserved payoff sequences.

To anticipate our main result, the post hoc absolute fit method revealed that all models provided an acceptable fit to two stylized data sets (i.e., constructed data sets that consist of homogeneous participants with small individual differences). In contrast, the simulation method revealed that the EV and PVL models failed to generate both types of choice patterns present in the two stylized data sets. However, the PVL-Delta model adequately generates all choice patterns. These results were confirmed by fitting five complete data sets. Our findings show that a model's ability to fit a particular choice pattern does not guarantee that the model is also able to generate that same choice pattern. This indicates that a good post hoc absolute fit performance may be caused by choice mimicry (see also Ahn et al., 2008; Erev & Haruvy, 2005; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008, and see Lewandowsky, 1995, for a similar phenomenon).

The outline of this article is as follows. In the first section we explain the IGT, outline the three RL models, and review methods to assess performance of RL models. In the second section, we compare the absolute performance of the three RL models using the post hoc absolute fit method and the simulation method. In particular, we compare the ability of the three RL models to fit and generate choice patterns present in two stylized IGT data sets, and investigate whether our results generalize to five IGT data sets from the review article of Steingroever, Wetzels, Horstmann, Neumann, and Wagenmakers (2013). In the last section, we summarize our findings and discuss their ramifications. Readers already familiar with the IGT, the RL models, and their Bayesian hierarchical implementation may skip the corresponding parts of the first and second section below.

1. The Iowa Gambling Task and Three Reinforcement-Learning Models

1.1. The Iowa Gambling Task

In this section we describe the IGT (see also Steingroever, Wetzels, & Wagenmakers, 2013). The purpose of the IGT is to measure decision-making deficits of clinical populations in an experimental setting. In the traditional IGT, participants are initially given $2000 facsimile money and are presented with four decks of cards with different payoffs. Participants are instructed to choose cards in order to maximize their long-term net outcome

Table 1
Payoff scheme of the traditional IGT as developed by Bechara et al. (1994).

	Deck A	Deck B	Deck C	Deck D
	Bad deck with frequent losses	Bad deck with infrequent losses	Good deck with frequent losses	Good deck with infrequent losses
Reward/trial	100	100	50	50
Number of losses/10 cards	5	1	5	1
Loss/10 cards	−1250	−1250	−250	−250
Net outcome/10 cards	−250	−250	250	250

(Bechara et al., 1994; Bechara, Damasio, Tranel, & Damasio, 1997). Unbeknownst to the participants, the task typically contains 100 trials. After each choice, participants receive feedback on the rewards and the losses (if any) associated with that card, and the running tally.

The task aims to determine whether participants learn to prefer the good, safe decks over the bad, risky decks because this is the only choice pattern that maximizes the long-term net outcomes. The good, safe decks are typically labeled as decks C and D, whereas the bad, risky decks are labeled as decks A and B. Table 1 presents the traditional payoff scheme as developed by Bechara et al. (1994). This table illustrates that decks A and B yield high immediate, constant rewards, but even higher unpredictable, occasional losses: hence, the long-term net outcome is negative. Decks C and D, on the other hand, yield low immediate, constant rewards, but even lower unpredictable, occasional losses: hence, the long-term net outcome is positive. In addition to the different payoff magnitudes, the decks also differ in the frequency of losses: Two decks yield frequent losses (decks A and C) and two decks yield infrequent losses (decks B and D).

1.2. The EV, PVL, and PVL-Delta Models

In this section, we describe the EV, PVL, and PVL-Delta models (see also Steingroever, Wetzels, & Wagenmakers, 2013). Table 2 contains the model equations, the psychological interpretation of the free parameters, and their ranges. In the following, we describe each model separately; the general idea, however, is that each model describes the

Table 2
Formalization of the EV, PVL, and PVL-Delta models.

Concept	Model(s)	Model equation	Free parameters	Range
Utility function	EV	$u_k(t) = (1-w) \cdot W(t) + w \cdot L(t)$	w: Attention weight	$[0,1]$
	PVL & PVL-Delta	$u_k(t) = \begin{cases} X(t)^A & if\ X(t) \geq 0 \\ -w \cdot \lvert X(t) \rvert^A & if\ X(t) < 0 \end{cases}$	A: Shape w: Loss aversion	$[0,1]$ $[0,5]$
Learning rule	EV & PVL-Delta	$Ev_k(t) = Ev_k(t-1) + a \cdot (u_k(t) - Ev_k(t-1))$	a: Updating	$[0,1]$
	PVL	$Ev_k(t) = a \cdot Ev_k(t-1) + \delta_k(t) \cdot u_k(t)$	a: Recency	$[0,1]$
Choice rule	All	$P[S_k(t+1)] = \frac{e^{\theta(t) Ev_k}}{\sum_{j=1}^{4} e^{\theta(t) Ev_j}}$		
Sensitivity	EV	$\theta(t) = (t/10)^c$	c: Consistency	$[-2,2]$
	PVL & PVL-Delta	$\theta(t) = 3^c - 1$	c: Consistency	$[0,5]$

Note. $W(t)$ and $L(t)$ are the rewards and losses, respectively, on trial t. $X(t)$ is the net outcome on trial t, $X(t) = W(t) - |L(t)|$. $\delta_k(t)$ is a dummy variable that takes the value 1 if deck k is chosen on trial t and 0 otherwise.

performance on the IGT through the interaction of distinct psychological processes captured by the model parameters.

The EV, PVL, and PVL-Delta models share the assumption that, following each choice, participants evaluate the rewards and losses (if any) associated with the just-chosen card by means of a utility function. These momentary utilities are used to update expectancies about the utilities of all decks. This updating process entails that, on every trial, participants adjust their expected utilities of the decks based on the new utility they just experienced, a process described by a learning rule. In the next step, the models assume that the expected utilities of all decks are used to guide the participants' choices on the next trial. This assumption is formalized by the softmax choice rule, also known as the ratio-ofstrength choice rule, that all models use to compute the probability of choosing a particular deck on a particular trial (Luce, 1959). This rule contains a sensitivity parameter $\theta(t)$ that indexes the extent to which trial-by-trial choices match the expected deck utilities. Values of $\theta(t)$ close to zero indicate a random choice behavior (i.e., strong exploration), whereas large values of $\theta(t)$ indicate a choice behavior that is strongly determined by the expected deck utilities (i.e., strong exploitation). As is customary, for all analyses in this paper, we scaled the traditional payoffs of the IGT as presented in Table 1 by dividing by 100 (cf. Ahn et al., 2011).

1.2.1. The EV model. The EV model uses three parameters to formalize its assumptions about participants' performance on the IGT (Busemeyer & Stout, 2002). The first model assumption is that after choosing a card from deck k, $k \in \{1, 2, 3, 4\}$ on trial t, participants compute a weighted mean of the experienced rewards $W(t)$ and losses $L(t)$ to obtain the utility of deck k on trial t, $u_k(t)$. The weight that participants assign to losses relative to rewards is the attention weight parameter w. A small value of w, that is, $w < .5$, is characteristic for decision makers who put more weight on the immediate rewards and can thus be described as reward-seeking, whereas a large value of w, that is, $w > .5$, is characteristic for decision makers who put more weight on the immediate losses and can thus be described as loss-averse (Ahn et al., 2008; Busemeyer & Stout, 2002).

The EV model assumes that decision makers use the utility of deck k on trial t, $u_k(t)$, to update only the expected utility of deck k, $Ev_k(t)$; the expected utilities of the unchosen decks are left unchanged. This updating process is described by the Delta learning rule, also known as the Rescorla-Wagner rule (Rescorla & Wagner, 1972). If the experienced utility $u_k(t)$ is higher than expected, the expected utility of deck k is adjusted upward. If the experienced utility $u_k(t)$ is lower than expected, the expected utility of deck k is adjusted downward. This updating process is influenced by the second model parameter—the updating parameter a. This parameter quantifies the memory for rewards and losses. A value of a close to zero indicates slow forgetting and weak recency effects, whereas a value of a close to one indicates rapid forgetting and strong recency effects. For all models, we initialized the expectancies of all decks to zero, $Ev_k(0) = 0$. This setting reflects an absence of prior knowledge about the payoffs of the decks.

According to the EV model, the sensitivity $\theta(t)$ changes over trials depending on the response consistency parameter c. If c is positive, successive choices become less random and more determined by the expected deck utilities; if c is negative, successive choices become more random and less determined by the expected deck utilities, a pattern that is clearly non-optimal. We restricted the consistency parameter of the EV model to the range $[-2, 2]$ instead of the proposed range $[-5, 5]$ (Busemeyer & Stout, 2002). This modification improved the estimation of the EV model and prevented the choice rule from producing numbers that exceed machine precision.

In sum, the EV model has three parameters: (1) The attention weight parameter w, which quantifies the weight of losses over rewards, (2) the updating parameter a, which determines the memory for past expectancies, and (3) the response consistency parameter c, which determines the amount of exploration versus exploitation.

1.2.2. The PVL model. The PVL model uses four parameters to formalize its assumptions about participants' performance on the IGT (Ahn et al., 2008, 2011). The PVL model assumes that decision makers only process the net outcome after choosing a card from deck k on trial t, $X(t) = W(t) - |L(t)|$. In contrast to the linear utility function of the EV model, the PVL model uses the Prospect Utility function—a non-linear utility function from prospect theory (Tversky & Kahneman, 1992). The Prospect Utility function contains the first two model parameters—the shape parameter A, that determines the shape of the utility function, and the loss aversion parameter w. As A approaches zero, the shape of the utility function approaches a step function. The implication of such a step function is that given a positive net outcome $X(t)$, all utilities are similar because they approach one, and given a negative net outcome $X(t)$, all utilities are also similar because they approach $-w$. On the other hand, as A approaches one, the subjective utility $u_k(t)$ increases in direct proportion to the net outcome, $X(t)$. A value of w larger than one indicates a larger impact of net losses than net rewards on the subjective utility, whereas a value of w of one indicates equal impact of net losses and net rewards. As w approaches zero, the model predicts that net losses will be neglected.

Unlike the EV model, the PVL model assumes that, on every trial t, decision makers update the expected utilities of every deck according to the Decay learning rule (Erev & Roth, 1998). This rule discounts expectancies of every deck on every trial to an extent depending on the recency parameter a. This means that, in contrast to the EV model, the expectancies of the unchosen decks are discounted. The dummy variable contained in the learning rule, δ_k, ensures that only the current utility of the chosen deck k is added to the expectancy of that deck. A small value of a indicates rapid forgetting and strong recency effects, whereas a large value of a indicates slow forgetting and weak recency effects.

