

## A Comparison of Reinforcement Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning

*Helen Steingroever,<sup>1</sup> Ruud Wetzels,<sup>2</sup> and Eric-Jan Wagenmakers<sup>1</sup>*

### Abstract

The Iowa gambling task (IGT) is one of the most popular tasks used to study decision-making deficits in clinical populations. In order to decompose performance on the IGT in its constituent psychological processes, several cognitive models have been proposed (e.g., the Expectancy Valence (EV) and Prospect Valence Learning (PVL) models). Here we present a comparison of three models—the EV and PVL models, and a combination of these models (EV-PU)—based on the method of parameter space partitioning. This method allows us to assess the choice patterns predicted by the models across their entire parameter space. Our results show that the EV model is unable to account for a frequency-of-losses effect, whereas the PVL and EV-PU models are unable to account for a pronounced preference for the bad decks with many switches. All three models underrepresent pronounced choice patterns that are frequently seen in experiments. Overall, our results suggest that the search of an appropriate IGT model has not yet come to an end.

### Keywords

decision making, loss aversion, Expectancy Valence model,  
Prospect Valence Learning model

---

<sup>1</sup>University of Amsterdam. Please direct correspondence to [helen.steingroever@gmail.com](mailto:helen.steingroever@gmail.com).

<sup>2</sup>The European Foundation for the Improvement of Living and Working Conditions, Dublin, Ireland

The Iowa gambling task (IGT; Bechara, Damasio, Damasio, & Anderson, 1994) is a popular neuropsychological paradigm used to assess decision-making deficits in clinical populations. The IGT was originally developed to detect decision-making deficits of patients with lesions to the ventromedial prefrontal cortex (vmPFC), but in the last two decades it has been applied to a variety of clinical populations, such as patients with pathological gambling disorder (Cavedini, Riboldi, Keller, D'Annuncci, & Bellodi, 2002), attention deficit hyperactivity disorder (Agay, Yechiam, Carmel, & Levkovitz, 2010; Toplak, Jain, & Tanock, 2005), obsessive-compulsive disorder (Cavedini, Riboldi, D'Annuncci, et al., 2002), psychopathic tendencies (Blair, Colledge, & Mitchell, 2001), bipolar disorder (Brambilla et al., 2012), and schizophrenia (Martino, Bucay, Butman, & Allegri, 2007; Premkumar et al., 2008). In addition, the IGT has been applied to cocaine addicts (Stout, Busemeyer, Lin, Grant, & Bonson, 2004), chronic cannabis users (Fridberg et al., 2010), traffic offenders (Lev, Hershkovitz, & Yechiam, 2008), and inmates (Yechiam, Kanz, et al., 2008).

The task contains four decks that differ in their payoff scheme. Participants are told to choose cards so as to maximize their net profit. The only way to accomplish this goal in the IGT is to avoid risky decks that contain high immediate rewards but even higher occasional losses, and instead prefer safe decks that contain small immediate rewards but even smaller occasional losses. Impaired performance might be caused by several factors, such as focusing on immediate rewards, avoidance of losses, or bad memory for experienced payoffs. Thus, the IGT intends to measure the extent to which participants learn to prefer decks that maximize their long-term outcomes.

In order to provide insights into the psychological processes that drive performance on the IGT, several reinforcement-learning (RL) models have been proposed. The most popular model of IGT data is the Expectancy Valence model (EV; Agay et al., 2010; Bass & Nussbaum, 2010; Bishara et al., 2009; Brambilla et al., 2012; Brown, Anderson, Symington, & Paul, 2012; Busemeyer & Stout, 2002; Cella, Dymond, Cooper, & Turnbull, 2012; de Visser et al., 2010; Escartin et al., 2012; Farah, Yechiam, Bekhor, Toledo, & Polus, 2008; Gullo & Stieger, 2011; Hochman, Yechiam, & Bechara, 2010; Isella et al., 2008; Johnson, Yechiam, Murphy, Queller, & Stout, 2006; Kester et al., 2006; Kjome et al., 2010; Lane, Yechiam, & Busemeyer, 2006; Lev et al., 2008; Lovallo, Yechiam, Sorocco, Vincent, & Collins, 2006; Peatfield, Parkinson, & Intriligator, 2012; Premkumar et al., 2008; Sevy et al., 2006, 2007; Stout et al., 2004; van den Bos, Homberg, Gijsbers, Heijer, & Cuppen, 2009; Wood, Busemeyer, Kolling, Cox, & Davis, 2005; Yechiam, Busemeyer, Stout, & Bechara, 2005; Yechiam, Hayden, Bodkins, O'Donnell, & Hetrick, 2008; Yechiam, Kanz, et al., 2008; Yechiam, Arshavsky, Shamay-Tsoory, Yaniv, & Aharon, 2010). Yechiam et al. (2005) fit the EV model to data of 10 clinical groups, and mapped these groups according to the differences in the model parameters. This analysis shows that modeling IGT data provides deeper insights in the reasons for the decision-making deficits of the 10 different disorders, and in relations between the disorders. The EV model has recently been chal-

lenged by the Prospect Valence Learning model (PVL; Ahn, Busemeyer, Wagenmakers, & Stout, 2008; Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011; a detailed description of the models can be found in the next section). Studies comparing these two models failed to find that one model was consistently superior to the other; instead they concluded that combinations of the two models should be used to fit IGT data (Ahn et al., 2008; Fridberg et al., 2010; Yechiam & Busemeyer, 2005). These studies used various methods, such as the post hoc fit criterion (Ahn et al., 2008; Fridberg et al., 2010; Yechiam & Busemeyer, 2005), the generalization criterion (Ahn et al., 2008; Yechiam & Busemeyer, 2005), and simulation methods (Fridberg et al., 2010; a detailed review of these studies can be found in the next section).

Here we present a different approach to compare RL models for the IGT. Using a method known as parameter space partitioning (PSP; Pitt, Kim, Navarro, & Myung, 2006; Pitt, Myung, Montenegro, & Pooley, 2008), we aim to compare the flexibility of three models of the IGT: the EV model, PVL model, and a combination of these two models (i.e., EV-PU; the EV model with Prospect Utility function of the PVL model, Ahn et al., 2008). We thus aim to answer the question: Can each model generate typical empirical choice patterns over a wide range of parameter settings? By applying PSP we can gain detailed insights in differences between the three models and possible reasons for inconsistent results of previous studies comparing these models.

The outline of this article is as follows. The first section explains the IGT, outlines the three RL models for the IGT, and reviews previous studies comparing this class of models. In the second section, we compare these three models with respect to model flexibility. Section three summarizes our findings and discusses their ramifications.

## The IGT and Three Reinforcement-Learning Models

### *The Iowa Gambling Task*

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. Participants are instructed to choose cards in order to maximize their long-term outcomes (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 outcomes. The good, safe decks are typically labelled as decks C and D, whereas the bad, risky decks are labelled 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 are associated with high immediate, constant rewards, but with even higher unpredictable, occasional losses resulting in negative long-term outcomes. Decks C and D, on the other hand, are associated with lower immediate, constant rewards, but with even lower unpredictable, occasional losses and thus result in positive long-term outcomes. In addition to the different magnitude of the immediate rewards and occasional losses resulting in different long-term outcomes, 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).

*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

### **Reinforcement-Learning Models**

In this section, we describe the three RL models we aim to compare with respect to model flexibility: the EV, PVL, and EV-PU models. Table 2 contains the equations of each model and the psychological interpretation of the free parameters including their ranges. In the following, we describe each model separately; the general idea, however, is that each model describes the performance on the IGT through the interaction of distinct psychological processes captured by the model parameters. The first assumption of all models is that, after 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 each deck. This updating process entails that participants adjust their expectancies based on the new utilities they experience, a process described by a learning rule. In the next step, the models assume that the expected utilities of each deck are used to guide the choices of the participants. This assumption is formalized by the softmax choice rule, also known as the ratio-of-strength choice rule, that all three models use to compute the probability of choosing a particular deck on a particular trial (Luce, 1959). This rule contains the sensitivity parameter,  $\theta$ , that indexes the extent to which trial-by-trial choices match the expected utilities of the decks. Values of  $\theta$  close to zero indicate a random choice behavior (i.e., strong exploration leading to many switches), whereas large values of  $\theta$  indicate a choice behavior that is strongly determined by the expected utilities (i.e., strong exploitation leading to few switches).

Table 2. Formalization of the three reinforcement-learning models.

Concept	Model(s)	Model equation	Free parameter	Range
Utility function	EV	$u_k(t) = (1 - w) \times W(t) + w \times L(t)$	w: Loss aversion parameter	[0, 1]
	PVL, EV-PU	$u_k(t) = \begin{cases} X(t)^A & \text{if } X(t) \geq 0 \\ -w \times  X(t) ^A & \text{if } X(t) < 0 \end{cases}$	A: Shape parameter w: Loss aversion parameter	[0, 1] [0, 5]
Learning rule	EV, EV-PU	$Ev_k(t + 1) = (1 - a) \times Ev_k(t) + a \times u_k(t)$	a: Recency parameter	[0, 1]
	PVL	$Ev_k(t + 1) = a \times Ev_k(t) + \delta_k(t) \times u_k(t)$	a: Recency parameter	[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, EV-PU	$\theta(t) = (t/10)^c$	c: Consistency parameter	[-5, 5]
	PVL	$\theta = 3^c - 1$	c: Consistency parameter	[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.

**The Expectancy Valence 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,  $w$ , is the first model parameter, and is called the loss aversion parameter. A small value of  $w$ , that is,  $w < .5$ , is characteristic for decision makers who put more weight on the 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 losses and can thus be described as loss-averse (Ahn et al., 2008).

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  $k$ ,  $Ev_k(t + 1)$ ; the expected utilities of the unchosen decks are left unchanged. If the experienced utility,  $u_k(t)$ , is higher than expected, the expected utility of deck  $k$  on trial  $t + 1$  is adjusted upward. If the experienced utility,  $u_k(t)$ , is lower than expected, the expected utility of deck  $k$  on trial  $t + 1$  is adjusted downward. This updating process is influenced by the second model parameter—the recency parameter,  $a$ . This parameter quantifies the memory for rewards and losses. A value of 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 under consideration, 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  $\theta(t)$  included in the choice rule, changes over trials depending on the response consistency parameter  $c$ . If  $c$  is positive, the sensitivity of trial-by-trial choices to the expected utilities of the decks increases over trials; otherwise, the sensitivity decreases.

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

### **The Prospect Valence Learning model**

The PVL model uses four parameters to formalize its assumptions about participants' performance on the IGT (Ahn et al., 2008). The PVL model assumes that decision makers only process the net outcome after choosing a card from deck  $k$ ,  $k \in \{1, 2, 3, 4\}$  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 losses than rewards on the subjective utility, whereas a value of  $w$  of one indicates equal impact of losses and rewards. As  $w$  approaches zero, the model predicts that 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-RL rule. 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 in the learning rule,  $\delta_k$ , 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. Note that the interpretation of the recency parameter,  $a$ , of the EV model is reversed in the PVL model.

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

To sum up, 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 losses over 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.

### **The Expectancy Valence model with Prospect Utility function**

The EV-PU model is a combination of the EV and PVL models because it uses the utility function of the PVL model, but all remaining equations of the EV model (i.e., the learning rule and the trial-dependent sensitivity parameter; Ahn et al., 2008). This construction results in a model with four parameters: (1) The shape parameter,  $A$ , (2) the loss aversion parameter,  $w$ , (3) the recency parameter,  $a$ , and (4) the response consistency parameter,  $c$ .

