Safe models for risky decisions

Steingröver, H.M.

Publication date
2017
Document Version
Other version
License
Other

Chapter 7

Bayes Factors for Reinforcement-Learning Models of the Iowa Gambling Task

This chapter has been published as:
Helen Steingroever, Ruud Wetzels, and Eric-Jan Wagenmakers (2016).
Bayes factors for reinforcement-learning models of the Iowa gambling task.
Decision, 3, 115–131.[1]

Abstract

The psychological processes that underlie performance on the Iowa gambling task (IGT) are often isolated with the help of reinforcement-learning (RL) models. The most popular method to compare RL models is the BIC post hoc fit criterion—a criterion that considers goodness-of-fit relative to model complexity. However, the current implementation of the BIC post hoc fit criterion considers only one dimension of model complexity, that is, the number of free parameters. A more sophisticated implementation of the BIC post hoc fit criterion, one that provides a coherent and complete discounting of complexity, is provided by the Bayes factor. Here we demonstrate an analysis in which Bayes factors are obtained with a Monte Carlo method, known as importance sampling, in order to compare four RL models of the IGT: the Expectancy Valence (EV), Prospect Valence Learning (PVL), PVL-Delta, and Value-Plus-Perseveration (VPP) models. We illustrate the method using a data pool of 771 participants from 11 different studies. Our results provide strong evidence for the VPP model and moderate evidence for the PVL model, but little evidence for the EV and PVL-Delta models—results that were not in line with a BIC post hoc fit analysis. We discuss how our results may be combined with results obtained from other model comparison studies in order to obtain a balanced and comprehensive assessment of model adequacy.

The Iowa gambling task (IGT; Bechara et al., 1994) is arguably the most popular neuropsychological paradigm to assess decision-making deficits in clinical populations. In order

[1]The final publication is available at `http://psycnet.apa.org/psycinfo/2015-49436-001/`.

to isolate and identify the psychological processes that drive performance on the IGT, behavioral analyses of IGT data are insufficient. A promising additional analysis approach is to use reinforcement-learning (RL) models that try to disentangle the psychological processes underlying performance on the IGT. Two of the most frequently used representatives include the Expectancy Valence model (EV), and the Prospect Valence Learning model (PVL; see Steingroever, Wetzels, & Wagenmakers, 2013a, for references). The parameters of these models correspond to distinct psychological processes such as motivation, learning/memory, and response consistency (Busemeyer et al., 2003). However, recent research suggests that there might be better IGT models; promising alternatives include the PVL-Delta model which is a hybrid version of the EV and PVL models (Ahn et al., 2008; Fridberg et al., 2010; Steingroever et al., 2014; Steingroever, Wetzels, & Wagenmakers, 2013b), and the Value-Plus-Perseveration (VPP) model which includes an additional perseveration process (Worthy, Pang, & Byrne, 2013).

RL models for the IGT have been tested and compared using a wide variety of methods that focus on different aspects of the models. These methods include, for instance: the post hoc fit criterion (i.e., Ahn et al., 2008; Busemeyer & Stout, 2002; Fridberg et al., 2010; Worthy, Hawthorne, & Otto, 2013; Yechiam & Busemeyer, 2005; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008);[2] the simulation method (i.e., Ahn et al., 2008; Fridberg et al., 2010; Steingroever et al., 2014; Worthy, Hawthorne, & Otto, 2013; Worthy, Pang, & Byrne, 2013); tests of generalizability (i.e., Ahn et al., 2008; Yechiam & Busemeyer, 2005; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008); tests of parameter consistency (i.e., Yechiam & Busemeyer, 2008); parameter space partitioning (i.e., Steingroever, Wetzels, & Wagenmakers, 2013a, 2013b); and tests of specific influence (i.e., Steingroever, Wetzels, & Wagenmakers, 2013b; Wetzels, Vandekerckhove, et al., 2010). In addition, several studies have carefully investigated the ability of the EV, PVL, PVL-Delta, and VPP models to recover the data-generating parameters (Ahn et al., 2011, 2014; Steingroever, Wetzels, & Wagenmakers, 2013b; Wetzels, Vandekerckhove, et al., 2010). A good parameter recovery is important because it signals that the estimated parameter values are reliable reflections of the underlying processes (Steingroever, Wetzels, & Wagenmakers, 2015).