The PVL model assumes a trial-independent sensitivity parameter θ, which depends on the final model parameter: the response consistency c. Small values of c cause a random choice pattern, whereas large values of c cause a deterministic choice pattern.

In sum, the PVL model has four parameters: (1) The shape parameter A, which determines the shape of the utility function, (2) the loss aversion parameter w, which quantifies the weight of net losses over net rewards, (3) the recency parameter a, which determines the memory for past expectancies, and (4) the response consistency parameter c, which determines the amount of exploitation versus exploration.

1.2.3. The PVL-Delta model. The PVL-Delta model is a hybrid version of the EV and PVL models because it uses the Delta learning rule of the EV model (Rescorla & Wagner, 1972), but all remaining equations of the PVL model (i.e., the Prospect Utility function and the trial-independent sensitivity parameter; Ahn et al., 2008; Fridberg et al., 2010). This construction results in a model with four parameters: (1) The shape parameter A, which determines the shape of the utility function, (2) the loss aversion parameter w, which quantifies the weight of net losses over net rewards, (3) the updating parameter a, which determines the memory for past expectancies, and (4) the response consistency parameter c, which determines the amount of exploitation versus exploration.

1.3 Methods to Assess Performance of RL Models

In this section, we review methods that previous studies have used to assess performance of RL models. We differentiate between applied studies (i.e., studies that fit an RL model to IGT data to compare model parameters across groups) and model comparison studies (i.e., studies that search for the best performing model among a set of competitor models). First of all, it is evident that many applied studies take model adequacy for granted; only about two thirds of the applied literature assessed model performance at all (Steingroever, Wetzels, & Wagenmakers, 2013). The standard measure to assess model performance has been the conventional fit index BIC or G^2. This index is a relative measure that compares the performance of two models (i.e., the accuracy of one-step-ahead predictions when provided with intermediate feedback on the observed choices and payoffs); the first model is an RL model that aims to explain trial-to-trial dependencies and learning effects; the second model is a baseline model that assumes constant choice probabilities across all trials (equal to the individual's overall choice proportions from each deck). This method is also called post hoc fit criterion or one-step-ahead prediction method (see for example, Farah, Yechiam, Bekhor, Toledo, & Polus, 2008; Yechiam et al., 2005, 2008). We call this measure post hoc *relative* fit criterion in the remainder of the article to stress the comparison against a baseline model.

The disadvantage of the post hoc relative fit criterion is that it is a relative measure, and thus provides no information on whether a given model is able to account for the data; relative performance measures can only be used to investigate whether a given model outperforms a reference model, but not to investigate whether it performs adequately in absolute terms. Thus, it is possible that a particular model makes more accurate one-step-ahead predictions than the reference model (i.e., a better performance according to the post hoc relative fit criterion), but nonetheless provides a poor fit to the data.

In contrast, methods used by model comparison studies cover a wider range of meticulous and sophisticated procedures of model checking. Nevertheless, they also used the post hoc relative fit (see for example Ahn et al., 2008; Busemeyer & Stout, 2002; Fridberg et al., 2010; Worthy, Hawthorne, & Otto, 2013; Yechiam & Busemeyer, 2005; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008). Since these studies used different RL models and different tasks (i.e., the IGT, but also gambling tasks with only two or three alternatives), it is difficult to draw strong conclusions from these studies, especially because the findings are rather equivocal: The study of Busemeyer and Stout (2002) showed that, among two competitor models, the EV model had the best post hoc relative fit. Fridberg et al. (2010), on the other hand, showed that the PVL-Delta model had a better post hoc relative fit than the EV model. In addition, Fridberg et al. (2010) mentioned that their main conclusions were not affected by whether they used the PVL model with Decay learning rule (labeled PVL model in this article) or the PVL-Delta model (see their footnote 1). Other studies, however, showed that models with a Decay learning rule resulted in a better post hoc relative fit than models with a Delta learning rule (Ahn et al., 2008; Worthy et al., 2013; Yechiam & Busemeyer, 2005; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008).

Model comparison studies have also investigated whether RL models can make generalizable predictions, that is, accurate predictions for experimental conditions that differ from the original ones (i.e., a different payoff sequence or task; see Busemeyer & Wang, 2000; Pitt, Kim, & Myung, 2003, for the importance of this type of tests). The least demanding test, that is, the test with the smallest difference from the original experiment, is the simulation method (e.g., Ahn et al., 2008; Fridberg et al., 2010; see also Laud & Ibrahim, 1995). This method assesses a model's ability to generate the observed choice pattern with parameter values obtained from model fitting. More specifically, the parameter estimates from model fitting are used to generate predictions for another payoff sequence that could have been observed (i.e., the underlying payoff structure remains the same, but the exact ordering of immediate wins and losses may differ). Typically, simulation performance is assessed by comparing the predicted choice probabilities from each deck averaged across all trials to the observed choice proportions from each deck averaged across all trials (Ahn et al., 2008; Fridberg et al., 2010; Worthy et al., 2013; but see Yechiam & Busemeyer, 2005). All studies that used the simulation method have shown that the EV model has poor simulation performance. In particular, Fridberg et al. (2010), Worthy et al. (2013), and Yechiam and Busemeyer (2005) pointed out that the EV model fails to generate a preference for the decks with infrequent losses over the decks with frequent losses (see also the parameter space partitioning study of Steingroever, Wetzels, & Wagenmakers, 2013). The PVL-Delta model, on the other hand, seems to be a model with good simulation performance (Ahn et al., 2008; Fridberg et al., 2010).

A more challenging test is the so called test of generalizability. This method assesses a model's predictions for a second, different task. This method can be implemented as a relative assessment (i.e., compared to a baseline model that makes random predictions for every trial; see Ahn et al., 2008; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008) or as an absolute assessment (i.e., compared to the observed choice proportions on the second task; see Ahn et al., 2008; Yechiam & Busemeyer, 2005).[1] In addition, model comparison studies also used parameter consistency tests (i.e., Yechiam & Busemeyer, 2008) –a method that compares the correlations between model parameters estimated in different tasks– and parameter space partitioning (PSP; Steingroever, Wetzels, & Wagenmakers, 2013). PSP assesses all choice patterns that a given model can generate over its entire parameter space.

Unfortunately, the model comparison studies mentioned above failed to identify an RL model that uniquely outperforms its competitors across the various methods and data sets. First, method dependency is apparent in the studies of Ahn et al. (2008), Yechiam and Ert (2007), and Yechiam and Busemeyer (2008): These studies showed that models with the Decay learning rule produced a better post hoc relative fit, whereas models with the Delta learning rule produced better long-term generalizability and higher parameter consistency (see also Erev & Haruvy, 2005). According to Yechiam and Ert (2007) and Yechiam and Busemeyer (2008), the better post hoc relative fit of the Decay learning rule is due to mimicry of past choices. The Delta learning rule, on the hand, relies more on past payoffs instead of past choices and therefore produces better generalizable predictions and

[1]Note that the procedure of Yechiam and Busemeyer (2005) cannot be considered as a strong test of generalization because the two tasks are relatively similar (i.e., both tasks used an implementation of the IGT, but the payoff scheme used in the second task differs by a constant factor of 1.5 from the payoff scheme of the first task).

parameter consistency. In line with our results below, this suggests that the Delta learning rule measures stable characteristics of an individual more successfully than does the Decay learning rule.

Second, data set dependency is apparent in the study of Fridberg et al. (2010): In their control group of healthy participants, the PVL-Delta model resulted in a better post hoc relative fit than the EV model and the Bernoulli baseline model; however, in the case of their experimental group of chronic cannabis abusers, the baseline model outperformed the EV model and the PVL-Delta model. Fridberg et al. (2010) explained the superiority of the baseline model by arguing that the experimental group does not learn on the IGT as indicated by a stable preference for the good decks across trials. However, a stable preference for the good decks may hide changes in deck preferences that occur on the level of individual decks; an inspection of the mean choice proportions from all decks separately suggests substantial changes across trials in the popularity of decks B and D (see Figure 2 in Fridberg et al., 2010).

Other examples of data set dependency include the studies of Ahn et al. (2008) and Yechiam and Busemeyer (2005). Participants of the former study showed a preference for the good decks on the IGT, whereas participants of the latter study showed a preference for the decks with infrequent losses. Yechiam and Busemeyer (2005) found that models with the Decay-RL rule and softmax choice rule predicted performance on a second task better than models with the Delta-RL rule and softmax choice rule. However, Ahn et al. (2008) used a similar test for a different data set and found the opposite result.[2]

Data set dependency has also been confirmed by the PSP study of Steingroever, Wetzels, and Wagenmakers (2013) showing that the EV model, PVL model, and a modified version of the EV model (i.e., the EV model with Prospect Utility function) all fail to

[2]It may be argued that Yechiam and Busemeyer (2005)'s implementation of the generalizability test is rather a simulation method because the two tasks resemble each other strongly (i.e., both tasks used an implementation of the IGT, but the payoff scheme used in the second task differs by a constant factor of 1.5 from the payoff scheme of the first task). However, the question whether Yechiam and Busemeyer (2005)'s method should be classified as a test of generalizability or as a simulation method does not affect our point that results of previous model comparison studies may depend on the data set analyzed because Ahn et al. (2008) showed that models with Delta-RL rule performed better than models with Decay-RL rule on both the simulation method and test of long-term generalizability.

generate the entire spectrum of choice patterns that are typically observed in experiments. In particular, the EV model fails to generate a pronounced preference for the decks with infrequent losses (see also Fridberg et al., 2010; Yechiam & Busemeyer, 2005)—a choice pattern that is often observed in healthy participants (e.g., Caroselli, Hiscock, Scheibel, & Ingram, 2006; Dunn, Dalgleish, & Lawrence, 2006; MacPherson, Phillips, & Della Sala, 2002; Lin, Chiu, Lee, & Hsieh, 2007; Steingroever, Wetzels, Horstmann, et al., 2013; Wilder, Weinberger, & Goldberg, 1998; Yechiam & Busemeyer, 2005). Such a dependency of the models' performance on the observed choice pattern presents a crucial limitation because a good RL model for the IGT should be able to generate choice patterns present in all groups that are typically tested on the IGT.