### ***Previous Comparisons of Reinforcement-Learning Models***

In this section, we review the main results of previous studies that compared components of the EV and PVL models. Despite the fact that these studies used similar approaches to compare the models (i.e., the post hoc fit criterion, the generalization criterion, and a simulation method), their conclusions are inconsistent. First, Yechiam and Busemeyer (2005) compared combinations of eight different learning rules with three different choice rules, resulting in 24 different models. All 24 models used an extended version of the utility function of the EV model in which the weight for wins and losses are estimated separately. To compare these models, Yechiam and Busemeyer (2005) used the post hoc fit criterion at the individual level and the generalization criterion at the group level. The post hoc fit criterion compares the ability of a given model to predict the choice on the next trial (based on the information obtained from all previous trials) to the same ability of a baseline model.<sup>1</sup> This comparison relies on the Bayesian Information Criterion (BIC) to account for the differences in the number of model parameters. The generalization criterion assesses the accuracy of a given model to predict the entire sequence of choices on one task based on parameters estimated from another task. The authors concluded that the learning rule of the PVL model is superior to the learning rule of the EV model.

Ahn et al. (2008), on the other hand, compared combinations of the utility functions, the learning rules, and the sensitivity functions of the EV and PVL models, resulting in eight different models. They used three different methods: (1) The post hoc fit criterion at the individual level, (2) the generalization criterion for one step ahead predictions for a second task at the individual level, (3) the generalization criterion for predictions of the entire sequence of choices of a second task.<sup>2</sup> The authors concluded that future studies that aim to disentangle psychological processes underlying the IGT should rely on a modified version of the EV model: The utility function of the EV model should be replaced

by the utility function of the PVL model, but the remaining components of the EV model should remain unchanged, a suggestion that gave rise to the EV-PU model. Note that in contrast to Yechiam and Busemeyer (2005), Ahn et al. (2008) did not suggest to replace the learning rule of the EV model by the learning rule of the PVL model.

Finally, Fridberg et al. (2010) used the post hoc fit criterion to compare the EV model to a modified version of the PVL model, that is, the PVL model with learning rule of the EV model. The authors showed that the modified PVL model had a better post hoc fit to data of the control group, whereas the EV model had a better post hoc fit to data of the experimental group of chronic cannabis abusers. But nevertheless, the EV model had a worse post hoc fit to the data of the experimental group than the baseline model. Apart from the post hoc fit criterion, Fridberg et al. (2010) also compared the EV model to the modified PVL model by means of a simulation method. In this simulation, the authors determined the best fitting parameters for each subject and used these parameters to generate predictions of that subject's trial-by-trial deck selections. The simulated proportions of deck selections were then compared to the observed proportions. The results showed that the modified PVL model simulated the observed choice patterns of both groups more accurately than the EV model. The authors thus concluded that the modified PVL model better accounts for psychological processes underlying the IGT than the EV model.

To sum up, previous studies failed to find a consistent advantage of one model over its competitors. However, they all showed that neither the pure EV model nor the pure PVL model should be accepted as the default model to describe IGT data. Instead, previous studies recommended to use different combinations of the EV and PVL models, but it is still unclear which exact combination is most favorable.

A possible reason for the inconsistencies in previous comparisons of RL models might be that these studies based their analyses on data sets that show different choice patterns. Whereas healthy participants in the study of Ahn et al. (2008) showed a preference for the safe decks (that is, decks C and D) over the risky decks (that is, decks A and B), healthy participants of Fridberg et al. (2010) and Yechiam and Busemeyer (2005) showed a preference for the decks with infrequent losses (decks B and D) over the decks with frequent losses (decks A and C). Hence it seems that comparisons of components of the EV and PVL models yield conclusions that depend on the observed choice pattern. This indicates that some choice patterns might be more accurately fitted by components of the EV model, whereas other choice patterns might be more accurately fitted by components of the PVL model. This indication has already been touched upon briefly by Yechiam and Busemeyer (2005) who demonstrated that the EV model fails to predict a strong preference for bad deck B. This finding has been supported by Fridberg et al. (2010)'s simulation showing that the EV model over-predicts the observed proportion of cards chosen from decks with high-frequent losses (decks A and C).



In order to explore systematically the ability of the different models to handle a preference for decks with infrequent losses, we employ the PSP method. This method systematically assesses the choice patterns predicted by a given model across its entire parameter space and allows us to investigate whether typical empirical choice patterns occupy a major region of the parameter space (Pitt et al., 2006, 2008). Hence the results of a PSP study describe how flexible a model is. An ideal model is parsimonious and generates only those choice patterns that are observed in experiments. A crucial advantage of PSP over previous methods used to compare RL models (Ahn et al., 2008; Fridberg et al., 2010; Yechiam & Busemeyer, 2005)<sup>3</sup> is that PSP is a global method enabling us to assess the full range of choice patterns a model can generate, whereas previous methods are local methods that are restricted to a particular parameter combination (Pitt et al., 2006). Our conclusions therefore do not depend on the idiosyncrasies of any single data set.

Before we can interpret the results of a PSP study, we need to know what choice patterns are observed empirically. When Bechara et al. (1994) developed the IGT, they explicitly assumed the existence of at least two choice patterns: Impaired decision makers are assumed to prefer risky decks over safe decks, whereas healthy participants are assumed to show the reversed choice pattern. However, findings of several studies go against this assumption (Caroselli, Hiscock, Scheibel, & Ingram, 2006; Chiu & Lin, 2007; Chiu et al., 2008; Dunn, Dalgleish, & Lawrence, 2006; MacPherson, Phillips, & Della Sala, 2002; Lin, Chiu, Lee, & Hsieh, 2007; Wilder, Weinberger, & Goldberg, 1998; Yechiam & Busemeyer, 2005). In a recent IGT review article, we showed that healthy participants often display a preference for decks with infrequent losses over decks with frequent losses (Steingroever, Wetzels, Horstmann, Neumann, & Wagenmakers, 2013). This choice pattern corresponds to the one shown in the studies of Fridberg et al. (2010) and Yechiam and Busemeyer (2005). This suggests that there are at least three different types of typical empirical choice patterns that a model of the IGT should produce across a wide range of parameter settings: (1) Preference for good decks over bad decks,  $\{C, D\} \succ \{A, B\}$ , (2) Preference for bad decks over good decks,  $\{A, B\} \succ \{C, D\}$ , and (3) Preference for decks with infrequent losses over decks with frequent losses,  $\{B, D\} \succ \{A, C\}$ . In addition to capturing the rank order of deck preferences, a good RL model of the IGT should also be able to capture the switches participants make on the IGT (Zhao & Costello, 2007). In a RL context (Sutton & Barto, 1998), it is assumed that participants first explore all decks, and then settle down and exploit the most profitable ones. This assumption implies that the number of switches decreases across trials; however, many data sets failed to show a systematic decrease in the number of switches across trials (Steingroever et al., 2013). It is interesting to see whether the models also generate a decrease in the number of switches, or whether they can generate an explorative behavior with many switches.

In this article, we aim to gain more detailed insights in differences between RL models for the IGT and possible reasons for inconsistent results of previous studies comparing these models. We will therefore compare the EV, PVL, and EV-PU models with respect to model flexibility.

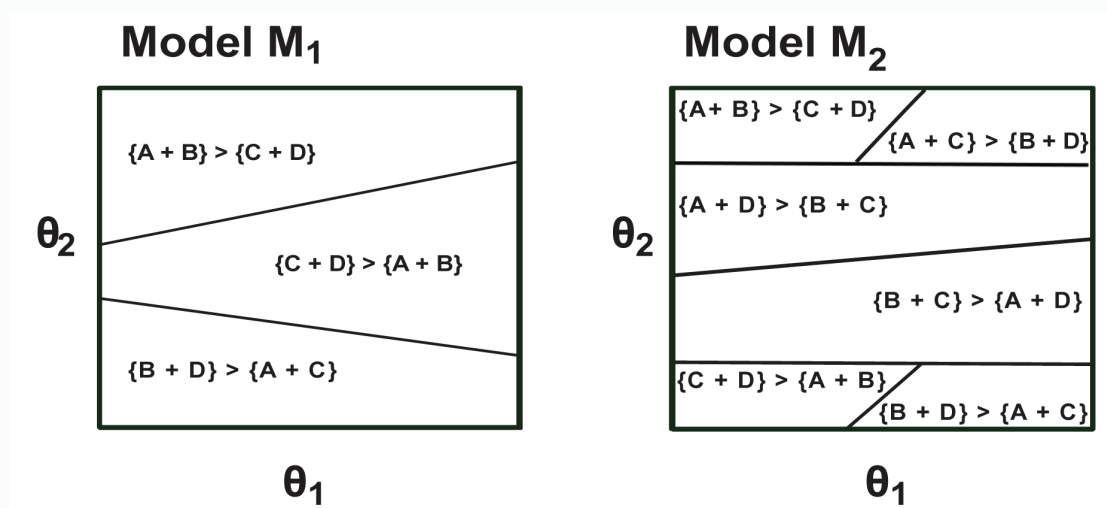
## Comparison of the EV, PVL, and EV-PU Models

### Methods

In order to evaluate the flexibility of the three RL models, we performed a PSP study for each model (Pitt et al., 2006, 2008). The PSP method systematically assesses the choice patterns predicted by the models across the entire parameter space. To explain PSP, we use the example of two hypothetical models with two parameters,  $\theta_1$  and  $\theta_2$ , as shown in Figure 1. The left panel shows a model that can generate three choice patterns, whereas the model presented in the right panel is more flexible because it can generate six different choice patterns. A model is overly flexible when it can generate not only all choice patterns that are observed empirically, but also choice patterns that are logically possible, but never observed. Instead, one should prefer a less flexible model that—ideally—only generates choice patterns that are also frequently observed in experiments (Pitt et al., 2006, 2008). In the case of the IGT, such a model is shown in the left panel of Figure 1.

Pitt et al. (2006) describe a new search algorithm to implement PSP. In our implementation we did not use their sophisticated search algorithm, but followed the conceptual idea of PSP, and used a grid search that works as follows: For each model and for each parameter, we chose 60 values that were equally spaced over the corresponding

*Figure 1.* The partitioned parameter space of two hypothetical models with two parameters,  $\theta_1$  and  $\theta_2$ . Model  $M_1$  generates three different patterns. Model  $M_2$  is more flexible because it generates six different patterns. See text for further explanations.



parameter range. All combinations of these parameter values were used to generate data for 100 synthetic participants completing a 100-trial IGT. The generated data were used to compute the average proportions of choices from each deck for each parameter combination. For each parameter combination, the choice proportions were then sorted in decreasing order to determine the generated rank order of deck preferences. We defined five possible choice patterns: (1) Preference for good decks over bad decks, that is,  $\{C, D\} > \{A, B\}$ , (2) preference for bad decks over good decks, that is,  $\{A, B\} > \{C, D\}$ , (3) preference for decks with infrequent losses over decks with frequent losses, that is,  $\{B, D\} > \{A, C\}$ , (4) preference for decks with frequent losses over decks with infrequent losses, that is,  $\{A, C\} > \{B, D\}$ , and (5) remaining choice patterns. This definition of choice patterns is labeled in the remainder of this article as “broad definition of choice patterns.” Finally, we computed for each model the proportion of the parameter space occupied by each generated choice pattern. Even though we defined five possible types of choice patterns, we assume based on the theory underlying the IGT (Bechara et al., 1994, 1997) and our IGT review (Steingroever et al., 2013) that a good model for the IGT should only generate the first three types of choice patterns, as presented in the left panel of Figure 1.