Among these different methods to compare and evaluate RL models for the IGT, the Bayesian information criterion (BIC; Schwarz, 1978) has proven to be the most popular. Since it is termed "post hoc fit criterion" (or "one-step-ahead prediction method") in the context of RL models for the IGT, in the remainder of this article we refer to this method as "BIC post hoc fit criterion" to emphasize that the BIC and post hoc fit criterion are essentially the same method. This criterion implements a tradeoff between goodness-of-fit (i.e., descriptive adequacy of a model) and parsimony (i.e., a model should be as simple as possible), as follows:

$$\text{BIC}_{\mathcal{M}_{(.)}} = \underbrace{-2\log\Big(L_{(.)}\Big)}_{\text{Descriptive adequacy}} + \underbrace{k_{(.)}\log(n)}_{\text{Penalty term}}, \tag{7.1}$$

where $L_{(.)}$ is the maximum likelihood of model $\mathcal{M}_{(.)}$, $k_{(.)}$ is the number of free parameters of model $\mathcal{M}_{(.)}$, and n is number of IGT trials. Equation 7.1 illustrates how the BIC post hoc fit criterion uses the penalty term $k_{(.)}\log(n)$ to discount descriptive adequacy $-2\log\Big(L_{(.)}\Big)$. It is also apparent that the BIC post hoc fit criterion only considers one dimension of model complexity, that is, the number of free parameters. On the other hand, the gold standard of model comparison in Bayesian statistics –the Bayes factor– also takes into account two additional dimensions of model complexity, that is, the functional form of the model (i.e., the way in which the parameters are combined in

[2] In the context of RL models for the IGT, the post hoc fit criterion is also known as the one-step-ahead prediction method.

the model equations), and the extension of the parameter space (i.e., the prior distributions over parameters; I. J. Myung & Pitt, 1997). In this way the Bayes factor implements the tradeoff between goodness-of-fit and parsimony in a manner that is more comprehensive than that used by the current implementation of the BIC post hoc fit criterion. However, the Bayes factor has not yet been used to compare RL models for the IGT.

In this article we use Bayes factors to compare four popular RL models: the EV, PVL, PVL-Delta, and VPP models.[3] Bayes factors are derived using importance sampling, a numerical technique developed by Stanislaw Ulam and John von Neumann (Eckhardt, 1987; Kass & Raftery, 1995; Hammersley & Handscomb, 1964; Robert & Casella, 2011; Vandekerckhove, Matzke, & Wagenmakers, 2015). We illustrate this method with a large IGT data pool containing 771 healthy participants from 11 studies.

The outline of this article is as follows. Sections $1 - 3$ explain the IGT, the different models (i.e., the EV, PVL, PVL-Delta, and VPP models), and the Bayes factor, respectively. In the fourth section we explain how Bayes factors can be obtained with importance sampling, and in the fifth section we report a Bayes factor analysis of the four RL models with IGT data from 771 healthy participants. This analysis allows us to compare the EV, PVL, PVL-Delta, and VPP models by means of the Bayes factor. In the last section, we summarize our findings and discuss their ramifications, and how our results may be combined with results obtained from other model comparison studies in order to obtain a balanced and comprehensive assessment of model adequacy. In Appendix E, we present a recipe on how to obtain Bayes factors with importance sampling, two tests to check our implementation of importance sampling: (1) a model-recovery study, and (2) the Savage-Dickey density ratio test for each model. In addition, Appendix E contains a robustness analysis, and a model comparison study using BIC for the same models and data pool as used in this article. To anticipate our main result, our model comparison shows that the data provide strong evidence for the VPP model and moderate evidence for the PVL model, but weak evidence for the EV and PVL-Delta models.