To sum up, previous applied studies and model comparison studies used a wide variety of methods to assess performance of RL models. Applied studies typically focused on relative measures, even though assessing absolute model performance is essential to confirm model adequacy and to legitimize inferences drawn from model parameters. We will therefore propose two straightforward and general methods that allow a relatively thorough assessment of absolute model performance. In addition, we will shed light on why results of previous model comparison studies may depend on the method and data set used.

2. Performance of the EV, PVL, and PVL-Delta Models

2.1. Methods

We fit the EV, PVL, and PVL-Delta models using a Bayesian hierarchical estimation procedure (detailed in section 2.1.1) to two data sets that were constructed from our IGT data pool of healthy participants (Steingroever, Wetzels, Horstmann, et al., 2013).[3] For the first data set, we selected 31 healthy participants with a pronounced preference for the good decks (i.e., participants with at least 75% choices from the good decks, $(C + D) \geq .75$); for

[3]See Steingroever, Wetzels, Horstmann, et al. (2013) for a description of the data sets. Note that we did not use the data of Fernie and Tunney (2006), Fridberg et al. (2010), Rodríguez-Sánchez et al. (2005), and Toplak et al. (2005) to construct the stylized data sets because we have received their data either only in bins of several trials, because we did not receive information on the payoff of each participant, or because fewer than 100 IGT trials were recorded.

the second data set, we selected 31 healthy participants with a pronounced preference for the decks with infrequent losses (i.e., participants with at least 75% choices from the decks with infrequent losses, $(B + D) \geq .75$).[4] All participants completed a 100-trial IGT.

We chose these two types of choice patterns because the first type is in line with Bechara et al. (1994)'s assumptions about the performance of healthy participants on the IGT. The second type goes against Bechara et al. (1994)'s assumptions, but it is frequently observed in healthy participants (see for example, Caroselli et al., 2006; Chiu & Lin, 2007; Chiu et al., 2008; Dunn et al., 2006; Fridberg et al., 2010; MacPherson et al., 2002; Lin et al., 2007; Steingroever, Wetzels, Horstmann, et al., 2013; Wilder et al., 1998; Yechiam & Busemeyer, 2005). Since healthy participants are typically used as a control group, it is important that the models can account for these two types of choice patterns.

By using the cutoff of .75 to construct the two groups (i.e., $(C+D) \geq .75$ or $(B+D) \geq .75$), we ensured that participants within each of these groups have similar deck preferences. This procedure minimizes the impact of individual differences within each group, creating optimal conditions for precise parameter estimation in the Bayesian hierarchical framework.

To visualize the representativeness of our two stylized data sets, Figure 1 displays the proportions of choice patterns shown by participants in our IGT data pool considered for this article (total $N = 359$). For this figure, we defined five different types of choice patterns based on the deck rank order: (1) Preference for the decks with infrequent losses (i.e., $\{B, D\} \succ \{A, C\}$); (2) preference for the good decks (i.e., $\{C, D\} \succ \{A, B\}$); (3) preference for the bad decks (i.e., $\{A, B\} \succ \{C, D\}$); (4) preference for the decks with frequent losses (i.e., $\{A, C\} \succ \{B, D\}$); (5) remaining choice patterns. From the figure it is evident that the choice patterns "preference for the decks with infrequent losses" and "preference for the good decks" are most central in our IGT data pool considered for this article. Even though we chose an arbitrary cutoff value of .75 to construct the two groups, Figure 1 suggests that the construction of the two groups is empirically well founded (i.e., 32.6% ($N = 31$) and 38.3% ($N = 57$) of the participants with a preference for the decks with

[4]These participants were selected at random out of all participants showing a pronounced preference for the decks with infrequent losses.

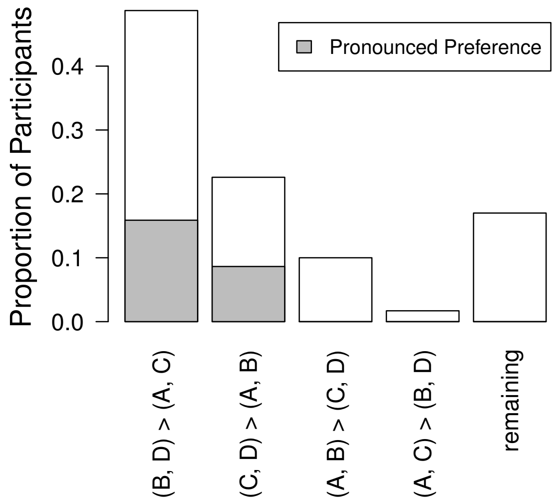


Figure 1. Proportions of choice patterns shown by participants of our IGT data pool considered for this article (total $N = 359$). The shaded areas represent the proportions of participants with pronounced deck preferences (i.e., $(B + D) \geq .75$ or $(C + D) \geq .75$).

infrequent losses and for the good decks, respectively, show a pronounced deck preference).

To assess the models' performance in absolute terms, we used two different methods: the post hoc absolute fit method and the simulation method. These two methods allow us to assess the models' ability to fit and generate the choice patterns present in the two stylized data sets. Our implementation of both methods relies on visually contrasting – separately for each deck as a function of 10 bins– the observed mean choice proportions from the experiment against the mean choice probabilities from a particular model. For the data sets at hand a visual inspection is sufficient; a more formal approach is provided by posterior predictive p-values (Gelman, Meng, & Stern, 1996; Meng, 1994; but see Bayarri & Berger, 1999, 2000).

The difference between the two methods lies in how the choice probabilities from a particular model were obtained (see Appendix for detailed recipes). Both methods start by sampling parameter values from the joint posterior distributions over the individual-level parameters (hereafter individual-level joint posteriors). For a given participant i, this sample represents a parameter value combination $\{w_i, a_i, c_i\}$ in the case of the EV model, and $\{A_i, w_i, a_i, c_i\}$ in the case of the PVL and PVL-Delta models. This parameter value combination is then provided to the model. In the case of the post hoc absolute fit method, the model is also provided with the actual choices and payoffs of participant i. Based on the information on the observed choices and payoffs up to and including the current trial, the post hoc absolute fit method computes the probability of choosing each deck on the next trial. The simulation method, on the other hand, relies on generating choices for another sequence of payoffs that could have been observed.[5] In particular, on each trial, the simulation method generates a choice based on the predicted choice probabilities. The model then uses the payoff that corresponds to the generated choice to compute the utility of the chosen deck, it updates the expected utilities of the decks, and computes the probability of choosing each deck on the next trial. These probabilities are then used to generate the next choice. Thus, the simulation method spawns synthetic participants who are confronted with the IGT just as the human participants. For both methods and for each participant, we repeated the process of obtaining the predicted choice probabilities 100 times to account for uncertainty in the individual-level joint posteriors.

We consider a model fit adequate whenever the observed choice proportions match the choice probabilities calculated from the model with access to the observed sequence of choices and payoffs (i.e., an adequate post hoc absolute fit). Similarly, we consider a model's predictions adequate whenever the observed choice proportions match the choice probabilities generated by the model without access to the observed sequence of choices and payoffs (i.e., an adequate simulation performance).

Our implementation of the post hoc absolute fit method and the simulation method differ from previously used tests in the following ways. First, our post hoc absolute fit method entails an absolute comparison to the data instead of a relative comparison to a baseline model (i.e., the post hoc relative fit; but see Busemeyer & Stout, 2002; Wood, Busemeyer, Koling, Cox, & Davis, 2005, for an absolute presentation featuring only the

[5]Note that we used the same payoff schedule as in the corresponding experiment.

good decks). Second, our implementation of the simulation method considers the choice probabilities for each deck and trial separately, instead of averaging across all trials (see Ahn et al., 2008; Fridberg et al., 2010; Worthy et al., 2013). Yechiam and Busemeyer (2005) also used this implementation, but they did not account for uncertainty in the parameter estimates.

To investigate whether our conclusions hold more generally, we also fit the EV, PVL, and PVL-Delta models to five complete data sets presented in the IGT review article of Steingroever, Wetzels, Horstmann, et al. (2013; see therein for further details on the data sets). These data sets were received from the authors upon request.[6]

2.1.1. Bayesian hierarchical estimation procedure. To fit the EV, PVL, and PVL-Delta models to the data, we used a Bayesian hierarchical estimation procedure (see Ahn et al., 2011; Wetzels, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2010, for advantages of the Bayesian hierarchical approach). The Bayesian graphical PVL (and PVL-Delta) model for a hierarchical analysis is shown in Figure 2. The Bayesian graphical EV model looks very similar; the only difference is that the EV model has one fewer parameter, that parameter w_i is immediately drawn from a group-level distribution instead of being obtained from w_i', and that the sensitivity parameter is trial-dependent (i.e., $\theta_{i,t}$). Figure 2 shows that the graphical model consists of two plates: The inner plate expresses the replications of the choices on $t = 1, \ldots, T$ trials of the IGT, and the outer plate expresses the replications for $i = 1, \ldots, N$ participants. For the sake of clarity, we omitted the notation that indexes the deck number k. The quantities $W_{i,t}$ (rewards of participant i on trial t), $L_{i,t}$ (losses of participant i on trial t), and $Ch_{i,t+1}$ (choice of participant i on trial $t+1$) can directly be obtained from the data, and the quantities $u_{i,t}$, $Ev_{i,t+1}$, and θ_i can be calculated with the equations presented in Table 2. Each individual-level parameter vector z_i, that is $\{w_i, A_i, a_i, c_i\}$ in the case of the PVL and PVL-Delta models, and $\{w_i, a_i, c_i\}$ in the case of

[6]Note that we did not fit the models to the data sets of Fernie and Tunney (2006), Rodríguez-Sánchez et al. (2005), and Toplak et al. (2005) because we received their data only in bins of several trials or because we did not receive information on the payoff of each participant. The data set labeled as "own data set" in Steingroever, Wetzels, Horstmann, et al. (2013) is here labeled as "Horstmann" because Annette Horstmann collected the data.

the EV model, is assumed to be drawn from a group-level beta distribution, Beta(α_z, β_z). Since beta distributions are restricted to the $[0, 1]$ interval, we transformed parameters with different ranges (see Table 2) to the $[0, 1]$ interval, and only transformed them back to their correct ranges after the analysis was complete. Beta distributions are typically defined by two shape parameters α and β. Here we reparameterize the two shape parameters in terms of the group-level mean μ_z and group-level precision λ_z as follows:

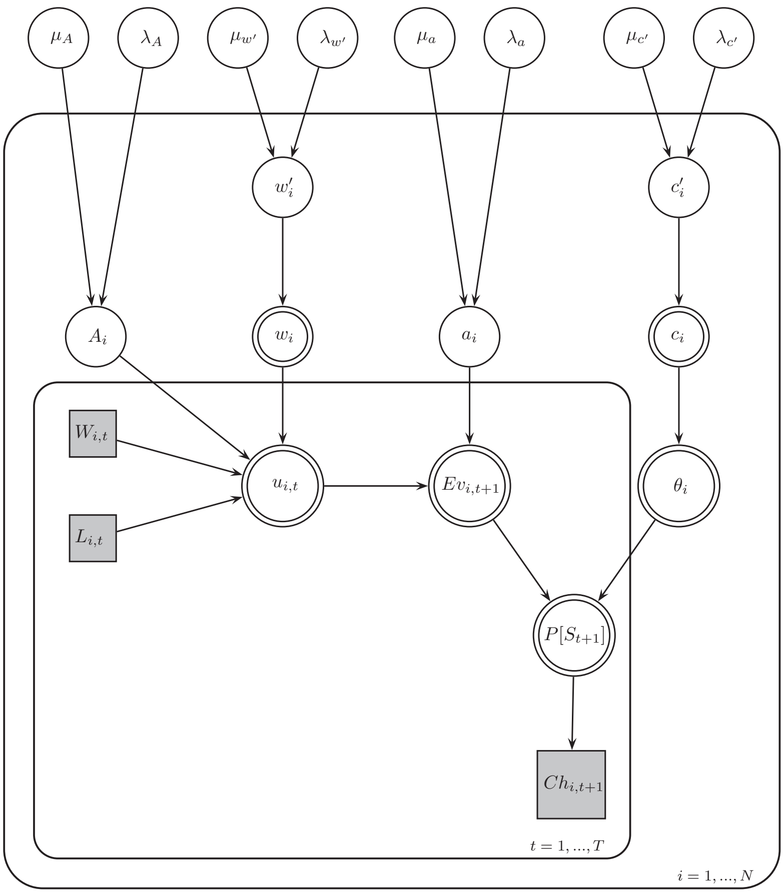
$$\alpha_z = \mu_z \lambda_z \tag{1}$$

$$\beta_z = \lambda_z (1 - \mu_z) \tag{2}$$

We assigned a uniform prior to the group-level means, $\mu_z \sim \text{U}(0, 1)$, and to the logarithm of the group-level precisions, $\log(\lambda_z) \sim \text{U}(\log(2), \log(600))$. Setting the lower limit of the prior on $\log(\lambda_z)$ to $\log(2)$ prevents the beta group-level distributions from being bimodal (Beta distribution, n. d.). However, to prevent numerical problems in the estimation program we had to increase this lower limit to a maximum of 21 for the most challenging stylized data set and to a maximum of 31 for the most challenging complete data set—a modification that can reduce the variance of the group-level distributions; here this increase had little effect on our inferences as the posterior distributions of the group-level precision parameters were not cut off at their lower limit.

We implemented the EV, PVL, and PVL-Delta models in the WinBUGS Development Interface (WBDev, Lunn, 2003)—an add-on program to WinBUGS (BUGS stands for Bayesian inference Using Gibbs Sampling; Lunn, Jackson, Best, Thomas, & Spiegelhalter, 2012). The advantage of WBDev over WinBUGS is that WBDev allows the implementation of user-defined functions and distributions, and requires less computational time (Wetzels, Lee, & Wagenmakers, 2010). The code for the fitting procedures of the EV, PVL, and PVL-Delta models in WBDev is available on `http://www.helensteingroever.com`.

For each parameter, we collected posterior samples using three Markov chain Monte



$$\mu_z \sim \text{Beta}(1, 1)$$
$$\log(\lambda_z) \sim \text{Uniform}(\log(2), \log(600))$$

$$w_i' \sim \text{Beta}(\mu_{w'}\lambda_{w'}, (1-\mu_{w'})\lambda_{w'})$$
$$c_i' \sim \text{Beta}(\mu_{c'}\lambda_{c'}, (1-\mu_{c'})\lambda_{c'})$$

$$A_i \sim \text{Beta}(\mu_A\lambda_A, (1-\mu_A)\lambda_A)$$
$$a_i \sim \text{Beta}(\mu_a\lambda_a, (1-\mu_a)\lambda_a)$$

Figure 2. Bayesian graphical PVL (and PVL-Delta) model for a hierarchical analysis.

Carlo (MCMC) chains that were run simultaneously. To assess whether the chains of all parameters had converged successfully from their starting values to their stationary distributions, we visually inspected the MCMC chains. In addition, we used the $\hat{R}$ statistic (Gelman & Rubin, 1992), a formal diagnostic measure of convergence that compares the between-chain variability to the within-chain variability. As a rule of thumb, values of $\hat{R}$ close to 1.0 indicate adequate convergence to the stationary distribution, whereas values greater than 1.1 indicate inadequate convergence.

We initialized all chains with different starting values that were generated from uniform distributions covering a wide range of possible parameter values (i.e., randomly overdispersed starting values). Fitting the PVL and PVL-Delta models with three chains often resulted in convergence difficulties: for instance, two chains may appear to have converged to their stationary distributions and gave the appearance of "hairy caterpillars" that are randomly intermixed, whereas the third chain behaved differently, often seemingly stuck at either the lower or upper parameter bound and consequently producing a much larger deviance (i.e., an inferior goodness of fit). In such situations we decided to run at least five chains simultaneously and to base inferences on three chains with the smallest deviance.[7] However, even this procedure resulted in convergence problems for a few participants (e.g., bimodal posterior distributions). We therefore excluded participants with such convergence issues and repeated the fitting procedure. This explains why the sample sizes presented in Table 3 are slightly smaller than stated earlier and than those reported by Steingroever, Wetzels, Horstmann, et al. (2013).

Table 3 also contains, for each data set separately, the number of samples we discarded as burn-in and the number of posterior samples that we collected for each chain. These specifications differ across data sets to ensure that all chains reached convergence. We based our inferences on these posterior samples.

2.3. Results

Visual inspection of the MCMC chains and consideration of the $\hat{R}$ statistics for all parameters suggested that all chains converged successfully (i.e., all parameters of the two stylized data sets and complete data sets had $\hat{R}$ values below 1.04 and 1.05, respectively). To illustrate how we assessed convergence visually, Figure 3 shows the chains of one individual-level parameter. From the figure it is evident that the chains have converged to their stationary distribution, giving the appearance of "fat hairy caterpillars" that are randomly

[7]Our convergence difficulties with the PVL and PVL-Delta models are not unique. Ahn et al. (2011) made available online two alternative fitting routines for the PVL model, and also reported convergence difficulties for their first code. They propose two solutions for the convergence difficulties: The first solution is the same as we proposed here, that is, basing inferences on chains that have converged successfully. The second solution is to use their second code that uses a different prior specification and model formulation in which the individual-level parameters are assumed to be drawn from truncated normal distributions. Also, Ahn et al. (2011) needed a large amount of burn-in samples and iterations to fit their data set with their first code, that is, they based their inferences on 25,000 samples that were drawn after 70,000 burn-in samples. The necessity of such a large amount of burn-in samples and iterations indicates the presence of convergence difficulties.

Table 3
Sample size of the two stylized and five complete data sets, number of samples discarded as burn-in, and number of posterior samples collected for each chain.

Data set	Sample size	EV model		PVL model		PVL-Delta model	
		Burn-in samples	Posterior samples	Burn-in samples	Posterior samples	Burn-in samples	Posterior samples
Good decks	30	2,000	2,000	1,000	5,000	3,000	2,000
Infrequent losses	31	1,000	3,000	1,000	2,333	12,000	2,000
Fridberg et al. (2010)	15	1,000	1,000	12,000	7,000	13,000	5,000
Horstmann	147	32,000	6,000	16,000	5,000	15,000	3,000
Kjome et al. (2010)	18	3,000	2,000	5,000	5,000	9,000	2,000
Premkumar et al. (2008)	25	1,000	1,000	4,000	3,800	12,000	4,333
Wood et al. (2005)	147	12,000	8,000	16,000	5,000	3,000	2,000

intermixed.

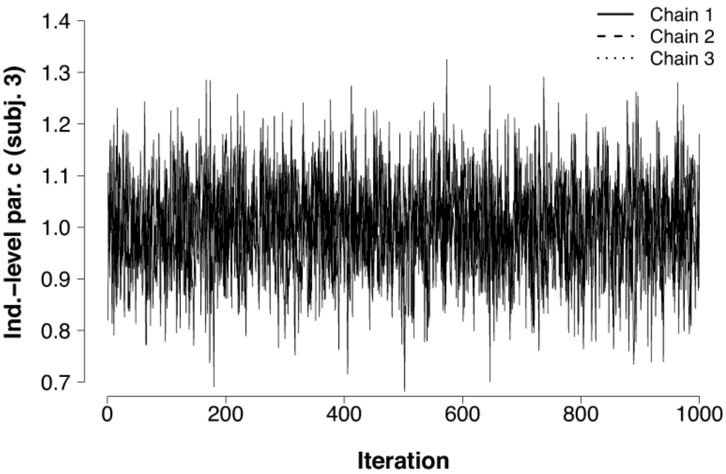


Figure 3. MCMC chains of the individual-level PVL parameter c of the third participant in the stylized data set featuring a pronounced preference for the good decks. We inspected this type of plot for every parameter to assess convergence visually, in addition to quantifying convergence through the formal diagnostic measure $\hat{R}$.