Note that the above definition of choice patterns only considers the rank order of the overall proportions of choices from each deck averaged over 100 repeated data generations with the same parameter combination. This means that it did not matter whether a model generated, for example, a very strong or a very weak preference for the good decks over the bad decks. Both generated choice patterns would have been classified as the choice pattern “good decks over bad decks,” that is,  $\{C, D\} > \{A, B\}$ . To go beyond this coarse classification, we also analyzed the models’ behavior when confronted with pronounced deck preferences. To get an indication of pronounced deck preferences shown on the IGT, we searched our IGT data pool ( $N = 394$ ) for participants that chose at least 65% cards from either the good decks, the bad decks, or the decks with infrequent losses (Steingroever et al., 2013).<sup>4</sup> We thus obtained three groups of participants with a pronounced deck preference. For each of these three groups, we computed the mean proportions of choices from each deck (data are presented in the next section). These mean proportions of choices from each deck determine our second definition of choice patterns that is labeled as “restricted definition of choice patterns” in the remainder of this article.

We then determined for each model, the proportion of the parameter space that produced pronounced deck preferences according to the “restricted definition of choice patterns.” We compared this theoretical popularity of pronounced deck preferences to the empirical popularity of deck preferences.

In addition, we determined the mean proportions of switches during the last 50 trials for all parameter combinations that produced pronounced deck preferences according to the “restricted definition of choice patterns,” and compared them to the empirical mean

proportions of switches. 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). The codes for the PSP studies are available on [www.helensteingroever.com](http://www.helensteingroever.com).

## Results

### Broad definition of choice patterns

Table 3 presents for each model the proportion of the parameter space occupied by each of the five different types of choice patterns. From this table, it is evident that the number of different choice patterns a model can produce does not differ between the three models under consideration. However, if we consider the partitioned parameter spaces, we detect substantial differences between the models: In the case of the EV model, the choice pattern “good decks over bad decks” is the most central to its overall performance, as it occupies the largest part of its parameter space. The second largest part of its parameter space is occupied by the choice pattern “bad decks over good decks.” However, the choice pattern “infrequent losses over frequent losses” is only generated over a very minor part of its parameter space.

Table 3. Proportions of choice patterns generated by each model.

Choice pattern		Proportion of all choice patterns		
		EV	PVL	EV-PU
Good > bad decks	{C, D}>{A, B}	.585	.427	.661
Bad > good decks	{A, B}>{C, D}	.153	.005	.005
Infr. > frequent losses	{B, D}>{A, C}	.075	.363	.181
Frequent > infr. losses	{A, C}>{B, D}	.099	.003	.003
Remaining		.089	.202	.151

Just as the EV model, the PVL model produces the choice pattern “good decks over bad decks” over the largest part of its parameter space. But in contrast to the EV model, the PVL model also generates the choice pattern “infrequent losses over frequent losses” over a large part of its parameter space. Note that it seems almost incidental that the PVL model produces a preference for the bad decks. Not only does this choice pattern occupy a small region of the parameter space, but larger regions are produced by choice patterns that are not empirically observed, that is, choice patterns labeled as “remaining.”

Among the three RL models, the EV-PU model shows the strongest preference for the choice pattern “good decks over bad decks.” Hence its parameter space is very unequally partitioned among the five different types of choice patterns. The two next largest parts of its parameter space are occupied by the choice pattern “infrequent losses over frequent losses” and by “remaining” choice patterns.

Since the EV model generates the choice pattern “infrequent losses over frequent losses,” and the PVL and EV-PU models the choice pattern “bad decks over good decks” only within a very small region of the parameter space, we have grounds to conclude that these choice patterns are uncharacteristic of the models, and are thus almost irrelevant to their overall performance (Pitt et al., 2006). These findings are especially relevant because the choice pattern “infrequent losses over frequent losses” is often observed in healthy participants (Steingroever et al., 2013), and the choice pattern “bad decks over good decks” is thought to be characteristic in patients with lesions to the vmPFC (Bechara et al., 1994, 1997).

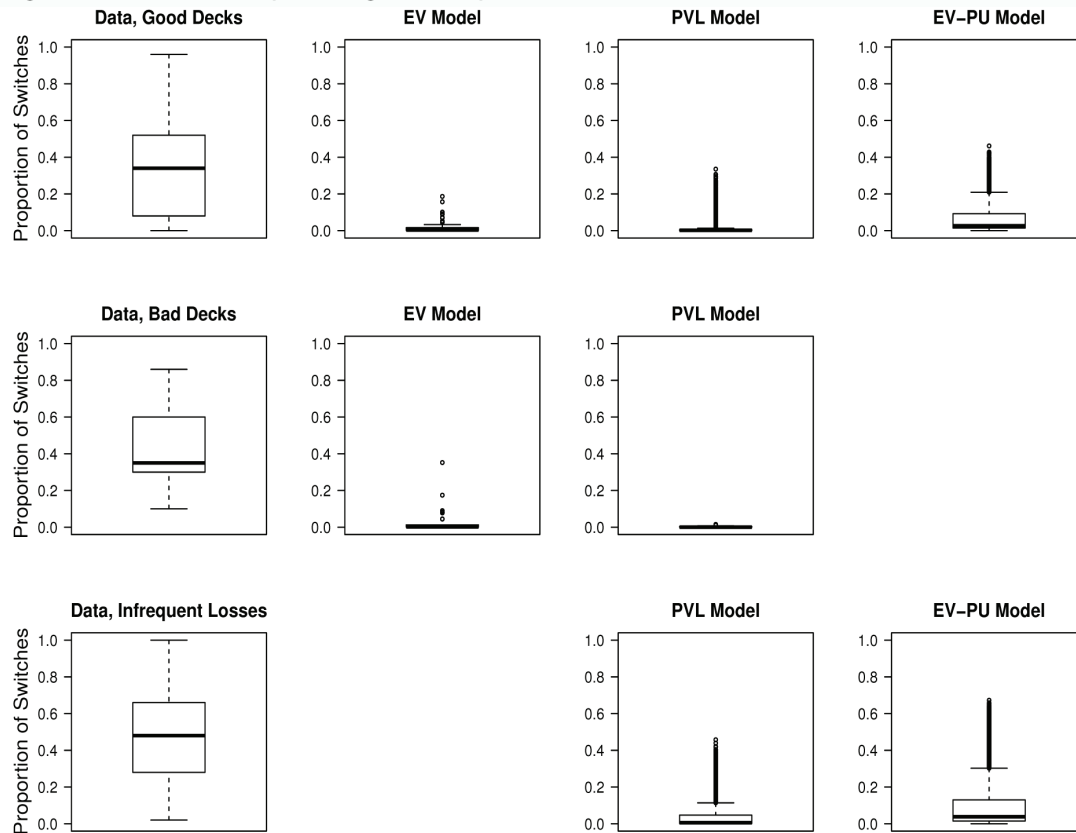
**Restricted definition of choice patterns**

Table 4 shows the mean proportions of choices from each deck and the mean proportions of switches during the last 50 trials of three groups of healthy participants who showed a pronounced preference for either the good decks ((C + D) ≥ .65), the bad decks ((A + B) ≥ .65), or the decks with infrequent losses ((B+D) ≥ .65). Participants classified into one of these three groups were selected out of our IGT data pool (N = 394) (Steingroever et al., 2013). By using the .65-criterion, we include participants with pronounced deck preferences and exclude participants with random choice behaviors. Note that 53.6% of all participants in our data pool showed a pronounced deck preference, and made at least 65% choices from the two most preferred decks. This empirical popularity of pronounced deck preferences underscores how important it is that an RL model for the IGT is able to produce such choice patterns. Table 4 also contains for each group the mean proportions of switches during the last 50 trials and statistics quantifying the distribution of switches in each group, that is, the interquartile range and the minimum and maximum proportions of switches during the last 50 trials. This information is visualized by the boxplots shown in the first column of Figure 2. From the table and the figure, it is evident that, in general, in all three groups participants switch frequently. However, the interquartile ranges and the

*Table 4.* Mean proportions of choices from each deck and mean proportions of switches during the last 50 trials of healthy participants showing a pronounced deck preference. Healthy participants are selected out of data pool (N = 394) reported by Steingroever et al. (2013).

Choice pattern	N	Deck A [sd]	Deck B [sd]	Deck C [sd]	Deck D [sd]	Switches during the last 50 trials [25%, 75% quantile] (min, max)
(C + D) ≥ .65	54	.10 [.05]	.14 [.05]	.36 [.17]	.40 [.14]	.35 [.08, .52] (0.00, 0.96)
(A + B) ≥ .65	18	.25 [.07]	.52 [.11]	.11 [.05]	.12 [.06]	.43 [.31, .58] (0.10, 0.86)
(B + D) ≥ .65	139	.12 [.05]	.37 [.12]	.13 [.05]	.39 [.12]	.47 [.28, .66] (0.02, 1.00)

Figure 2. Boxplots of observed and generated proportions of switches during the last 50 trials. Each row presents the results for different choice patterns: First row: Pronounced preference for the good decks; second row: Pronounced preference for the bad decks; third row: Pronounced preference for the decks with infrequent losses. The first column presents the switches of 211 healthy participants selected out of our IGT data pool (Steingroever et al., 2013) (cf. Table 4). The remaining columns present the switches generated by the EV, PVL, and EV-PU models (cf. Table 5). A cell is unfilled if the corresponding model fails to generate the corresponding choice pattern.



minimum and maximum proportions of switches during the last 50 trials show that there is also a large variability in the proportions of switches, such that the switches of healthy participants vary between no switches at all to switches on every trial. This tendency to switch frequently, but also the large individual differences in the switch behavior of healthy participants are illustrated by Figures 3, 7, and 10 which show the trial-by-trial choices of healthy participants with a pronounced preference for the good decks, bad decks, and decks with infrequent losses, respectively. The deck selection profiles of all healthy participants that showed a pronounced deck preference, that is, at least 65% choices from the two most preferred decks, can be found in the online appendix.<sup>5</sup>

The mean proportions of choices shown in Table 4 were used to investigate whether the models can also generate these pronounced deck preferences. This means that we

Table 5. Proportion of choice patterns generated by each model that satisfy the “restricted definition of choice patterns”. Note that this definition is only based on the mean proportions of switches of the two strongest preferred decks (second column). For the selected choice patterns, the corresponding mean and standard deviation of the model parameters, and the mean proportions of switches during the last 50 trials are presented.