7.1 The Iowa Gambling Task

In this section we describe the IGT (see also Steingroever, Wetzels, Horstmann, et al., 2013; Steingroever, Wetzels, & Wagenmakers, 2013a, 2013b; Steingroever et al., 2014). 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 (Bechara et al., 1994, 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 7.1 presents the main characteristics of 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,

[3] Even though it might be interesting to include additional models obtained from different combinations of the utility function, learning rule and sensitivity function, we decided to focus on these four RL models because of results from previous model comparison studies (e.g., Ahn et al., 2008, 2014; Busemeyer & Stout, 2002; Yechiam & Busemeyer, 2005) and because of their popularity in the field of RL models for the IGT (e.g., Ahn et al., 2011; Fridberg et al., 2010; Steingroever et al., 2014; Stout et al., 2004; Wood et al., 2005; Yechiam et al., 2005).

Table 7.1: *Main characteristics of the 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

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).

7.2 The EV, PVL, PVL-Delta, and VPP Models

In this section, we describe the EV, PVL, PVL-Delta, and VPP models (see also Steingroever, Wetzels, & Wagenmakers, 2013a, and Steingroever et al., 2014). Table 7.2 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 performance on the IGT through the interaction of distinct psychological processes captured by the model parameters.

The four RL 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-of-strength 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 the trial-by-trial choice probabilities of the decks 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 article, we scaled the traditional payoffs of the IGT as presented in Table 7.1 by dividing by 100 (cf. Ahn et al., 2011).

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

Table 7.2: *Formalization of the EV, PVL, PVL-Delta, and VPP 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, & VPP	$u_k(t) = \begin{cases} X(t)^A & \text{if } X(t) \geq 0 \\ -w \cdot \lvert X(t)\rvert^A & \text{if } X(t) < 0 \end{cases}$	A: Shape w: Loss aversion	$[0,1]$ $[0,5]$
Learning rule	EV, PVL-Delta, & VPP	$Ev_k(t) = Ev_k(t-1) + a \cdot (u_k(t) - Ev_k(t-1))$	a: Updating	$[0,1]$
	PVL	$Ev_k(t) = a \cdot Ev_k(t-1) + \delta_k(t) \cdot u_k(t)$	a: Recency	$[0,1]$
Perseveration	VPP	$P_k(t) = \begin{cases} d \cdot P_k(t-1) + \delta_k(t) \cdot \epsilon_{pos} & \text{if } X(t) \geq 0 \\ d \cdot P_k(t-1) + \delta_k(t) \cdot \epsilon_{neg} & \text{if } X(t) < 0 \end{cases}$	d: Decay ϵ_{pos} ϵ_{neg}	$[0,1]$ $[-1,1]$ $[-1,1]$
		$Ev_k(t) = w_{Ev} \cdot Ev_k(t) + (1-w_{Ev}) \cdot P_k(t)$	w_{Ev}: Expectancy weight	$[0,1]$
Choice rule	All	$Pr[S_k(t+1)] = \frac{e^{\theta(t)Ev_k(t)}}{\sum_{j=1}^{4} e^{\theta(t)Ev_j(t)}}$		
Sensitivity	EV	$\theta(t) = (t/10)^c$	c: Consistency	$[-2,2]$
	PVL, PVL-Delta, & VPP	$\theta(t) = 3^c - 1$	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.

reward *W(t)* and loss *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, $Ev_k(0) = 0$ ($k \in \{1, 2, 3, 4\}$). 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 (see also Steingroever et al., 2014).

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 balance between exploitation and exploration.

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 balance between exploitation and exploration.

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; Steingroever, Wetzels, & Wagenmakers, 2013b; Steingroever et al., 2014). 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 balance between exploitation and exploration.