2.3.1. Ability to fit. Figure 4 presents the post hoc absolute fit of the three RL models with respect to the two stylized data sets. The first column presents the observed choice proportions from each deck as a function of 10 bins; the second, third, and fourth column present the mean probabilities of choosing each deck on each trial as calculated with the EV, PVL, and PVL-Delta models, respectively. The participants from the first

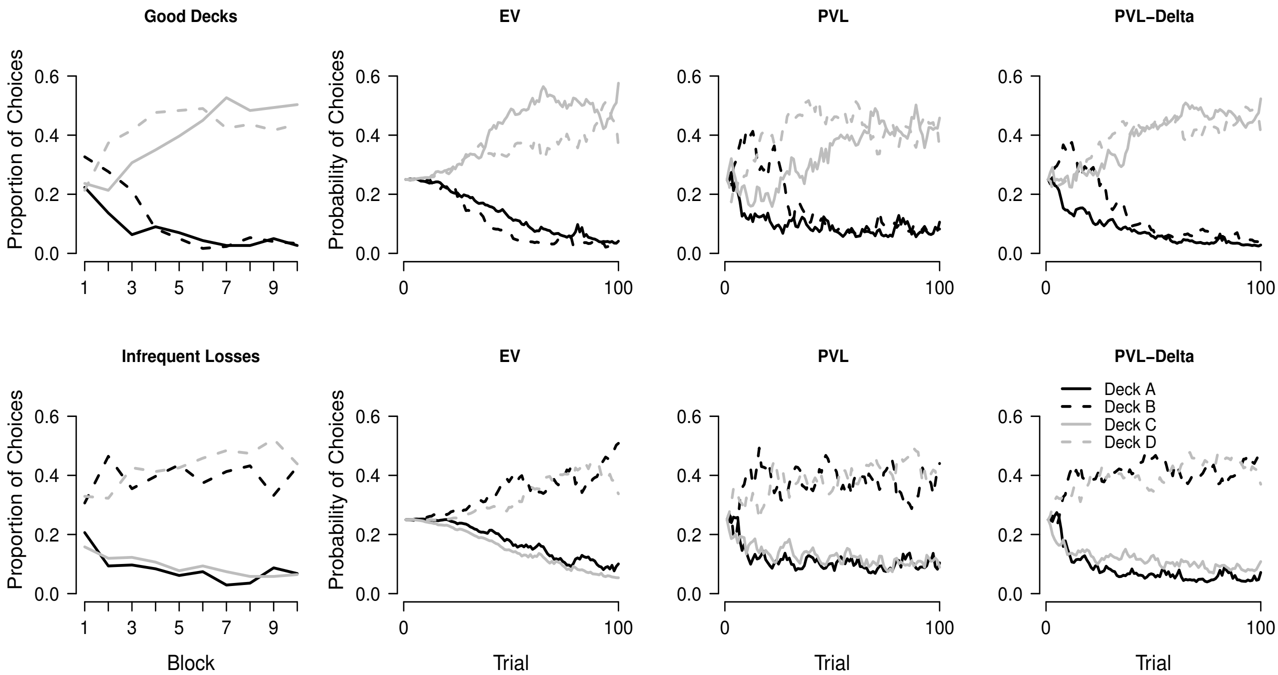


Figure 4 . Post hoc absolute fit performance of the three RL models with respect to the two stylized data sets. The first column presents the observed mean proportions of choices from each deck within 10 blocks. Each block contains 10 trials. The second, third, and fourth column present the mean probabilities of choosing each deck on each trial as calculated with the EV, PVL, and PVL-Delta model, respectively.

data set show a pronounced preference for the good decks (first panel of the first row); the participants from the second data set show a pronounced preference for the decks with infrequent losses (first panel of the second row).[8]

It is evident from Figure 4 that all models provide an acceptable fit to the observed data. All models capture the qualitative choice patterns (i.e., the rank order of the decks) shown by the two stylized data sets. In addition, the models also adequately capture the size of the choice proportions. Nonetheless, the EV model seems to fit the two stylized data sets slightly worse than the PVL and PVL-Delta models: According to the EV model, it takes a few trials at the beginning of the IGT until participants start learning and develop

[8]The deck selection profiles of all 61 participants included in the two stylized data sets can be downloaded here: `https://dl.dropbox.com/u/12798592/DeckSelectionProfilesFit.zip`. In these profiles, filled dots indicate the occurrence of rewards and losses together, whereas unfilled dots indicate the occurrence of rewards only.

the pronounced deck preferences—a pattern that is inconsistent with the observed data. But altogether, Figure 4 suggests that only small qualitative differences exist in the models' ability to fit the two stylized data sets.

To ascertain that our results generalize to other data sets, Figure 5 shows the post hoc absolute fit performance of the three RL models with respect to the five complete data sets. The first column presents the observed choice proportions from each deck as a function of 10 bins; the second, third, and fourth column present the mean probabilities of choosing each deck on each trial as calculated with the EV, PVL, and PVL-Delta models, respectively. At the behavioral level, Figure 5 illustrates that only the data set of Premkumar et al. (2008) shows a preference for the good decks. The remaining four data sets show a frequency-of-losses effect (i.e., a preference for the decks with infrequent losses)—an effect that differs in its extent across the four data sets: The data sets of Fridberg et al. (2010) and Horstmann show a pronounced frequency-of-losses effect with a clear preference for both decks with infrequent losses (i.e., decks B and D) over both decks with frequent losses, whereas the remaining two data sets show a less pronounced frequency-of-losses effect, indicating that, at the end of the IGT, participants choose about equally often from decks B, C, and D, while clearly avoiding deck A (i.e., Kjome et al., 2010; Wood et al., 2005). In general, it is evident that the choice patterns shown by the five complete data sets are less pronounced than those of the two stylized data sets presented in Figure 4.

It is evident that Figure 5 corroborates the conclusions from Figure 4: All models provide an acceptable fit to the data, but the EV model fits the five complete data sets slightly worse than the PVL and PVL-Delta models.

In addition to visually assessing the models' ability to fit the two stylized and five complete data sets, we also compared the deviance measure provided by WinBUGS of all models and data sets (Table 4). The deviance is defined as $D(\boldsymbol{\theta}) = -2 \log p(y|\boldsymbol{\theta})$, where $p(y|\boldsymbol{\theta})$ is the likelihood of the data y given the parameters $\boldsymbol{\theta}$. Thus, the smaller the deviance, the better the fit. In line with Figures 4 - 5, Table 4 shows that the EV model provides the worst fit to the two stylized and five complete data sets. In addition, the PVL model has

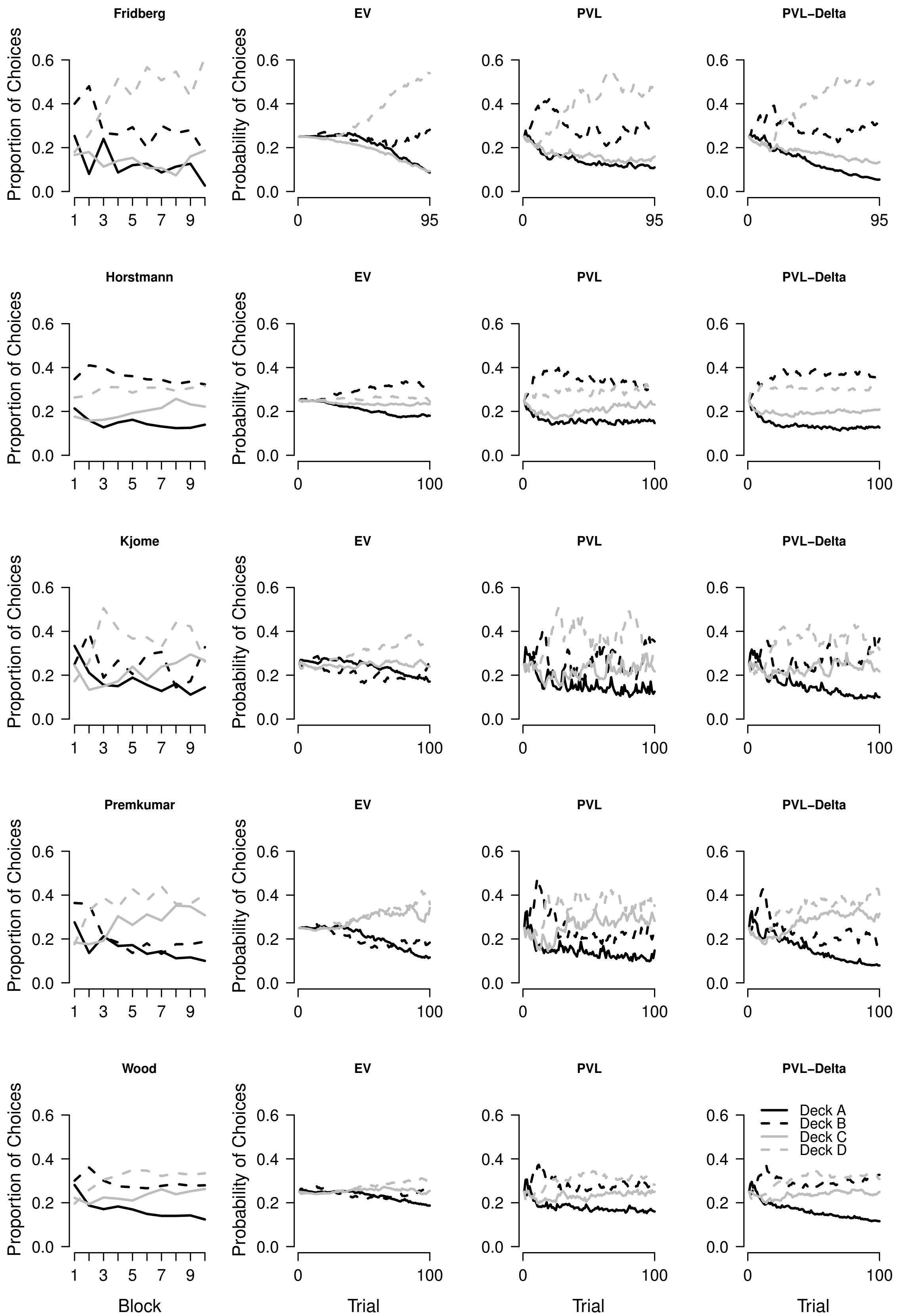


Figure 5. Post hoc absolute fit performance of the three RL models with respect to the five complete data sets. The first column presents the observed mean proportions of choices from each deck within 10 blocks. Each block contains 10 trials, except the last block of Fridberg et al. (2010, 5-trials). The second, third, and fourth column present the mean probabilities of choosing each deck on each trial as calculated with the EV, PVL, and PVL-Delta model, respectively.

Table 4
Deviance measures provided by WinBUGS of all models and data sets. For each data set, we printed in bold the lowest deviance value to identify the model with the best fit.

Data set	EV	PVL	PVL-Delta
Good decks	5,905	**4,667**	5,213
Infrequent losses	6,823	**5,450**	5,757
Fridberg et al. (2010)	3,545	3,368	**3,356**
Horstmann	37,210	33,510	**32,830**
Kjome et al. (2010)	4,597	**4,010**	4,258
Premkumar et al. (2008)	5,915	**5,223**	5,653
Wood et al. (2005)	38,390	**34,690**	35,820

a smaller deviance (i.e., a better fit) than the PVL-Delta model for five out of seven data sets.