Model	Choice pattern	Proportion of all choice patterns	A [sd]	w [sd]	a [sd]	c [sd]	Switches during the last 50 trials [25%, 75% quantile] (min, max)
EV	$C \geq .36, D \geq .40$	.00027		.45 [.04]	.04 [.02]	3.76 [0.67]	.020 [.0016, .017] (.00020, .19)
	$A \geq .25, B \geq .52$	.00013		.04 [.03]	.15 [.07]	2.66 [0.49]	.032 [.00060, .013] (.00, .35)
	$B \geq .37, D \geq .39$	none					
PVL	$C \geq .36, D \geq .40$	.010	.66 [.22]	.74 [.41]	.78 [.16]	3.21 [1.14]	.0065 [.0002, .0054] (.00, .34)
	$A \geq .25, B \geq .52$	.0000035	.92 [.07]	.02 [.03]	.95 [.05]	0.79 [0.15]	.0018 [.00, .0034] (.00, .015)
	$B \geq .37, D \geq .39$	.016	.29 [.25]	.46 [.76]	.50 [.29]	2.41 [1.24]	.042 [.0020, .0468] (.00, .46)
EV-PU	$C \geq .36, D \geq .40$	.0019	.66 [.21]	.63 [.34]	.29 [.19]	3.41 [1.00]	.067 [.014, .092] (.00, .46)
	$A \geq .25, B \geq .52$	none					
	$B \geq .37, D \geq .39$	.0080	.17 [.19]	.52 [.42]	.30 [.22]	2.70 [0.97]	.10 [.0150, .130] (.00, .67)

searched the results of the PSP studies for those parameter combinations that yielded pronounced deck preference as presented in Table 4, that is, a pronounced preference for the good decks ( $C \geq .36, D \geq .40$ ), a pronounced preference for the bad decks ( $A \geq .25, B \geq .52$ ), and a pronounced preference for the decks with infrequent losses ( $B \geq .37, D \geq .39$ ). Table 5 presents for each model the proportion of all choice patterns that satisfied this “restricted definition of choice patterns.” The table also presents the mean and standard deviation of the parameter combinations that generated these pronounced deck preferences, and the corresponding generated proportions of switches during the last 50 trials.

Figure 3. Deck selection profiles of four healthy participants showing a preference for the good decks. The filled circles indicate the occurrence of rewards and losses together; the empty circles indicate the occurrence of only rewards.

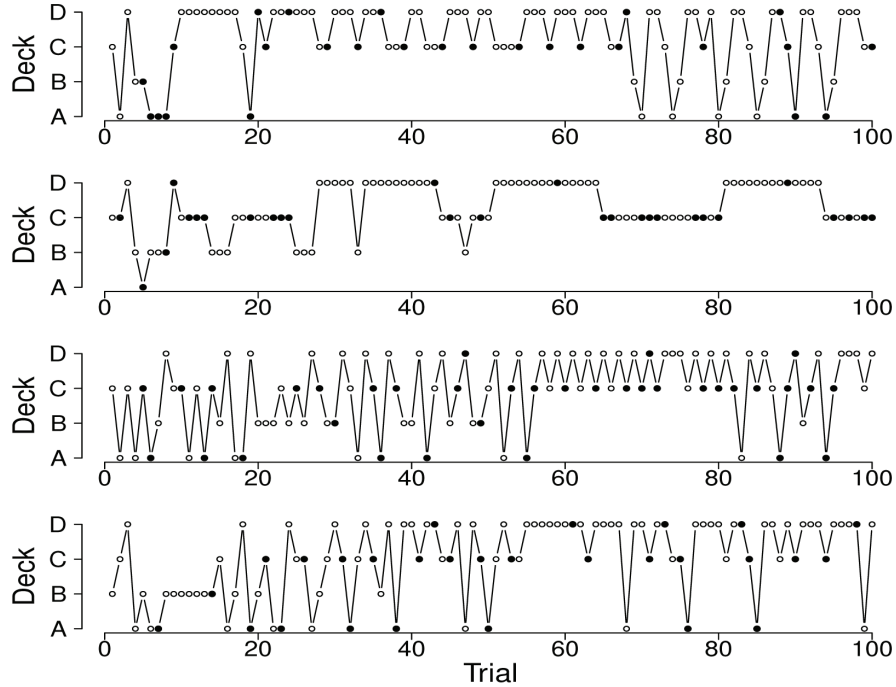


Figure 4. Deck selection profiles of four synthetic participants showing a preference for the good decks (generated by the EV model;  $w = 0.42$ ,  $a = 0.03$ ,  $c = 5.00$ ).

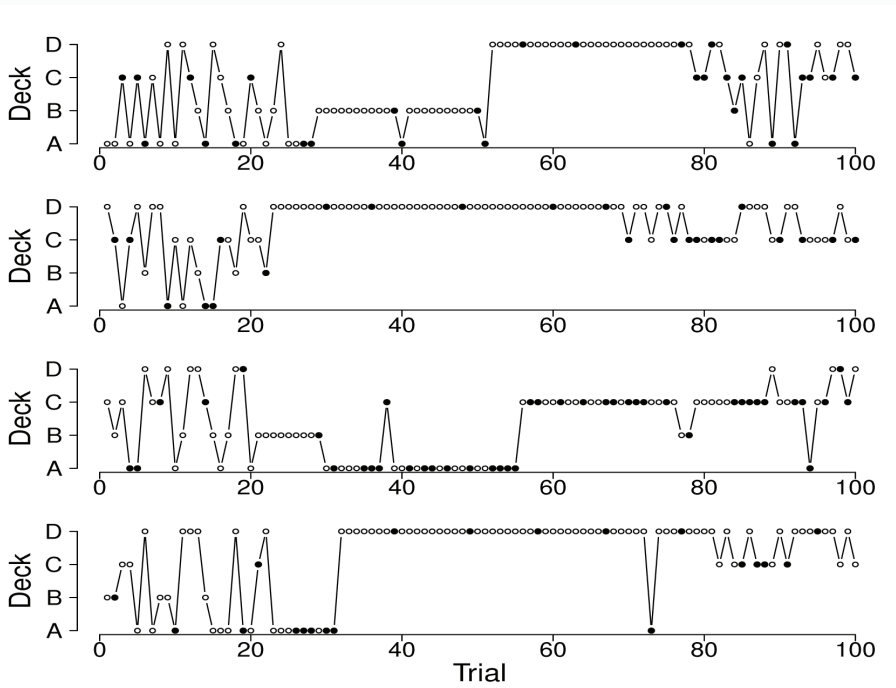




Figure 5. Deck selection profiles of four synthetic participants showing a preference for the good decks (generated by the PVL model;  $A = .97, w = .93, a = .93, c = .59$ ).

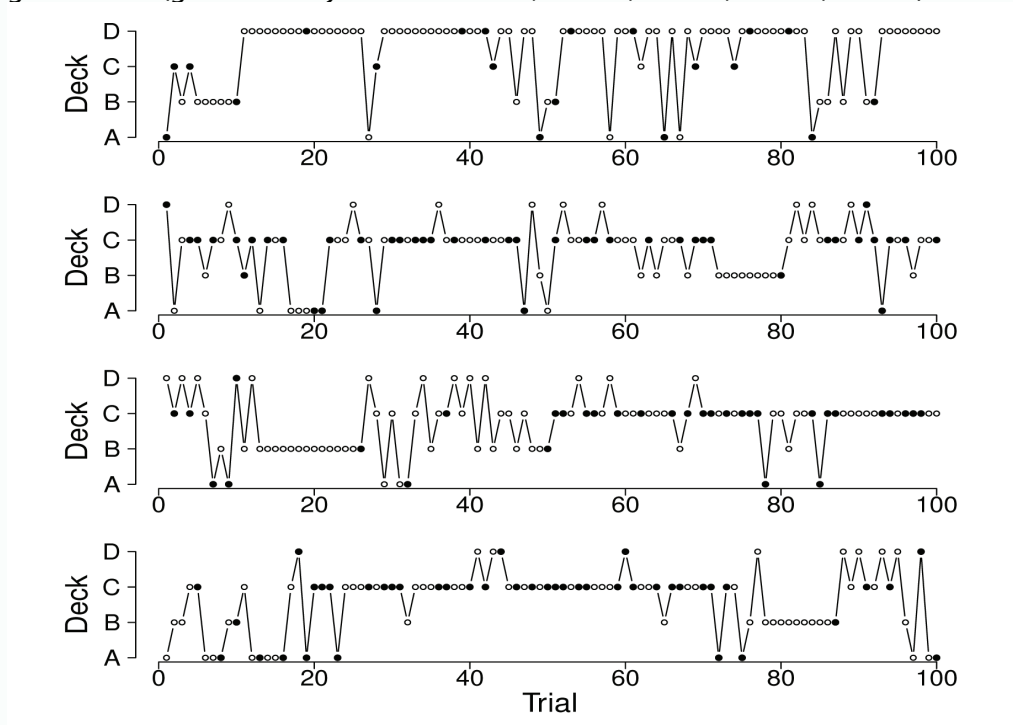


Figure 6. Deck selection profiles of four synthetic participants showing a preference for the good decks (generated by the EV-PU model;  $A = 0.81, w = 0.34, a = 0.24, c = 5.00$ ).

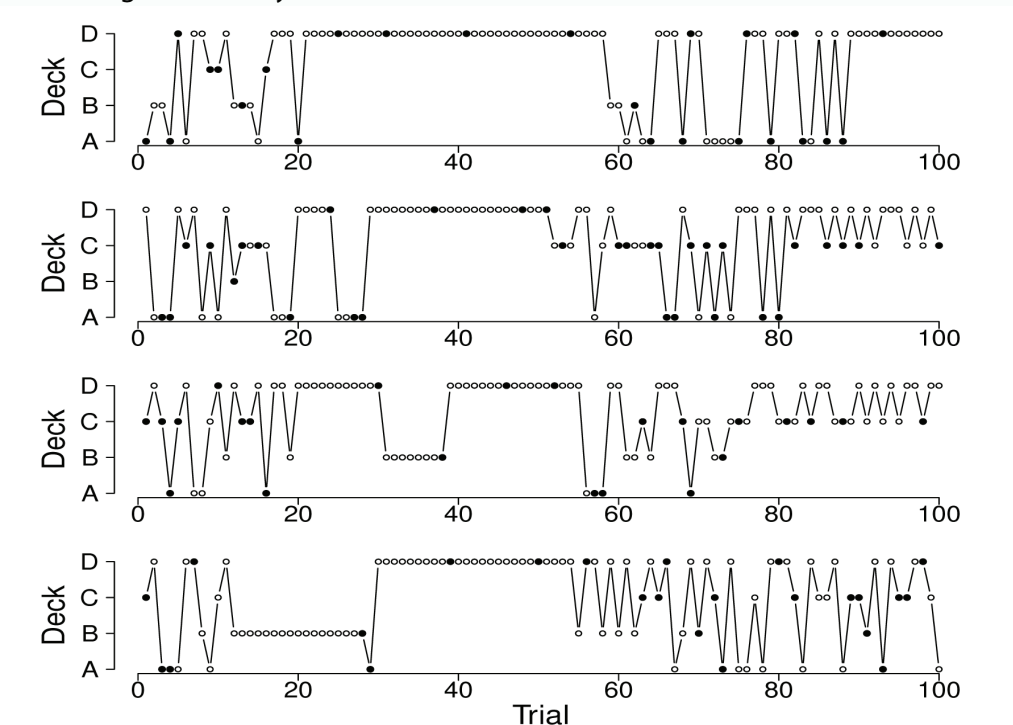


Figure 7. Deck selection profiles of four healthy participants showing a preference for the bad decks.