The VPP Model

The VPP model uses eight parameters to formalize its assumptions about participants' performance on the IGT (Worthy, Pang, & Byrne, 2013). The VPP model is a more complex version of the PVL-Delta model that consists of the same utility function, learning rule, and sensitivity parameter as the PVL-Delta model, but also includes preservation—an additional psychological process representing participants' tendency to stay with the same option (i.e., to persevere), or to switch. The VPP model assumes that a participant's tendency to persevere on deck k decays on each trial by d, and if deck k is chosen on trial t an additional value ϵ_{pos} or ϵ_{neg} –depending on whether the net outcome on trial t is positive or negative– is added to the perseveration of deck k. Positive values of ϵ_{pos} or ϵ_{neg} indicate a tendency of persevere (i.e., to take the same option on the following trial), whereas negative values indicate a tendency to switch.

The VPP model considers perseveration and maximization of expected utility as two fundamental, but separate psychological processes involved in decision making. A weighted average of these two processes (i.e., w_{Ev} represents the weight for the expected utilities of the decks, and $1 - w_{Ev}$ represents the weight for the perseveration process) is used in the softmax choice rule to determine the choice probabilities of all decks on the next trial. A large value of w_{Ev}, that is, $w_{Ev} > .5$, is characteristic for decision makers who put more weight on the expected utility of the decks, whereas a small value of w, that is, $w_{Ev} < .5$, is characteristic for decision makers who put more weight on the perseveration of the decks. As in Worthy, Pang, and Byrne (2013), we initialized the perseveration of all decks to zero (i.e., $P_k(0) = 0$ for $k \in \{1, 2, 3, 4\}$).

In sum, the VPP model has eight 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; (4) the decay parameter d, which determines how strongly the perseveration of each deck decays; (5) ϵ_{pos} and (6) ϵ_{neg}, which quantify the tendency to persevere given positive and negative net outcomes, respectively; (7) the weight of the expected utility w_{Ev}, which quantifies the weight given to the expected utility relative to the perseveration of each deck, and (8) the response consistency parameter c, which determines the balance between exploitation and exploration.

7.3 The Bayes Factor

In this section we describe how two models, $\mathcal{M}_1$ and $\mathcal{M}_2$, can be compared using the Bayes factor (e.g., Jeffreys, 1961; Kass & Raftery, 1995). The Bayes factor BF_{12} is defined as the change from prior model odds $p(\mathcal{M}_1)/p(\mathcal{M}_2)$ to posterior model odds $p(\mathcal{M}_1 \mid y)/p(\mathcal{M}_2 \mid y)$ brought about by the data y:

$$\underbrace{\frac{p(\mathcal{M}_1 \mid y)}{p(\mathcal{M}_2 \mid y)}}_{\text{Posterior model odds}} = \underbrace{\frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}}_{\text{Prior model odds}} \times \underbrace{\frac{m(y \mid \mathcal{M}_1)}{m(y \mid \mathcal{M}_2)}}_{\text{Bayes factor}} \tag{7.2}$$

Thus, the Bayes factor is the ratio of the marginal likelihoods of the two models: $\text{BF}_{12} = m(y \mid \mathcal{M}_1)/m(y \mid \mathcal{M}_2)$. The Bayes factor can range from zero to infinity. For ease of interpretation the Bayes factor is often transformed to the log scale, where $\log(\text{BF}_{12}) = 0$ indicates that observed data are equally likely to occur under both models, $\log(\text{BF}_{12}) = \log(10)$ indicates that the data are 10 times more likely to occur under model $\mathcal{M}_1$ than model $\mathcal{M}_2$, and $\log(\text{BF}_{12}) = \log(.10)$ indicates that the data are 10 times more likely to occur under model $\mathcal{M}_2$ than model $\mathcal{M}_1$. To get an indication of the strength of evidence measured with the Bayes factor, Jeffreys (1961) proposed to divide the continuous scale of the Bayes factor into discrete categories of evidential strength. For example, a $\log(\text{BF}_{12})$ between $\log(3)$ and $\log(10)$ is characterized as moderate evidence for $\mathcal{M}_1$, whereas a $\log(\text{BF}_{12})$ between $\log(10)$ and $\log(30)$ is characterized as strong evidence for $\mathcal{M}_1$.