2.3.2. Ability to generate. Figure 6 illustrates how the three models perform on the simulation method with respect to the two stylized data sets. The first column presents the observed choice proportions from each deck as a function of 10 bins (i.e., identical to the first column of Figure 4); the second, third, and fourth column present the probabilities of choosing each deck on each trial as generated with the EV, PVL, and PVL-Delta models, respectively. From the figure it is evident that neither the EV nor the PVL model succeeds to generate both observed choice patterns. Specifically, the EV model fails to generate a choice pattern featuring a pronounced preference for the decks with infrequent losses (second stylized data set), whereas the PVL model fails to generate a choice pattern featuring a pronounced preference for the good decks (first stylized data set).

In the case of the first stylized data set, the EV model correctly generates the empirical rank order of the decks. However, the model strongly overestimates the mean choice proportions from deck C. In addition, the EV model predicts that the probability of choosing deck D increases until trial 50, but then decreases to chance level—a prediction that is not in line with the data; the observed choice proportions from deck D are above chance level across all trials. The PVL model, on the other hand, performs acceptably on the simulation method in the case of the second stylized data set; it correctly generates that the decks with infrequent losses are preferred over the decks with frequent losses, but

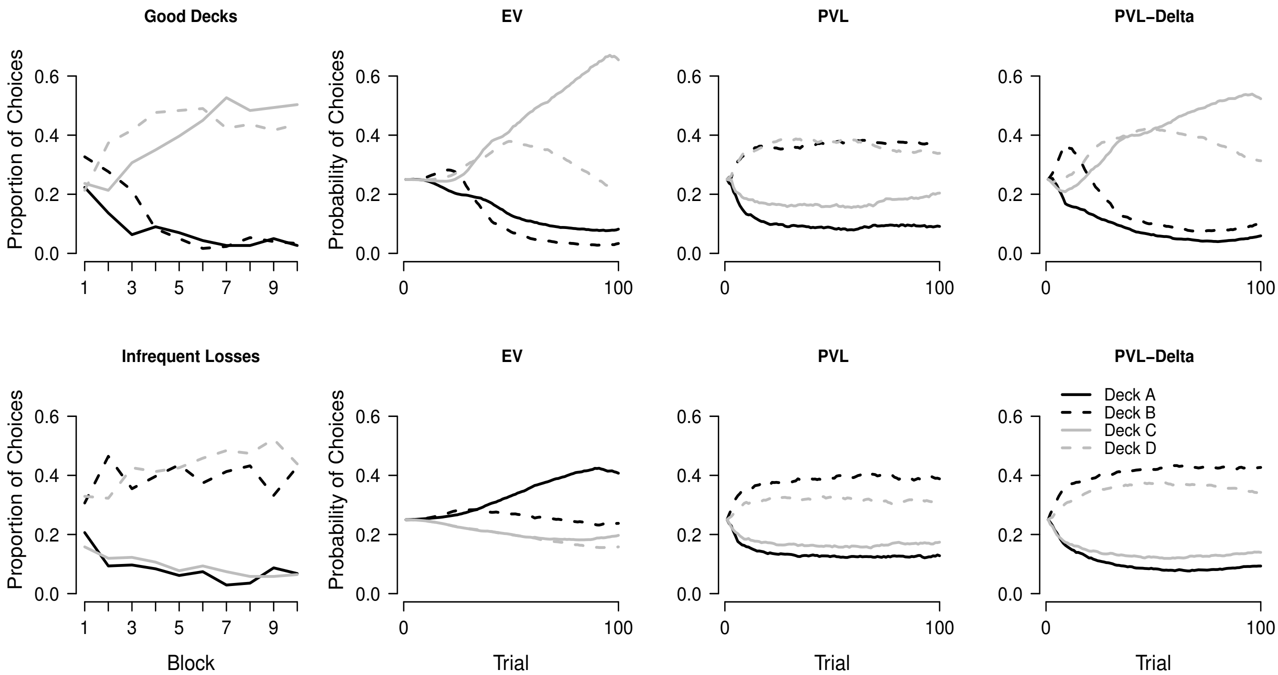


Figure 6. Simulation performance of the three RL models with respect to the two stylized data sets. The first column presents the observed mean proportions of choices from each deck within 10 blocks. Each block contains 10 trials. The second, third, and fourth column present the mean probabilities of choosing each deck on each trial as generated with the EV, PVL, and PVL-Delta model, respectively.

it fails to generate that deck D is on average preferred over deck B.

The PVL-Delta model, on the other hand, is the only model considered in this article that adequately generates the qualitative choice patterns of both stylized data sets. Yet, a few discrepancies exist between the observed and generated choice patterns: In the case of the first stylized data set, the PVL-Delta model slightly underestimates the mean choice proportions from deck D, and slightly overestimates the mean choice proportions from deck B. Just as the EV model, the PVL-Delta model predicts that the mean probability of choosing deck D increases until trial 50 and then decreases, even though the observed data do not show this decrease. In the case of the second stylized data set, the PVL-Delta model –just as the PVL model– fails to generate that deck D is on average preferred over deck B.

To ascertain that our results generalize to other data sets, Figure 7 shows how the three models perform on the simulation method with respect to the five complete data sets. The first column presents the observed choice proportions from each deck as a function of 10 bins (i.e., identical to the first column of Figure 5); the second, third, and fourth column present the probabilities of choosing each deck on each trial as generated with the EV, PVL, and PVL-Delta models, respectively.

It is evident that Figure 7 corroborates the conclusions from Figure 6: The EV model fails to generate choice patterns featuring a preference for the decks with infrequent losses as shown by four data sets; for these data sets the EV model makes almost random predictions. Interestingly, the EV model makes very similar predictions for the data sets of Fridberg et al. (2010) and Horstmann, and the data sets of Kjome et al. (2010) and Wood et al. (2005), even though there are pronounced differences in the choice patterns at the behavioral level. In the case of the choice pattern featuring a preference for the good decks as shown by the data set of Premkumar et al. (2008), the EV model correctly predicts a preference for the good decks over the bad decks, but –as in the case of the stylized data set with a preference for the good decks– the EV model underestimates the preference for deck D and overestimates the preference for deck A.

As already suggested by Figure 6, Figure 7 underscores that the PVL model fails to generate a choice pattern featuring a preference for the good decks (i.e., as present in the data set of Premkumar et al., 2008). However, the PVL model makes acceptable predictions for the four data sets with a frequency-of-losses effect; for all four data sets the PVL model correctly generates that the decks with infrequent losses are preferred over the decks with frequent losses. Yet, it is evident that, for the data sets of Fridberg et al. (2010) and Wood et al. (2005), the PVL model fails to generate that deck D is on average preferred over deck B—a discrepancy between the data and the predictions that was already apparent in Figure 6. In addition, both Figures 6 and 7 illustrate that the PVL model generates learning curves that are relatively flat.

Moreover, as already suggested by Figure 6, the PVL-Delta model demonstrates adequate simulation performance: Figure 7 illustrates that the PVL-Delta model correctly

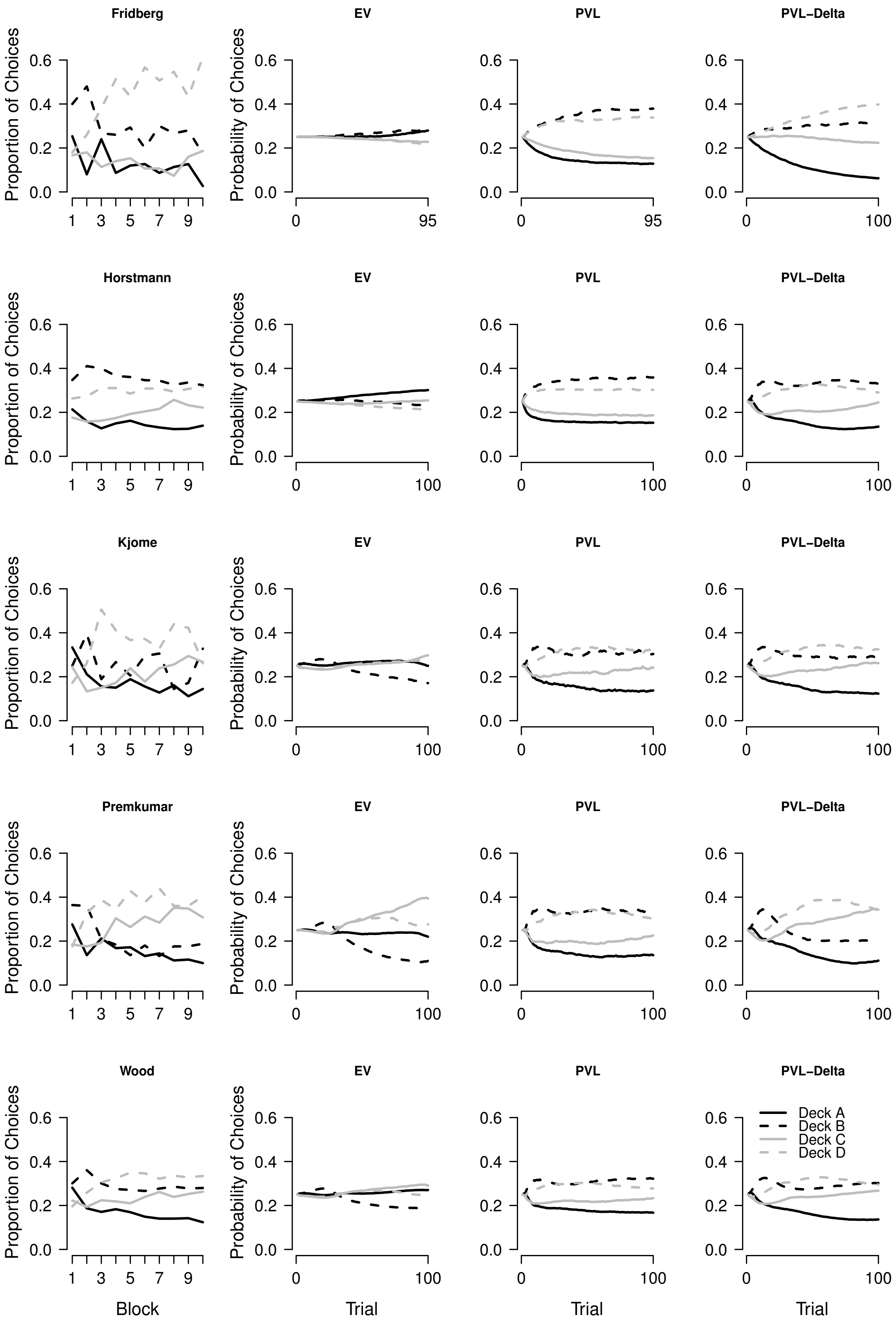


Figure 7. Simulation performance of the three RL models with respect to the two stylized data sets. The first column presents the observed mean proportions of choices from each deck within 10 blocks. Each block contains 10 trials, except the last block of Fridberg et al. (2010, 5-trials). The second, third, and fourth column present the mean probabilities of choosing each deck on each trial as generated with the EV, PVL, and PVL-Delta model, respectively.

generates the choice patterns shown by most data sets. Only in the case of the data set of Fridberg et al. (2010) does the PVL-Delta model slightly underestimate the mean choice proportions from deck D and slightly overestimate the choice proportions from deck C.