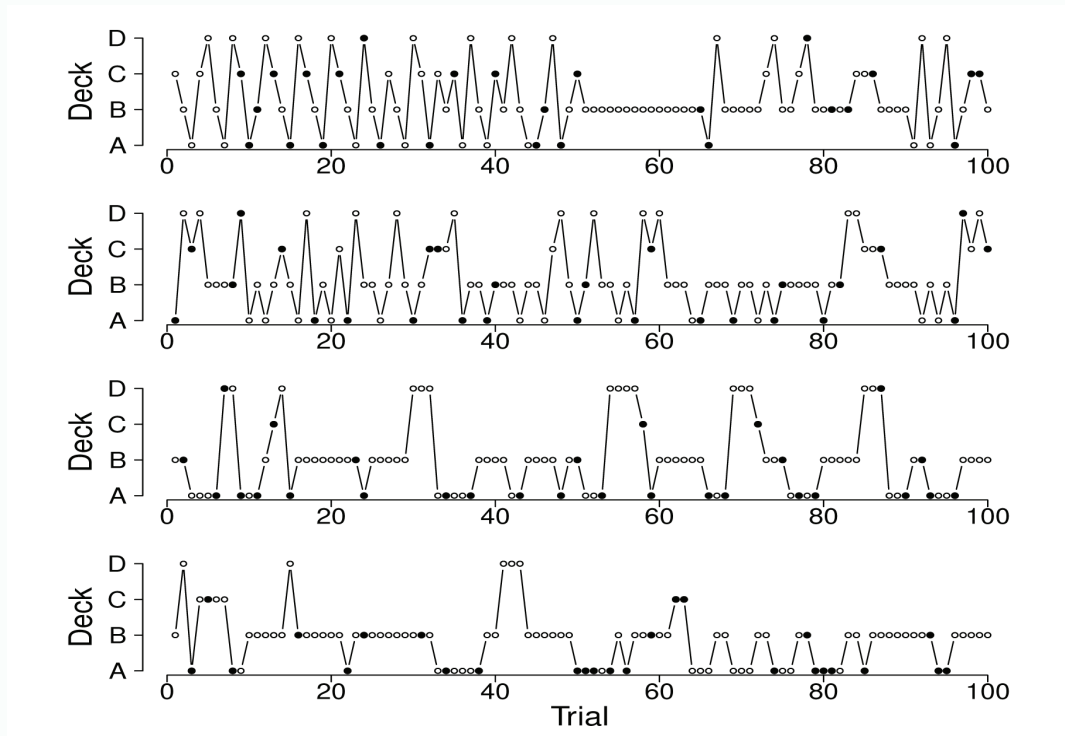


Figure 8. Deck selection profiles of four synthetic participants showing a preference for the bad decks (generated by the EV model;  $w = 0.00$ ,  $a = 0.25$ ,  $c = 1.61$ ).

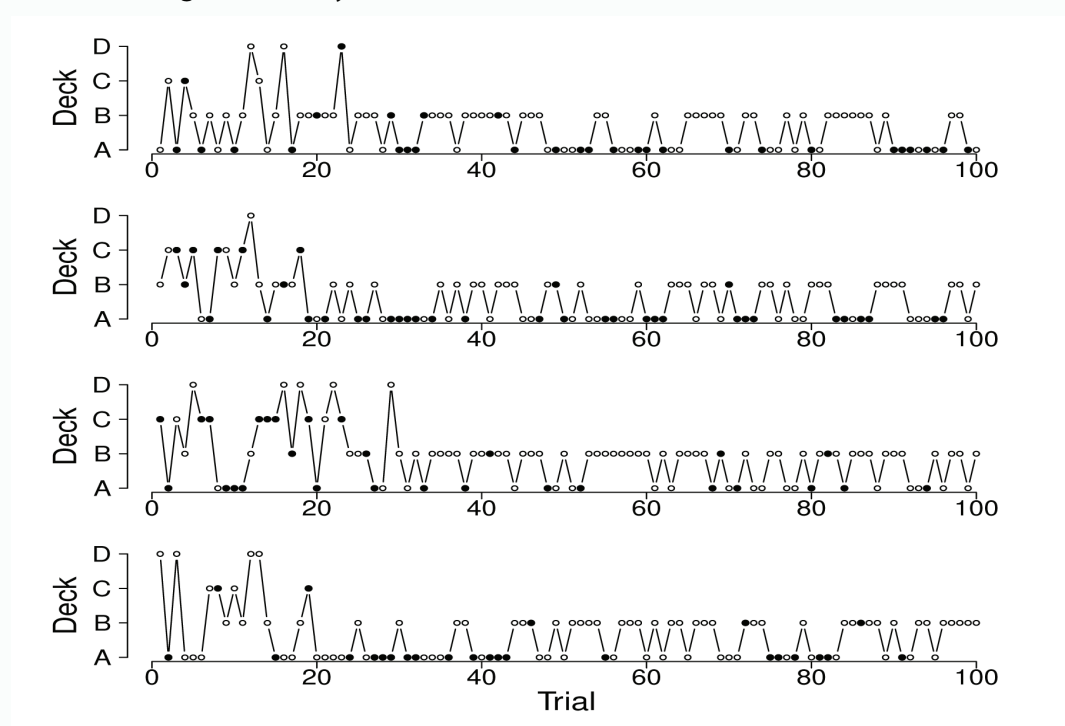


Figure 9. Deck selection profiles of four synthetic participants showing a preference for the bad decks (generated by the PVL model;  $A = 1.00, w = 0.00, a = 0.88, c = 0.76$ ).

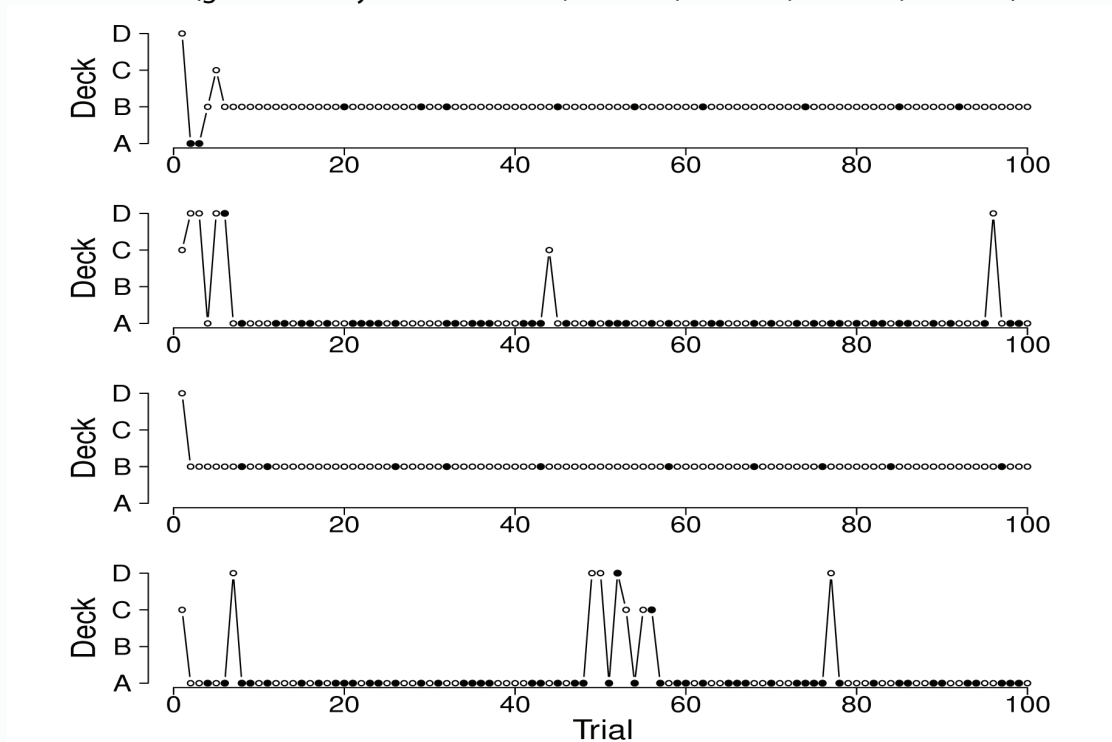


Figure 10. Deck selection profiles of four healthy participants showing a preference for the decks with infrequent losses.

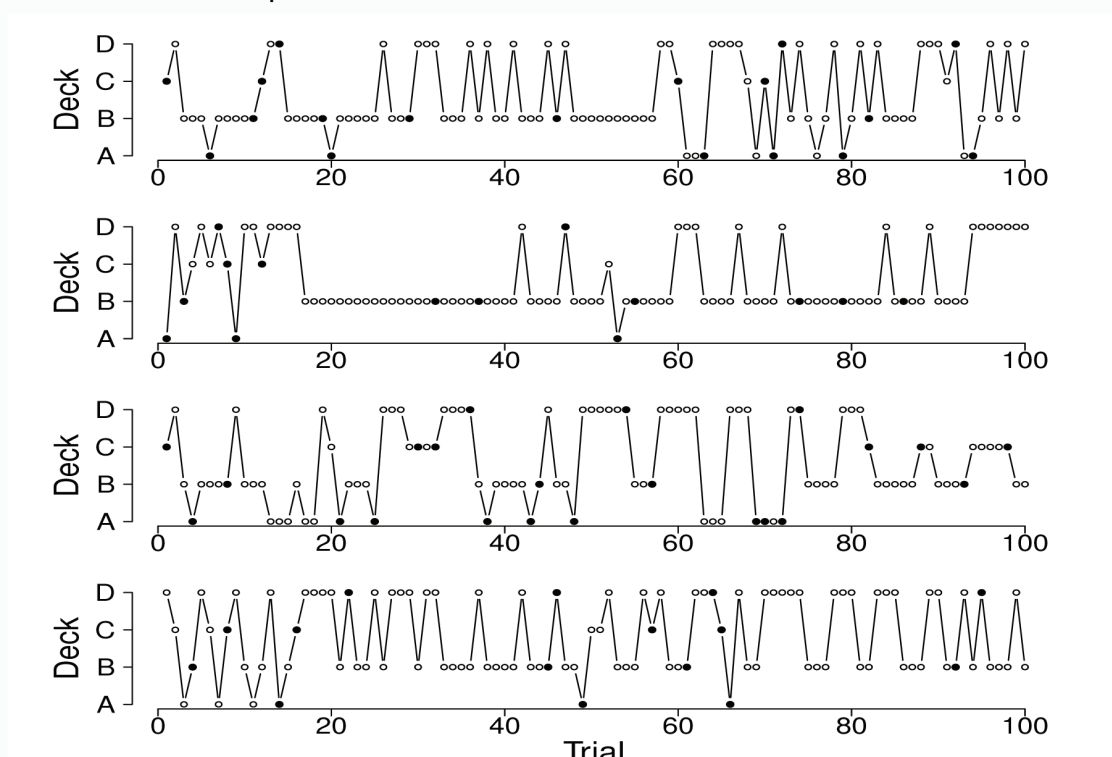


Figure 11. Deck selection profiles of four synthetic participants showing a preference for the decks with infrequent losses (generated by the PVL model;  $A = .07, w = .25, a = .88, c = .34$ ).