Note that the marginal likelihood is the likelihood of the data averaged across the entire parameter space, with the prior acting as averaging weights. That is, the marginal likelihood of the data y given model $\mathcal{M}_{(.)}$ is obtained by integrating the likelihood over the prior:

$$\underbrace{m(y \mid \mathcal{M}_{(.)})}_{\text{Marginal likelihood}} = \int \underbrace{p(y \mid \theta, \mathcal{M}_{(.)})}_{\text{Likelihood}} \underbrace{p(\theta \mid \mathcal{M}_{(.)})}_{\text{Prior}} \mathrm{d}\theta, \tag{7.3}$$

with θ being a vector containing the model parameters. Equation 7.3 illustrates why the Bayes factor automatically accounts for the tradeoff between parsimony and goodness-of-fit (e.g., I. J. Myung & Pitt, 1997): Although complex models may provide a good fit to the data as indicated by a high maximum likelihood, these models may have a high-dimensional parameter space with large regions that yield a poor fit to the data (i.e., a low likelihood). Parsimonious models, on the other hand, use only a small part of the parameter space, and do not contain large parameter regions that yield a poor fit to the data; consequently, the average likelihood of parsimonious models may be higher than that of overly complex models. Thus, by considering the entire parameter space the Bayes factor penalizes overly complex models; models that make vague predictions by including extra parameters, by assigning very wide prior distributions to the model parameters, and/or by using parameters that affect the likelihood through a complicated

functional form of the model (Busemeyer, Wang, & Shiffrin, in press; I. J. Myung & Pitt, 1997; Vandekerckhove et al., 2015; Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010). By averaging the likelihood over the prior distribution, the Bayes factor focuses on the extent to which models make good predictions, thereby taking into account, automatically and simultaneously, all three different dimensions of model complexity (i.e., number of free parameters, functional form of the model, and the extension of the parameter space). Thus, the main thrust of our work is that the current version of the BIC post hoc fit method for comparing RL models can be made more sophisticated by using Bayes factors instead of only penalizing maximum likelihood through the number of free parameters. The next section outlines a method that shows how this can be accomplished in a way that is practical and efficient.

7.4 Obtaining Bayes Factors for RL Models Using Importance Sampling

In the last section, we introduced the Bayes factor as the ratio of the marginal likelihood of two models. The question that now arises is how a model's marginal likelihood $m(y \mid \mathcal{M}_{(.)})$ can be obtained. Unfortunately, we cannot analytically solve the integral shown in Equation 7.3 because of the complex functional form of the likelihood of the four models. An obvious alternative is to use brute force integration to approximate $m(y \mid \mathcal{M}_{(.)})$:

$$\underbrace{m(y \mid \mathcal{M}_{(.)})}_{\text{Marginal likelihood}} \approx \underbrace{\frac{1}{N}\sum_{i=1}^{N} p(y \mid \theta_i, \mathcal{M}_{(.)})}_{\text{Average likelihood}}, \qquad \underbrace{\theta_i \sim p(\theta \mid \mathcal{M}_{(.)})}_{\text{Samples from the prior}} . \tag{7.4}$$