Overall, the results of this section showed that the simulation method –in contrast to the post hoc absolute fit method– allows for a good discrimination between the models. In addition, a comparison of the two methods resulted in conflicting findings: Even though all models provided an adequate fit to the observed choice patterns, only the PVL-Delta model was able to also generate these choice patterns.

4. General Discussion

In this article, we compared two methods that assess absolute model performance: the post hoc absolute fit method and the simulation method. We used these methods to investigate whether three RL models of the IGT –the popular EV and PVL models, and a hybrid version of both models, the PVL-Delta model– can fit and generate choice patterns present in two stylized and five complete data sets.

Our results showed that all models provided an acceptable fit to all data sets and that only small differences existed in the models' ability to fit the different choice patterns. Thus, our results suggest that the post hoc absolute fit method allows for limited qualitative discrimination between the models. The simulation method, on the other hand, revealed important performance differences between the models: When provided with no intermediate feedback on the observed choices and payoffs, the EV model failed to generate a choice pattern featuring a preference for the decks with infrequent losses, whereas the PVL model failed to generate a choice pattern featuring a preference for the good decks. Only the PVL-Delta model adequately generated the choice patterns shown by the two stylized and five complete data sets.

Our results clearly illustrate that a model's ability to fit a particular choice pattern does not guarantee that the model is also able to generate that same choice pattern. This conflicting finding is supported by findings from previous model comparison studies (e.g., Ahn et al., 2008; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008, see Lewandowsky, 1995, for a similar phenomenon). Specifically, Yechiam and Ert (2007) and Yechiam and Busemeyer (2008) compared two RL models that only differed in the learning rule (i.e., either the Delta learning rule or the Decay learning rule) using post hoc relative fit, a long-term generalization test, and a parameter consistency test. In both studies, the model with the Decay learning rule had a better post hoc relative fit, but the model with the Delta learning rule performed better on the latter two tests. The authors explain these conflicting findings by arguing that the Decay learning rule produces a better post hoc relative fit because it relies more on past choices (i.e., mimicry of past choices), whereas the Delta learning rule relies more on past payoffs. According to these authors, the increased reliance on past payoffs explains why the Delta learning rule is superior in producing predictions that generalize to other tasks, and parameters that are consistent across different tasks. These results relate to ours because post hoc relative fit and post hoc absolute fit share the same foundation; both methods assess the accuracy of an RL model for the exact sequences of observed choices and payoffs. The only difference is that the post hoc relative fit compares the model's accuracy to that of a baseline model, whereas the post hoc absolute fit features an absolute comparison to the observed choice proportions from each deck.

However, in contrast to earlier work (Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008), our results suggest that a model's generalizability is determined not only by the learning rule but rather by the combination of different model equations. In particular, even though both the EV and the PVL-Delta model use the Delta learning rule, only the PVL-Delta model adequately generated all choice patterns considered in this article (see Fridberg et al., 2010; Yechiam & Busemeyer, 2005; Worthy et al., 2013, for studies that also report poor simulation performance of the EV model).

It should be noted that, in this article, we did not rule out the possibility that other parameter combinations may result in better simulation performance: to assess simulation performance, we used parameter values that were obtained with a likelihood-based estimation procedure that optimizes the fit for the exact sequences of observed choices and payoffs. Nevertheless, at least in the case of the EV model, we can be certain that no matter which parameter combination we choose the EV model will never generate a frequency of losses effect (see the PSP study of Steingroever, Wetzels, & Wagenmakers, 2013, and the simulation performance of the EV model reported by Fridberg et al., 2010, Worthy et al., 2013, and Yechiam & Busemeyer, 2005). However, instead of searching a model's entire parameter space for those parameter values that produce the best simulation performance, it is conventional to assess simulation performance with parameter values obtained from a likelihood-based estimation procedure because researchers typically base their inferences on these parameter values when they wish to draw conclusions about psychological processes underlying performance on the IGT.

Our results suggest that, among the two methods compared in this article, the simulation method is more indicative of whether or not a model captures psychological processes underlying the IGT: "the goal of model selection is to choose the model that generalizes best across all samples, because the one that does has probably captured the cognitive process of interest, and not the random fluctuations (i.e., error) that any one sample will exhibit. This is the essence of generalizability, and a model should be judged on its ability to generalize correctly, not on its adeptness (i.e., flexibility) in fitting only the data in hand." (Pitt et al., 2003, p. 31). Thus, the risk is that a good descriptive adequacy (i.e., a good post hoc absolute fit) is caused by choice mimicry; it is possible that a model strongly relies on past choices instead of past payoffs when making one-step-ahead predictions. Thus, our results suggest that models can fine-tune their parameters to obtain an accurate fit for the exact sequences of observed payoffs and choices, but a model's ability to make accurate one-step-ahead predictions cannot be taken as sufficient evidence to decide whether or not the model has successfully estimated psychological processes that drive performance on the IGT—an ambition that applied studies typically have. This also means that the conventional BIC or G^2 fit index is insufficient to decide whether model parameters are a valid reflection of psychological processes (see also Laud & Ibrahim, 1995).

Instead of using the conventional fit index as the standard measure of model performance in applied studies, our results suggest that applied researchers should carefully assess *absolute* model performance to avoid premature conclusions about the psychological processes that drive performance on the IGT. In particular, the simulation method seems to represent a more stringent and challenging test of absolute model performance than the post hoc absolute fit method because the simulation method relies on predicting the entire sequence of choices for another payoff sequence that could have been observed. Since one assumes that participants show a similar choice pattern on the IGT independently of the exact ordering of the payoffs, a model for the IGT should be able to make accurate predictions for a new payoff sequence, especially because the changes in the payoff sequence are trivial (i.e., the underlying payoff structure remains the same, but the exact ordering of immediate wins and losses differs): "It seems clear that good models, among those under consideration, should make predictions close to what has been observed for an identical experiment." (Laud & Ibrahim, 1995; p. 249). A requirement for accurate predictions is that the model is sensitive to the payoff—not to previous choices. Our advice for future applications of RL models to IGT data is therefore that both proposed tests should pass a minimum threshold of adequacy.

It stands to reason that model performance cannot be summarized with only one measure. Previous model comparison studies proposed other sophisticated and sound methods to assess model performance; in particular, to investigate whether a given model captures the underlying decision-making processes (e.g., parameter consistency, generalization to another task, test of specific influence; see for example Ahn et al., 2008; Wetzels, Vandekerckhove, et al., 2010; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008). Even though we support these additional methods, they require data from another task and hence it may be not be realistic to advocate their use in applied work. Thus, we recommend applied researchers to choose a model based on results from previous model comparison studies that used these tests, and then to use the post hoc absolute fit method and the simulation method to assess absolute model performance. If both methods pass a minimum threshold of adequacy, we can be relatively confident that conclusions from model parameters are trustworthy.

Our results suggest that in future applications of the RL models to IGT data, researchers should carefully assess absolute model performance using the post hoc absolute fit method and especially the simulation method. Only a careful assessment of absolute model performance will help prevent applied researchers from drawing conclusions that may be unwarranted and premature. Our results also suggest that future studies should consider applying the PVL-Delta model instead of the popular EV and PVL models.

Acknowledgements

We thank Annette Horstmann, Fridberg et al. (2010), Kjome et al. (2010), Premkumar et al. (2008), and Wood et al. (2005) for providing the data used in this article. We are grateful to Woo-Young Ahn, Jerome Busemeyer, and Eldad Yechiam for their constructive comments on a previous version.

References

Agay, N., Yechiam, E., Carmel, Z., & Levkovitz, Y. (2010). Non-specific effects of methylphenidate (Ritalin) on cognitive ability and decision-making of ADHD and healthy adults. *Psychopharmacology*, *210*, 511 - 519.

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.

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience Psychology and Economics*, *4*, 95 - 110.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*, 215 - 233.

Bayarri, M., & Berger, J. O. (1999). Quantifying surprise in the data and model verification. In A. P. D. J. M. Bernardo J. O. Berger & A. F. M. Smith (Eds.), *Bayesian Statistics* (Vol. 6, p. 53 - 82). Oxford University Press, Oxford.

Bayarri, M., & Berger, J. O. (2000). P values for composite null models. *Journal of the American Statistical Association*, *95*, 1127 - 1142.

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

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*, 1293 - 1295.

Beta distribution. (n. d.). *In Wikipedia.* Retrieved January 17, 2013, from `http://en.wikipedia.org/wiki/Beta_distribution`.

Blair, R. J. R., Colledge, E., & Mitchell, D. G. V. (2001). Somatic markers and response reversal: Is there orbitofrontal cortex dysfunction in boys with psychopathic tendencies? *Journal of Abnormal Child Psychology*, *29*, 499 - 511.

Brambilla, P., Perlini, C., Bellani, M., Tomelleri, L., Ferro, A., Cerruti, S., ... Frangou, S. (2012). Increased salience of gains versus decreased associative learning differentiate bipolar disorder from schizophrenia during incentive decision making. *Psychological Medicine*, *1*, 1 - 10.

Busemeyer, J. R., Stout, J., & Finn, P. (2003). Using computational models to help explain decision making processes of substance abusers. In D. Barch (Ed.), *Cognitive and affective neuroscience of psychopathology.* New York: Oxford University Press.

Busemeyer, J. R., & 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.

Caroselli, J. S., Hiscock, M., Scheibel, R. S., & Ingram, F. (2006). The simulated gambling paradigm applied to young adults: An examination of university students' performance. *Applied Neuropsychology*, *13*, 203 - 212.