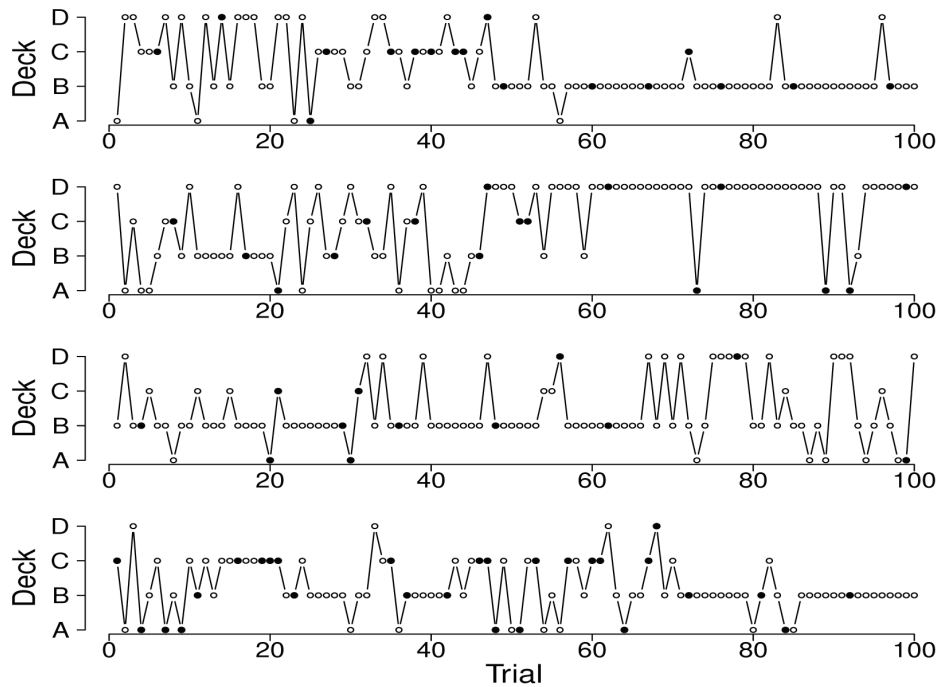
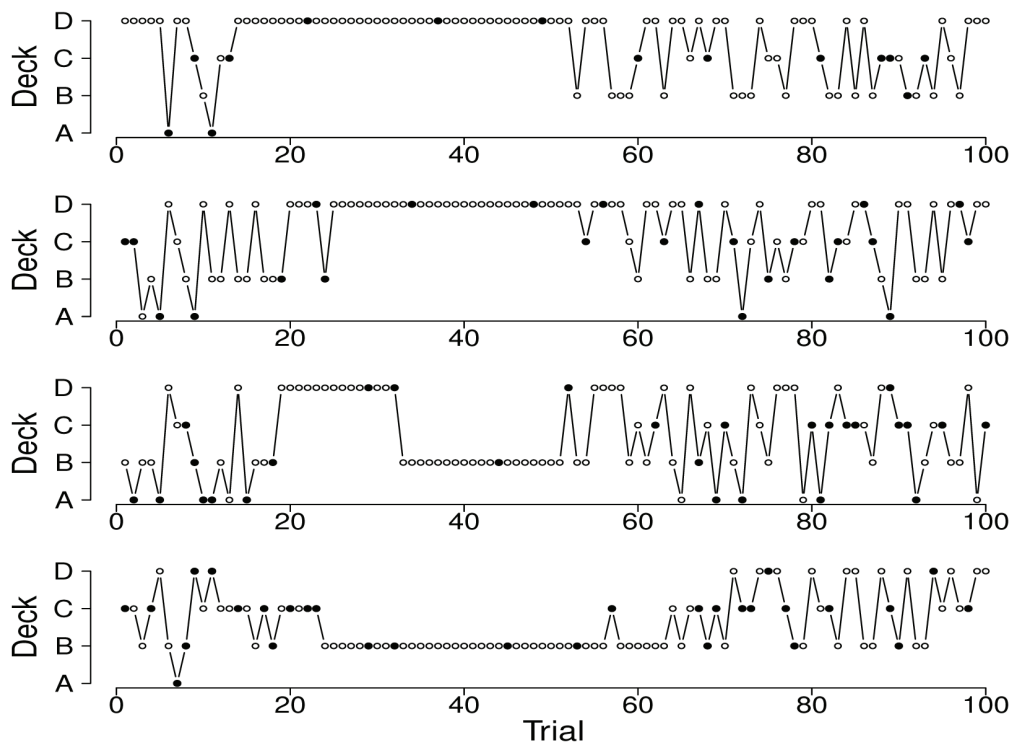


Figure 12. Deck selection profiles of four synthetic participants showing a preference for the decks with infrequent losses (generated by the EV-PU model;  $A = 0.03, w = 0.08, a = 0.36, c = 4.66$ ).



From this table, it seems at a first glance that all models are able to generate all pronounced deck preferences over a minor part of their parameter space, except for the EV model that fails to generate a pronounced preference for the decks with infrequent losses and for the EV-PU model that fails to generate a pronounced preference for the bad decks.<sup>6</sup> For the remaining two pronounced choice patterns of the EV and EV-PU models and for all three pronounced choice patterns of the PVL model, it is important to note that only minor parts of the parameter spaces of all models are occupied by the choice patterns featuring pronounced deck preferences even though they are frequently observed in experiments. For instance, 139 healthy participants out of our data pool (35.3%) show a pronounced preference for the decks with infrequent losses ( $(B + D) \geq .65$ ). However, the EV model completely fails to generate such a pronounced preference for the decks with infrequent losses, the PVL model only generates this choice pattern over 1.6% of its parameter space, and the EV-PU model only over 0.8% of its parameter space.

Next to considering the proportion of the parameter space occupied by each choice pattern featuring pronounced deck preferences, it is crucial to consider the proportions of switches during the last 50 trials predicted by each model. The last column of Table 5 shows for each model and each choice pattern separately, the proportions of switches during the last 50 trials averaged over all selected parameter combinations, and statistics quantifying the corresponding distribution of switches, that is, the interquartile range and the minimum and maximum proportions of switches during the last 50 trials. This information is visualized by the last three columns of Figure 2. When comparing the generated and observed mean proportions of switches during the last 50 trials, it is apparent that the models dramatically underestimate the observed proportions of switches, that is, the generated mean proportions of switches equal or fall below .10 for all models and all choice patterns, whereas the observed mean proportions of switches equal or exceed .35 for all choice patterns (Tables 4 and 5). In addition, only in the case of the EV-PU model and the choice pattern featuring a pronounced preference for the good decks do the interquartile range of the observed proportions of switches and generated proportions of switches overlap; for all other choice patterns and models, the interquartile ranges of the generated proportions of switches lie below the interquartile ranges of the observed proportions of switches (Figure 2, Table 4 and 5).

Figure 2 and Table 5 also reveal important differences between the models: First, the EV model performs worse among the three models; even if we neglect the generated proportions of switches during the last 50 trials and only consider the generated rank order of deck preferences, the EV model fails to generate a frequency-of-losses effect. In addition, it predicts, among the three models, the smallest number of switches when generating a pronounced preference for the good decks over the bad decks. In fact, the deck selection profiles that were generated with the EV model and those parameter values that yielded the largest number of switches and a pronounced preference for the good decks, display

few switches during the last 50 trials and long stays on the same deck for four synthetic participants showing a pronounced preference for the good decks (Figure 4). This strong exploitation is due to the high consistency parameter,  $c$ , and the low recency parameter,  $a$ , suggesting almost no forgetting and very weak recency effects. However, with respect to the choice pattern featuring a pronounced preference for the bad decks, the EV model outperforms its competitors. Even though the EV model in general underestimates the mean proportions of switches during the last 50 trials, it best approximates the observed mean proportion of choices among the three models (Figure 2, Tables 4 and 5). In fact, the deck selection profiles that were generated with the EV model and those parameter values that yielded the largest number of switches and a pronounced preference for the bad decks, strongly resemble the deck selection profiles of healthy participants (Figures 7 and 8).

Second, the PVL model outperforms the EV model because it can generate a frequency-of-losses effect, and for pronounced preferences from the good decks or the decks with infrequent losses, the PVL model's largest generated proportions of switches lies in interquartile ranges of the observed proportions of switches (cf. Figure 2, Tables 4 and 5). This suggests that for very rare parameter combinations the PVL model meets both empirical regularities: the rank order of deck preferences and the proportions of switches. This is illustrated by Figures 5 and 11 showing deck selection profiles featuring pronounced preferences for the good decks and decks with infrequent losses, respectively. These deck selection profiles were generated with the PVL model and those parameter values that yielded the largest number of switches and a pronounced preference for the good decks or the decks with infrequent losses. However, the deck selection profiles featuring a pronounced preference for the bad decks show that, for this choice pattern, the PVL fails to meet both empirical regularities in that it generates almost no switches during the last 50 trials (Figure 9). This is supported by Figure 2 and Tables 4 and 5 showing that, for choice patterns featuring a pronounced preference from the bad decks, the largest generated proportions of switches is smaller than the 25% quantile of the observed proportions of switches.

Third, the EV-PU model is clearly inferior to the EV and PVL models with respect to the choice pattern featuring a pronounced preference for the bad decks. But it outperforms its competitors with respect to the choice patterns featuring a pronounced preference of the good decks and the decks with infrequent losses; even though all models underestimate the observed proportions of switches during the last 50 trials, the EV-PU model comes closest to the observed proportions of switches. This indicates that for very rare parameter combinations the EV-PU model can account for both empirical regularities: the rank order of deck preferences and the proportions of switches. This is illustrated by Figures 6 and 12 showing the deck selection profiles of synthetic participants with a pronounced preference for the good decks and for the decks with infrequent losses, respectively, that were generated with the EV-PU model and with those parameter values that yielded the

largest number of switches. A surprising finding of these deck selection profiles, is that the EV-PU model predicts many switches during the last 50 trials, but a long stay at the same deck at the first half of the task. This might be due to the interaction between the high consistency parameter,  $c$ , that predicts strong exploitation, and the small recency parameter,  $a$ , that strongly discounts all past expectancies and disables strong learning effects.

To conclude, many healthy participants out of our data pool (53.6%) showed pronounced deck preferences, that is, a pronounced preference for good decks ( $(C + D) \geq .65$ ), a pronounced preference for bad decks ( $(A + B) \geq .65$ ), or a pronounced preference for decks with infrequent losses ( $(B + D) \geq .65$ ) (Table 4). This empirical popularity of pronounced deck preferences is not reflected by the three RL models; the models produce choice patterns that satisfy the “restricted definition of choice patterns” only within minor parts of their parameter spaces (Table 5). In addition, healthy participants in general show many switches during the last 50 trials. However, the RL models in general predict that participants, who show pronounced deck preferences switch vary rarely during the last 50 trials; all generated mean proportions of switches during the last 50 trials equal or fall below .10 whereas the observed mean proportions of switches lie around .40. Moreover, there are large individual differences in the proportions of switches of healthy participants, such that the switches of healthy participants vary between no switches at all to switches on every trial. However, the RL models fail to generate large proportions of switches, that is, none of the models can generate more than 67% switches during the last 50 trials.

## Discussion

This article compared the EV, PVL, and EV-PU models with respect to model flexibility using parameter space partitioning. We used two different definitions of choice patterns; the broad definition allowed us to get an indication of how central each of the choice patterns are to the models’ overall performance, and the restricted definition allowed us to assess the models’ data-fitting potential when confronted with data featuring pronounced deck preferences.