Note that for the sake of clarity we suppressed the notation that indexes a specific participant and model. Equation 7.4 illustrates that the brute force integration consists of drawing N samples from the prior $p(\theta \mid \mathcal{M}_{(.)})$, and then averaging the corresponding values for $p(y \mid \theta_i, \mathcal{M}_{(.)})$. However, brute force integration might be very inefficient, especially when the posterior distribution is highly peaked relative to the prior. In such a scenario, most draws from the prior result in low likelihoods, whereas only a few draws result in high likelihoods, increasing the variability of the estimator. A more efficient way to estimate a model's marginal likelihood $m(y \mid \mathcal{M}_{(.)})$ is to use a numerical method known as importance sampling (for a detailed description see Eckhardt, 1987; Kass & Raftery, 1995; Hammersley & Handscomb, 1964; Vandekerckhove et al., 2015). This method consists of sampling from an importance density $g(\theta \mid \mathcal{M}_{(.)})$ instead of the prior:

$$\begin{aligned}
\underbrace{m(y \mid \mathcal{M}_{(.)})}_{\text{Marginal likelihood}} &= \int p(y \mid \theta, \mathcal{M}_{(.)}) p(\theta \mid \mathcal{M}_{(.)}) \mathrm{d}\theta \\
&= \int p(y \mid \theta, \mathcal{M}_{(.)}) p(\theta \mid \mathcal{M}_{(.)}) \frac{g(\theta \mid \mathcal{M}_{(.)})}{g(\theta \mid \mathcal{M}_{(.)})} \mathrm{d}\theta \\
&= \int \frac{p(y \mid \theta, \mathcal{M}_{(.)}) p(\theta \mid \mathcal{M}_{(.)})}{g(\theta \mid \mathcal{M}_{(.)})} g(\theta \mid \mathcal{M}_{(.)}) \mathrm{d}\theta \\
&\underbrace{\approx \frac{1}{N}\sum_{i=1}^{N} \frac{p(y \mid \theta_i, \mathcal{M}_{(.)}) p(\theta_i \mid \mathcal{M}_{(.)})}{g(\theta_i \mid \mathcal{M}_{(.)})}}_{\text{Average adjusted likelihood}}, \qquad \underbrace{\theta_i \sim g(\theta \mid \mathcal{M}_{(.)}).}_{\text{Samples from the importance density}}
\end{aligned} \tag{7.5}$$

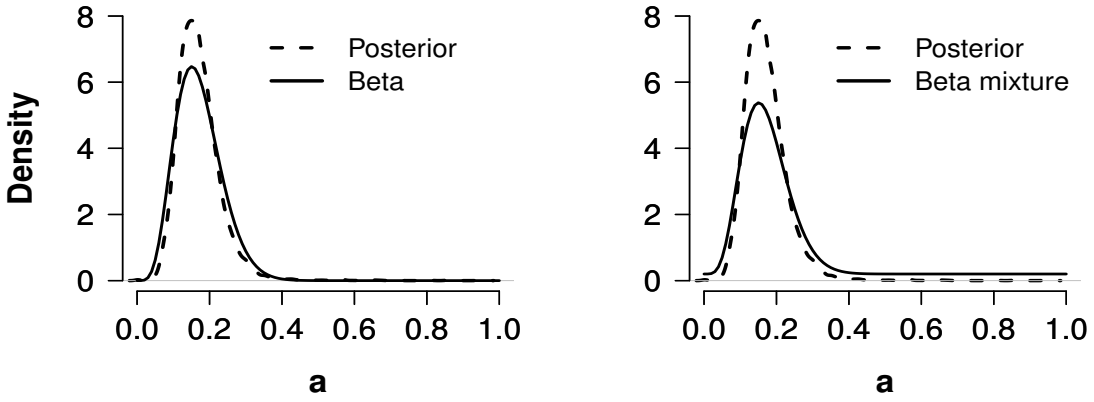

Figure 7.1: Two different importance sampling densities (solid lines) for the posterior distribution (dashed lines) of the a parameter in the EV model. Left panel: a Beta posterior importance density (i.e., a Beta distribution that provides the best fit to the posterior); right panel: a Beta mixture importance density (i.e., a mixture of the uniform Beta density and the Beta posterior density, with a mixture weight $w_{IS} = 0.2$ on the uniform component).