Cavedini, P., Riboldi, G., D'Annucci, A., Belotti, P., Cisima, M., & Bellodi, L. (2002). Decision-making heterogeneity in obsessive-compulsive disorder: Ventromedial prefrontal cortex function predicts different treatment outcomes. *Neuropsychologia*, *40*, 205 - 211.

Cavedini, P., Riboldi, G., Keller, R., D'Annucci, A., & Bellodi, L. (2002). Frontal lobe dysfunction in pathological gambling patients. *Biological Psychiatry*, *51*, 334 - 341.

Chiu, Y.-C., & Lin, C.-H. (2007). Is deck C an advantageous deck in the Iowa gambling task? *Behavioral and Brain Functions*, *3*, 37.

Chiu, Y.-C., Lin, C.-H., Huang, J.-T., Lin, S., Lee, P.-L., & Hsieh, J.-C. (2008). Immediate gain is long-term loss: Are there foresighted decision makers in the Iowa gambling task? *Behavioral* *and Brain Functions*, *4*, 13.

Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience & Biobehavioral Reviews*, *30*, 239 - 271.

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

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, *88*, 848 - 881.

Farah, H., Yechiam, E., Bekhor, S., Toledo, T., & Polus, A. (2008). Association of risk proneness in overtaking maneuvers with impaired decision making. *Transportation Research Part F: Traffic Psychology and Behaviour*, *11*, 313 - 323.

Fernie, G., & Tunney, R. J. (2006). Some decks are better than others: The effect of reinforcer type and task instructions on learning in the Iowa gambling task. *Brain and Cognition*, *60*, 94 - 102.

Fridberg, D. J., Queller, S., Ahn, W.-Y., Kim, W., Bishara, A. J., Busemeyer, J. R., ... Stout, J. C. (2010). Cognitive mechanisms underlying risky decision-making in chronic cannabis users. *Journal of Mathematical Psychology*, *54*, 28 - 38.

Gelman, A., Meng, X.-L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, *6*, 733 - 760.

Gelman, A., & Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*, 457 - 472.

Gullo, M. J., & Stieger, A. A. (2011). Anticipatory stress restores decision-making deficits in heavy drinkers by increasing sensitivity to losses. *Drug and Alcohol Dependence*, *117*, 204 - 210.

Johnson, S. A., Yechiam, E., Murphy, R. R., Queller, S., & Stout, J. C. (2006). Motivational processes and autonomic responsivity in Asperger's disorder: Evidence from the Iowa gambling task. *Journal of the International Neuropsychological Society*, *12*, 668 - 676.

Kjome, K. L., Lane, S. D., Schmitz, J. M., Green, C., Ma, L., Prasla, I., ... Moeller, F. G. (2010). Relationship between impulsivity and decision making in cocaine dependence. *Psychiatry Research*, *178*, 299 - 304.

Laud, P. W., & Ibrahim, J. G. (1995). Predictive model selection. *Journal of the Royal Statistical Society. Series B (Methodological)*, 247 - 262.

Lev, D., Hershkovitz, E., & Yechiam, M. (2008). Decision making and personality in traffic offenders:

A study of Israeli drivers. *Accident Analysis and Prevention*, *40*, 223 - 230.

Lewandowsky, S. (1995). Base-rate neglect in ALCOVE: A critical reevaluation. *Psychological Review*, *102*, 185 - 191.

Lin, C.-H., Chiu, Y.-C., Lee, P.-L., & Hsieh, J.-C. (2007). Is deck B a disadvantageous deck in the Iowa gambling task? *Behavioral and Brain Functions*, *3*, 16.

Luce, R. (1959). *Individual choice behavior.* New York: Wiley.

Lunn, D. (2003). Winbugs development interface (WBDev). *ISBA Bulletin*, *10*, 10 - 11.

Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis.* Boca Raton (FL): Chapman & Hall/CRC.

MacPherson, S. E., Phillips, L. H., & Della Sala, S. (2002). Age, executive function, and social decision making: A dorsolateral prefrontal theory of cognitive aging. *Psychology and Aging*, *17*, 598 - 609.

Martino, D. J., Bucay, D., Butman, J. T., & Allegri, R. F. (2007). Neuropsychological frontal impairments and negative symptoms in schizophrenia. *Psychiatry Research*, *152*, 121 - 128.

Meng, X.-L. (1994). Posterior predictive p-values. *The Annals of Statistics*, 1142 - 1160.

Pitt, M. A., Kim, W., & Myung, I. J. (2003). Flexibility versus generalizability in model selection. *Psychonomic Bulletin & Review*, *10*, 29 - 44.

Premkumar, P., Fannon, D., Kuipers, E., Simmons, A., Frangou, S., & Kumari, V. (2008). Emotional decision-making and its dissociable components in schizophrenia and schizoaffective disorder: A behavioural and MRI investigation. *Neuropsychologia*, *46*, 2002 - 2012.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning ii: Current research and theory* (p. 64 - 99). New York.

Rodríguez-Sánchez, J. M., Crespo-Facorro, B., Iglesias, R. P., Bosch, C. G.-B., Álvarez, M., Llorca, J., & Vázquez-Barquero, J. L. (2005). Prefrontal cognitive functions in stabilized first-episode patients with schizophrenia spectrum disorders: A dissociation between dorsolateral and orbitofrontal functioning. *Schizophrenia Research*, *77*, 279 - 288.

Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., & Wagenmakers, E.-J. (2013). Performance of healthy participants on the Iowa gambling task. *Psychological Assessment*, *25*, 180 - 193.

Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2013). A comparison of reinforcement-learning models for the Iowa gambling task using parameter space partitioning. *The Journal of Problem*

Solving, *5*, Article 2.

Stout, J., Busemeyer, J., Lin, A., Grant, S., & Bonson, K. (2004). Cognitive modeling analysis of decision-making processes in cocaine abusers. *Psychonomic Bulletin & Review*, *11*, 742 - 747.

Toplak, M., Jain, U., & Tannock, R. (2005). Executive and motivational processes in adolescents with attention-deficit-hyperactivity disorder (ADHD). *Behavioral and Brain Functions*, *1*, 1 - 12.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297 - 323.

Wetzels, R., Lee, M., & Wagenmakers, E.-J. (2010). Bayesian inference using WBDev: A tutorial for social scientists. *Behavior Research Methods*, *42*, 884 - 897.

Wetzels, R., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2010). Bayesian parameter estimation in the Expectancy Valence model of the Iowa gambling task. *Journal of Mathematical Psychology*, *54*, 14 - 27.

Wilder, K. E., Weinberger, D. R., & Goldberg, T. E. (1998). Operant conditioning and the orbitofrontal cortex in schizophrenic patients: Unexpected evidence for intact functioning. *Schizophrenia Research*, *30*, 169 - 174.

Wood, S., Busemeyer, J., Koling, A., Cox, C. R., & Davis, H. (2005). Older adults as adaptive decision makers: Evidence from the Iowa gambling task. *Psychology and Aging*, *20*, 220 - 225.

Worthy, D. A., Hawthorne, M. J., & Otto, A. R. (2013). Heterogeneity of strategy use in the Iowa gambling task: A comparison of win-stay/lose-shift and reinforcement learning models. *Psychonomic Bulletin & Review*, *20*, 364 - 371.

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

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

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. *Psychological Science*, *16*, 973 - 978.

Yechiam, E., & Ert, E. (2007). Evaluating the reliance on past choices in adaptive learning models.

Journal of Mathematical Psychology, *51*, 75 - 84.

Yechiam, E., Kanz, J. E., Bechara, A., Stout, J. C., Busemeyer, J. R., Altmaier, E. M., & Paulsen, J. S. (2008). Neurocognitive deficits related to poor decision making in people behind bars. *Psychonomic Bulletin & Review*, *15*, 44 - 51.

Appendix

Recipe for obtaining choice probabilities according to the post hoc absolute fit method

1. For a given participant i, take a random draw from the individual-level joint posterior (i.e., use a random chain and iteration). This random draw results in a parameter value combination (i.e., $\{w_i, A_i, a_i, c_i\}$ for the PVL and the PVL-Delta models, and $\{w_i, a_i, c_i\}$ for the EV model) that is then provided to the model. Alternatively, use the maximum likelihood estimates.

2. Initialize the expectancies of all decks to zero, $Ev_k(0) = 0$. Therefore, $P[S_k(1)] = 0.25$ for each deck k, $k \in \{1, 2, 3, 4\}$ (i.e., on the first trial, all decks are equally likely to be chosen).

3. Execute steps $4 - 7$ for trial $t = 1$ up to and including $t = T - 1$ where T is the maximum number of trials used in the corresponding experiment.

4. Provide the model with the observed choice $S_k(t)$, and payoff on trial t, $W(t)$ and $L(t)$.

5. Use the payoff observed on trial t to compute the utility of the chosen deck.

6. Update the expected utility of all decks (or only of the chosen deck, in the case of the EV and PVL-Delta models).

7. Compute the probability that deck k will be chosen on the next trial $P[S_k(t + 1)]$. Save the probabilities.

8. Repeat steps $1-7$ for each subject 100 times to account for the posterior uncertainty. This step is omitted if maximum likelihood estimates were used.

Recipe for obtaining choice probabilities according to the simulation method

1. For a given participant i, take a random draw from the individual-level joint posterior (i.e., use a random chain and iteration). This random draw results in a parameter

value combination (i.e., $\{w_i, A_i, a_i, c_i\}$ for the PVL and the PVL-Delta models, and $\{w_i, a_i, c_i\}$ for the EV model) that is provided to the model. Alternatively, use the maximum likelihood estimates.

2. Initialize the expectancies of all decks to zero, $Ev_k(0) = 0$. Therefore, $P[S_k(1)] = 0.25$ for each deck k, $k \in \{1, 2, 3, 4\}$ (i.e., on the first trial, all decks are equally likely to be chosen).

3. Execute steps $4 - 7$ for trial $t = 1$ up to and including $t = T - 1$ where T is the maximum number of trials used in the corresponding experiment.

4. Generate a choice on trial t using $P[S_k(t)]$.

5. Use the payoff corresponding to the choice on trial t to compute the utility of the chosen deck. Make sure to use the same payoff schedule as in the corresponding experiment.

6. Update the expected utility of all decks (or only of the chosen deck, in the case of the EV and PVL-Delta models).

7. Compute the probability that deck k will be chosen on the next trial $P[S_k(t + 1)]$. Save the probabilities.

8. Repeat steps $1 - 7$ for each subject 100 times to account for the posterior uncertainty. This step is omitted if maximum likelihood estimates were used.