The analysis based on the broad definition of choice patterns showed that the partitioned parameter spaces of the three models differ. We thus learned that the choice pattern “good decks over bad decks” is the most central to all models. But the popularity of this choice pattern differs between the models; the part of the parameter space occupied by this choice pattern is the largest in the case of the the EV-PU model (66.1%), and the smallest in the case of the PVL model (42.7%). Another important difference between the three models is that the choice pattern “infrequent losses over frequent losses” is almost irrelevant to the overall performance of the EV model, whereas the choice pattern “bad decks over good decks” is almost irrelevant to the overall performance of the PVL and EV-PU models. We were thus able to detect inconsistencies between the empirical frequency of each choice pattern and the frequency predicted by each model.

The analysis based on the restricted definition of choice patterns showed that all three models dramatically underestimate the empirical popularity of pronounced deck preferences, that is, a pronounced preference for good decks ( $C \geq .36$ ,  $D \geq .40$ ), a pronounced preference for bad decks ( $A \geq .25$ ,  $B \geq .52$ ), and a pronounced preference for decks with infrequent losses ( $B \geq .37$ ,  $D \geq .39$ ). Another finding of the analysis based on the restricted definition of choice patterns was that the EV model is the worst performing model among the three models that were compared in this article; it fails to generate a frequency-of-losses effect, and predicts very few switches when generating a pronounced preference for the good decks. Yet, a major advantage of the EV model is that it outperforms its competitors with respect to the choice pattern featuring a pronounced preference for the bad decks because the EV model can generate the largest range of switches during the last 50 trials. The EV-PU model, on the other hand, is the best model; comparing the ability of the three models to generate choice patterns featuring pronounced deck preferences, we found that all models underestimate the empirical popularity of such choice patterns, but also the proportions of switches during the last 50 trials; however, among the three models the EV-PU model best approximates the observed proportions of switches. Yet, a major disadvantage of the EV-PU model is that it fails to generate a pronounced preference for the bad decks ( $A \geq .25$ ,  $B \geq .52$ ). Thus, none of the three RL models met all empirical regularities, that is, pronounced deck preferences from either the good decks, the bad decks, or the decks with infrequent losses, and many switches during the last 50 trials.

The results of the PSP studies reveal important differences in the data-fitting potential of the three RL models. These findings are useful for future applications of the models. Based on the behavioral results, researchers can decide which model is most appropriate. For instance, our findings suggest that the EV model should not be used if the data set is characterized by a strong preference for the decks with infrequent losses.

The differences in the data-fitting potential of the three models might explain why previous comparisons of RL models found inconsistent results. Our finding that the EV model fails to generate a frequency-of-losses effect is in line with the studies of Fridberg et al. (2010) and Yechiam and Busemeyer (2005). These two studies show that the EV model is outperformed by its competitors when participants show a strong preference for decks with infrequent losses (decks B and D) over decks with frequent losses (deck A and C). However, in the study of Ahn et al. (2008) participants prefer the good decks over the bad decks and the authors recommend to use the EV-PU model instead of the PVL and EV models. This finding is also in line with our PSP studies because the choice pattern “good decks over bad decks” is the most central in the case of the EV-PU model.

An important advantage of PSP is that it is a global analysis technique augmenting local methods that have previously been used to compare RL models (Pitt et al., 2006, 2008). Whereas local methods, such as the post hoc fit criterion or the generalization criterion, evaluate a model's performance at a single point of a model's parameter space, global methods



such as PSP help us to determine the full range of choice patterns that a model can generate by varying its parameter values (see also Vanpaemel, 2009). This means that we can obtain a global perspective on a model's data-fitting potential. In addition, PSP allows us to draw conclusions about model flexibility that transcend the idiosyncrasies of any particular data set. We can use the results of the presented PSP studies to predict how well each of the three models will fit a certain data set only based on the behavioral findings. Thus, based on the observed choice pattern, we can decide which model is most appropriate.

One might argue that it is not surprising that the EV model fails to outperform its competitors, a failure that is due to the EV model having one free parameter less than its competitors. This implies that our comparison of RL models might not be fair because PSP does not correct for the number of free parameters. Even though we agree that PSP might disadvantage the EV model, it is important to bear in mind that previous comparisons of RL models based on methods that corrected for the number of free parameters (e.g., the post hoc fit criterion), also failed to find that the EV model is superior to its competitors (Ahn et al., 2008; Fridberg et al., 2010; Yechiam & Busemeyer, 2005). In addition, our goal was not a model-selection exercise in which the emphasis is on relative performance. Instead, we were primarily concerned with absolute measure of the models' performance, that is do the models account for the key regularities of the experimental data?

It should also be noted that the PSP results of this paper should be interpreted carefully. PSP gives an indication of how central choice patterns are to the overall performance of the model. But, when a model generates a certain choice pattern only over a small part of the parameter space, it is premature to conclude that the model cannot generate this choice pattern at all. Instead, we can only conclude that this choice pattern is not central to the model's overall performance. Moreover, a model that generates nonhuman-like choice patterns for certain parameter combinations is not necessarily inadequate.

The findings of this article offer several suggestions for future research. First, we found that the three RL models have difficulties in generating pronounced deck preferences, even though these choice patterns are frequently found in empirical research on the IGT. This suggests that a better model is required if researchers aim to decompose performance on the IGT in its constituent psychological processes using RL models. Such a model should be able to produce the large diversity of choice patterns shown on the IGT, and it should also be able to produce the switches made on the IGT. However, developing a model that well accounts for IGT data might be very challenging because there are large individual differences in the data, and, at the same time, there are only a small number of trials providing a limited amount of information (Steingroever et al., 2013).

A second possible approach would be to work on the level of the task. The IGT is a popular neuropsychological task that has been applied in many studies to assess decision-making deficits of clinical populations (Bowman, Evans, & Turnbull, 2005; Toplak, Sorge, Benoit, West, & Stanovich, 2010). Yet, the IGT has been confronted by a substantial

number of criticism (Chiu & Lin, 2007; Chiu et al., 2008; Dunn et al., 2006; Lin et al., 2007; Steingroever et al., 2013). It is therefore desirable to develop a neuropsychological task that measures risky decision making in an experimental context and that circumvents some of the problems of the IGT. As such, the new task should reduce the within-group variability resulting in more homogeneous response profiles.

Third, no matter in which direction future research proceeds, it is important that researchers rigorously assess absolute model fit. To date, more than 30 studies have fit the EV or PVL models to IGT data to compare two or more groups (Agay et al., 2010; Ahn et al., 2011; Bass & Nussbaum, 2010; Bishara et al., 2009; Brambilla et al., 2012; Brown et al., 2012; Busemeyer & Stout, 2002; Cella et al., 2012; de Visser et al., 2010; Escartin et al., 2012; Farah et al., 2008; Fridberg et al., 2010; Gullo & Stieger, 2011; Hochman et al., 2010; Isella et al., 2008; Johnson et al., 2006; Kester et al., 2006; Kjome et al., 2010; Lane et al., 2006; Lev et al., 2008; Lovallo et al., 2006; Peatfield et al., 2012; Premkumar et al., 2008; Sevy et al., 2006, 2007; Stout et al., 2004; van den Bos et al., 2009; Wood et al., 2005; Yechiam et al., 2005; Yechiam, Hayden, et al., 2008; Yechiam, Kanz, et al., 2008; Yechiam et al., 2010). About one third of these studies did not report the model fit at all. The remaining studies assessed the model fit by comparing the EV or PVL model to a baseline model and thus obtained a fit statistic, e.g.,  $G^2$  or BIC (for example, Farah et al., 2008; Yechiam, Kanz, et al., 2008). The problem of such fit statistics is that they only assess the fit of one model relative to the fit of another model. Another approach is to plot the observed and predicted choices as a function of trial number (Busemeyer & Stout, 2002; Wood et al., 2005). These two studies, however, collapsed choice proportions over the two good decks, a procedure that leads to a loss of potentially diagnostic information. To determine whether a model fits the data well, it is important to assess the model fit for each deck separately as a function of the trial number. Only a good model fit allows us to draw valid conclusion from the model parameters. Especially comparisons of several groups are only meaningful if the model fits the data of all groups well.

The findings of this paper showed that the search for an IGT model has not yet come to an end. Applying PSP to three different RL models, we have obtained a deeper understanding of the models' behavior. The results of the PSP studies allow us to predict each model's data-fitting potential by only considering behavioral characteristics of the data, that is, mean proportions of choices from each deck and mean proportions of switches during the last 50 trials. Each model was demonstrated to provide a good fit to only a restricted number of choice patterns. In particular, the EV model provided a poor fit to choice patterns featuring a preference for decks with infrequent losses, whereas the PVL and EV-PU models provided a poor fit to choice patterns featuring a pronounced preference for the bad decks with many switches. This work suggests that a deeper understanding of risky decision making requires a better interplay between the current version of the IGT and RL models.

## Acknowledgments

We thank Fernie and Tunney (2006), Fridberg et al. (2010), Kjome et al. (2010), Premkumar et al. (2008), and Wood, Busemeyer, Koling, Cox, and Davis (2005) for providing the data used in this article.

## Notes

1. The baseline model assumes that a participant's probability of choosing a given deck on a given trial equals the overall proportion of choices the participant actually made from that deck.
2. For the third method, data was generated for each subject individually, but eventually averaged over all subjects to compute the predicted proportions of choices from each deck on each trial.
3. See Yechiam and Busemeyer (2008) and Yechiam and Ert (2007) for newly developed methods to compare RL models applied to experiments with two and three choice options, respectively.
4. Note that we did not use the data of Rodríguez-Sánchez et al. (2005), and Toplak et al. (2005) because we have received their data only in bins of several trials.
5. The deck selection profiles can be downloaded here: <https://dl.dropbox.com/u/12798592/DeckSelectionProfilesPSP.zip>
6. The EV model already fails to generate a weak frequency-of-losses effect with 28% choices from each decks B and D. The EV-PU model already fails to produce a weak preference for the bad decks with 27% choices from each decks A and B.