The advantage of sampling from an importance density is that, if the importance density resembles the posterior distribution, most samples come from regions with high likelihood causing the estimator to have low variability. In our application to RL models, we used an independent Beta mixture importance density (see solid line in the right panel of Figure 7.1), that is, a density that is a mixture between a Beta(1, 1) density, and a Beta density that provides the best fit to the posterior (see left panel of Figure 7.1). The Beta mixture importance density has several advantages: It is an importance density that strongly resembles the posterior, is easy to evaluate, and is easy to sample from (Vandekerckhove et al., 2015). In addition, this importance density has tails that are fatter than those of the posterior, which is beneficial because thin tails cause the estimate to have high variance (Vandekerckhove et al., 2015). Finally, the Beta mixture importance density is a good candidate distribution because it is restricted to the [0, 1] range, just as the parameters from the models under scrutiny.[4] In our subsequent analyses, we drew $N = 20,000$ samples from the Beta mixture importance density, and set the mixture weight for the uniform component to $w_{IS} = 0.10$. Appendix E provides a recipe on how to obtain Bayes factors with importance sampling.

7.5 Application to IGT Data from 771 Healthy Participants

In the previous section we described how Bayes factors can be obtained with importance sampling in a relatively straightforward fashion. We now apply this methodology to a large IGT data pool consisting of 771 healthy participants from 11 different experiments in order to compare four RL models. Most of the data are published in Steingroever, Davis, et al. (2015). Below we first describe the details of the analysis procedure and then present the results.

[4] Parameters with different ranges (see Table 7.2) were transformed to the $[0, 1]$ interval, and were only transformed back to their correct ranges after the analysis was complete.

Table 7.3: *Data sets used in this article.*

Study	Number of participants	Number of IGT trials
Busemeyer and Stout (2002)	30	100
Fridberg et al. (2010)	15	95
Horstmann[a]	162	100
Kjome et al. (2010)	19	100
Maia and McClelland (2004)	40	100
Premkumar et al. (2008)	25	100
Steingroever, Pachur, Šmíra, and Lee (submitted)	70	100
Wetzels, Vandekerckhove, et al. (2010)[b]	165	150
Wood et al. (2005)	153	100
Worthy, Pang, and Byrne (2013)	35	100
Own unpublished data set	57	150
Total	771	

[a] Data collected by Annette Horstmann. These data were first published in Steingroever, Wetzels, Horstmann, et al. (2013).
[b] Data of four different conditions: Standard condition ($N = 41$), rewards condition ($N = 42$), updating condition ($N = 41$), and consistency condition ($N = 41$).

Method

We fit the four RL models to individual IGT data of 771 participants from 11 different studies (Table 7.3; for more details on the data sets, see the original studies, Steingroever, Davis, et al., 2015, or Steingroever, Wetzels, Horstmann, et al., 2013, for details on five of the data sets). We fit the data of each participant using Stan (Stan Development Team, 2014b, 2014c; Hoffman & Gelman, 2014). For each parameter, we ran two Hamiltonian Monte Carlo (HMC) chains simultaneously. We used random starting values, and collected 5,000 samples of each chain after having discarded the first 1,000 samples of each chain as burn-in. We assigned uniform priors (i.e., $\mathrm{U}(0, 1)$) to all model parameters; parameters with different ranges (see Table 7.2) were transformed to the $[0, 1]$ interval, and were only transformed back to their correct ranges after the analysis was complete. All code is available on `www.helensteingroever.com`.

To assess whether the chains of all parameters had converged successfully from their starting values to their stationary distributions, we assessed convergence with the $\hat{R}$ statistic (Gelman & Rubin, 1992). The $\hat{R}$ statistic is a formal diagnostic measure of convergence that compares the between-chain variability to the within-chain variability. Values close to 1.0 indicate convergence to the stationary distribution, whereas values greater than 1.1 generally indicate inadequate convergence.

If a data set resulted in posterior distributions with $\hat{R}$ statistics larger than 1.05, we fitted the corresponding data set again with an additional 5,000 samples per chain. We repeated this process until all $\hat{R}$ statistics of that data set were smaller than 1.05.