## 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*(4), 511–19. <http://dx.doi.org/10.1007/s00213-010-1853-4>
- 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*(8), 1376–1402. <http://dx.doi.org/10.1080/0364021080235299>
- 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*(2), 95–110. <http://dx.doi.org/10.1037/a0020684>
- Bass, S., & Nussbaum, D. (2010). Decision making and aggression in forensic psychiatric inpatients. *Criminal Justice and Behavior*, *37*(4), 365–83. <http://dx.doi.org/10.1177/0093854809360043>

- Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*(5304), 1293–95. <http://dx.doi.org/10.1126/science.275.5304.1293>
- Bishara, A. J., Pleskac, T. J., Fridberg, D. J., Yechiam, E., Lucas, J., Busemeyer, J. R., et al. (2009). Similar processes despite divergent behavior in two commonly used measures of risky decision making. *Journal of Behavioral Decision Making*, *22*(4), 435–54. <http://dx.doi.org/10.1002/bdm.641>
- 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*(6), 499–511. <http://dx.doi.org/10.1023/A:1012277125119>
- Bowman, C. H., Evans, C. E. Y., & Turnbull, O. H. (2005). Artificial time constraints on the Iowa gambling task: The effects on behavioural performance and subjective experience. *Brain and Cognition*, *57*, 21–25. <http://dx.doi.org/10.1016/j.bandc.2004.08.015>
- Brambilla, P., Perlini, C., Bellani, M., Tomelleri, L., Ferro, A., Cerruti, S., et al. (2012). Increased salience of gains versus decreased associative learning differentiate bipolar disorder from schizophrenia during incentive decision making. *Psychological Medicine*, *1*, 1–10. <http://dx.doi.org/10.1017/S0033291712001304>
- Brown, W., Anderson, L., Symington, M., & Paul, L. (2012). Decision-making in individuals with agenesis of the corpus callosum: Expectancy-Valence in the Iowa gambling task. *Archives of Clinical Neuropsychology*, *27*, 532–44. <http://dx.doi.org/10.1093/arclin/acs052>
- 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*, 253–62. <http://dx.doi.org/10.1037/1040-3590.14.3.253>
- 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*, 203–12. [http://dx.doi.org/10.1207/s15324826an1304\\_1](http://dx.doi.org/10.1207/s15324826an1304_1)
- Cavedini, P., Riboldi, G., D'Annunzi, 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*, 205–11. [http://dx.doi.org/10.1016/S0028-3932\(01\)00077-X](http://dx.doi.org/10.1016/S0028-3932(01)00077-X)
- Cavedini, P., Riboldi, G., Keller, R., D'Annunzi, A., & Bellodi, L. (2002). Frontal lobe dysfunction in pathological gambling patients. *Biological Psychiatry*, *51*, 334–41. [http://dx.doi.org/10.1016/S0006-3223\(01\)01227-6](http://dx.doi.org/10.1016/S0006-3223(01)01227-6)
- Cella, M., Dymond, S., Cooper, A., & Turnbull, O. H. (2012). Cognitive decision modelling of emotion-based learning impairment in schizophrenia: The role of awareness. *Psychiatry Research*, *196*, 15–19. <http://dx.doi.org/10.1016/j.psychres.2011.08.015>

- Chiu, Y.-C., & Lin, C.-H. (2007). Is deck C an advantageous deck in the Iowa gambling task? *Behavioral and Brain Functions, 3*, 1–11. <http://dx.doi.org/10.1186/1744-9081-3-37>
- 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*, 1–10. <http://dx.doi.org/10.1186/1744-9081-4-13>
- de Visser, L., van der Knaap, L., van de Loo, A., van der Weerd, C., Ohl, F., & van den Bos, R. (2010). Trait anxiety affects decision-making differently in healthy men and women: Towards gender-specific endophenotypes of anxiety. *Neuropsychologia, 48*, 1598–1606. <http://dx.doi.org/10.1016/j.neuropsychologia.2010.01.027>
- Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience & Biobehavioral Reviews, 30*, 239–71. <http://dx.doi.org/10.1016/j.neubiorev.2005.07.001>
- Escartin, G., Junqué, C., Juncadella, M., Gabarrós, A., Miquel, M. de, & Rubio, F. (2012). Decision-making impairment on the Iowa gambling task after endovascular coiling or neurosurgical clipping for ruptured anterior communicating artery aneurysm. *Neuropsychology, 26*, 172–80. <http://dx.doi.org/10.1037/a0024336>
- 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*, 313–23. <http://dx.doi.org/10.1016/j.trf.2008.01.005>
- 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., et al. (2010). Cognitive mechanisms underlying risky decision-making in chronic cannabis users. *Journal of Mathematical Psychology, 54*, 28–38. <http://dx.doi.org/10.1016/j.jmp.2009.10.002>
- 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*, 204–10. <http://dx.doi.org/10.1016/j.drugalcdep.2011.02.002>
- Hochman, G., Yechiam, E., & Bechara, A. (2010). Recency gets larger as lesions move from anterior to posterior locations within the ventromedial prefrontal cortex. *Behavioural Brain Research, 213*, 27–34. <http://dx.doi.org/10.1016/j.bbr.2010.04.023>
- Isella, V., Mapelli, C., Morielli, N., Pelati, O., Franceschi, M., & Appollonio, I. (2008). Age-related quantitative and qualitative changes in decision making ability. *Behavioural Neurology, 19*, 59–63.
- 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*, 668–76. <http://dx.doi.org/10.1017/S1355617706060802>

- Kester, H. M., Sevy, S., Yechiam, E., Burdick, K. E., Cervellione, K. L., & Kumra, S. (2006). Decision-making impairments in adolescents with early-onset schizophrenia. *Schizophrenia Research, 85*, 113–23. <http://dx.doi.org/10.1016/j.schres.2006.02.028>
- Kjome, K. L., Lane, S. D., Schmitz, J. M., Green, C., Ma, L., Prasla, I., et al. (2010). Relationship between impulsivity and decision making in cocaine dependence. *Psychiatry Research, 178*, 299–304. <http://dx.doi.org/10.1016/j.psychres.2009.11.024>
- Lane, S. D., Yechiam, E., & Busemeyer, J. R. (2006). Application of a computational decision model to examine acute drug effects on human risk taking. *Experimental and Clinical Psychopharmacology, 14*, 254–64. <http://dx.doi.org/10.1037/1064-1297.14.2.254>
- Lev, D., Hershkovitz, E., & Yechiam, M. (2008). Decision making and personality in traffic offenders: A study of Israeli drivers. *Accident Analysis and Prevention, 40*, 223–30. <http://dx.doi.org/10.1016/j.aap.2007.05.009>
- 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*, 1–10. <http://dx.doi.org/10.1186/1744-9081-3-16>
- 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.
- Lovallo, W. R., Yechiam, E., Sorocco, K. H., Vincent, A. S., & Collins, F. L. (2006). Working memory and decision-making biases in young adults with a family history of alcoholism: Studies from the Oklahoma family health patterns project. *Alcoholism: Clinical and Experimental Research, 30*, 763–73. <http://dx.doi.org/10.1111/j.1530-0277.2006.00089.x>
- Luce, R. (1959). *Individual choice behavior*. New York: Wiley.
- 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. <http://dx.doi.org/10.1037/0882-7974.17.4.598>
- Martino, D. J., Bucay, D., Butman, J. T., & Allegri, R. F. (2007). Neuropsychological frontal impairments and negative symptoms in schizophrenia. *Psychiatry Research, 152*, 121–28. <http://dx.doi.org/10.1016/j.psychres.2006.03.002>
- Peatfield, N., Parkinson, J., & Intriligator, J. (2012). Emotion-based learning is biased by brand logos. *Applied Cognitive Psychology, 26*, 694–701. <http://dx.doi.org/10.1002/acp.2847>
- Pitt, M., Kim, W., Navarro, D., & Myung, J. (2006). Global model analysis by parameter space partitioning. *Psychological Review, 113*, 57–83. <http://dx.doi.org/10.1037/0033-295X.113.1.57>
- Pitt, M., Myung, J., Montenegro, M., & Pooley, J. (2008). Measuring model flexibility with parameter space partitioning: An introduction and application example. *Cognitive Science, 32*, 1285–303. <http://dx.doi.org/10.1080/03640210802477534>
- 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*, 2002–12. <http://dx.doi.org/10.1016/j.neuropsychologia.2008.01.022>

- 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*, 279–88.
- Sevy, S., Burdick, K. E., Visweswarajah, H., Abdelmessih, S., Lukin, M., Yechiam, E., et al. (2007). Iowa gambling task in schizophrenia: A review and new data in patients with schizophrenia and co-occurring cannabis use disorders. *Schizophrenia Research*, *92*, 74–84. <http://dx.doi.org/10.1016/j.schres.2007.01.005>
- Sevy, S., Hassoun, Y., Bechara, A., Yechiam, E., Napolitano, B., Burdick, K., et al. (2006). Emotion-based decision-making in healthy subjects: short-term effects of reducing dopamine levels. *Psychopharmacology*, *188*, 228–35. <http://dx.doi.org/10.1007/s00213-006-0450-z>
- Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., & Wagenmakers, E.-J. (2013). Performance of healthy participants on the Iowa gambling task. *Psychological Assessment*, *25*, 180–93.
- 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*, 742–47. <http://dx.doi.org/10.3758/BF03196629>
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: The MIT Press.
- 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. <http://dx.doi.org/10.1186/1744-9081-1-8>
- Toplak, M., Sorge, G., Benoit, A., West, R., & Stanovich, K. (2010). Decision-making and cognitive abilities: A review of associations between Iowa gambling task performance, executive functions, and intelligence. *Clinical Psychology Review*, *30*, 562–81. <http://dx.doi.org/10.1016/j.cpr.2010.04.002>
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*, 297–323. <http://dx.doi.org/10.1007/BF00122574>
- van den Bos, R., Homberg, J., Gijsbers, E., Heijer, E. den, & Cuppen, E. (2009). The effect of COMT Val<sup>158</sup>Met genotype on decision-making and preliminary findings on its interaction with the 5-HTTLPR in healthy females. *Neuropharmacology*, *56*, 493–98. <http://dx.doi.org/10.1016/j.neuropharm.2008.10.002>
- Vanpaemel, W. (2009). Measuring model complexity with the prior predictive. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22: 23<sup>rd</sup> Annual Conference on Neural Information Processing Systems* (p. 1919–27). Red Hook, NY: Curran Associates Inc.
- 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*, 169–74. [http://dx.doi.org/10.1016/S0920-9964\(97\)00135-7](http://dx.doi.org/10.1016/S0920-9964(97)00135-7)

- Wood, S., Busemeyer, J., Kolling, A., Cox, C. R., & Davis, H. (2005). Older adults as adaptive decision makers: Evidence from the Iowa gambling task. *Psychology and Aging, 20*, 220–25. <http://dx.doi.org/10.1037/0882-7974.20.2.220>
- Yechiam, E., Arshavsky, O., Shamay-Tsoory, S. G., Yaniv, S., & Aharon, J. (2010). Adapted to explore: Reinforcement learning in Autistic Spectrum Conditions. *Brain and Cognition, 72*, 317–24. <http://dx.doi.org/10.1016/j.bandc.2009.10.005>
- Yechiam, E., & Busemeyer, J. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin & Review, 12*, 387–402. <http://dx.doi.org/10.3758/BF03193783>
- Yechiam, E., & Busemeyer, J. R. (2008). Evaluating generalizability and parameter consistency in learning models. *Games and Economic Behavior, 63*, 370–94. <http://dx.doi.org/10.1016/j.geb.2007.08.011>
- 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*, 973–78. <http://dx.doi.org/10.1111/j.1467-9280.2005.01646.x>
- Yechiam, E., & Ert, E. (2007). Evaluating the reliance on past choices in adaptive learning models. *Journal of Mathematical Psychology, 51*, 75–84. <http://dx.doi.org/10.1016/j.jmp.2006.11.002>
- Yechiam, E., Hayden, E. P., Bodkins, M., O'Donnell, B. F., & Hetrick, W. P. (2008). Decision making in bipolar disorder: A cognitive modeling approach. *Psychiatry Research, 161*, 142–52. <http://dx.doi.org/10.1016/j.psychres.2007.07.001>
- Yechiam, E., Kanz, J. E., Bechara, A., Stout, J. C., Busemeyer, J. R., Altmaier, E. M., et al. (2008). Neurocognitive deficits related to poor decision making in people behind bars. *Psychonomic Bulletin & Review, 15*, 44–51. <http://dx.doi.org/10.3758/PBR.15.1.44>
- Zhao, J., & Costello, F. (2007). Computational modelling of switching behaviour in repeated gambles. *Artificial Intelligence Review, 27*, 209–22. <http://dx.doi.org/10.1007/s10462-008-9083-4>