For each of the 771 participants, we quantified the relative support for each of the four RL models by computing the Bayes factor using the method of importance sampling.[5] For each participant we obtained $\binom{4}{2} = 6$ Bayes factors (where 4 is the number of RL models that we compare here). We summarize the results by presenting histograms of the Bayes factors across all participants, and violin plots of the posterior model probabilities. Using equal model priors, the

[5] A model comparison study using BIC for the same models and data pool as used in this article can be found in Appendix E.

Table 7.4: *Median, 25% to 75% quantile range (i.e., interquartile range), and 2.5% to 97.5% quantile range of the standard error of the Bayes factor estimates. The standard error is presented as percentage of the Bayes factor.*

	median [%]	25%, 75% quantile [%]	2.5%, 97.5% quantile [%]
$\text{SE}(\text{BF}_{\text{EV,PVL}})$	1.89	[1.36, 2.77]	[0.93, 7.92]
$\text{SE}(\text{BF}_{\text{EV,PVL-D}})$	2.21	[1.54, 3.55]	[1.05, 12.32]
$\text{SE}(\text{BF}_{\text{EV,VPP}})$	5.69	[4.03, 8.06]	[2.14, 23.70]
$\text{SE}(\text{BF}_{\text{PVL,PVL-D}})$	2.13	[1.54, 3.26]	[1.10, 8.71]
$\text{SE}(\text{BF}_{\text{PVL,VPP}})$	5.47	[4.00, 7.97]	[2.09, 22.34]
$\text{SE}(\text{BF}_{\text{PVL-D,VPP}})$	5.74	[4.27, 8.27]	[2.25, 22.83]

posterior model probability of model $\mathcal{M}_{(.)}$ given data y is defined by:

$$P(\mathcal{M}_{(.)} \mid y) = \frac{\text{BF}_{(.)B}}{\sum_{j=1}^{4} \text{BF}_{jB}}, \tag{7.6}$$

where B indicates the reference model which can be any of the four RL models (Berger & Molina, 2005). We also used these subject-specific posterior model probabilities to determine the most likely model for each participant. This allowed us to report the percentage of participants for whom each of the four RL models has the largest posterior model probability.

Results

Before applying our implementation of importance sampling to each of the 771 individual data sets, we checked the adequacy of our implementation by comparing it to the Savage-Dickey density ratio test,[6] and by conducting a model recovery study (see Appendix E). In addition, we checked the adequacy of our implementation by varying the mixture weight w_{IS}, and the number of draws N from the Beta mixture importance density.

Visual inspection of a sample of the HMC chains, and consideration of the $\hat{R}$ statistics for all parameters (all parameters had $\hat{R}$ values below 1.05) suggest that the chains of all parameters had converged successfully from their starting values to their stationary distributions. In order to quantify the extent to which sampling error may perturb the Bayes factor estimates, we computed bootstrapped standard errors for each participant and each Bayes factor. For each Bayes factor comparison separately, Table 7.4 presents summary information based on the 771 individual standard errors. Specifically, Table 7.4 shows the median standard error, the standard error interquartile range, and the standard error 2.5% – 97.5% quantile range. Most standard errors are small relative to the Bayes factor, underscoring the precision of the estimates. However, for a few participants and model comparisons there remains some sampling uncertainty about the estimate of the Bayes factor as reflected by a large value of the 97.5% quantile. The results of our Bayes factor model comparison efforts are presented in Figures 7.2 and 7.3, and Table 7.5.

Figure 7.2 shows the distribution of the log Bayes factors of all participants for the six possible model comparisons. A positive $\log(\text{BF}_{12})$ indicates that the data are more likely to occur under the first model than under the second model, whereas a negative $\log(\text{BF}_{12})$ indicates that the data

[6] Note that the Savage-Dickey density ratio test offers a method to compute Bayes factors, but only for *nested* models. We therefore had to construct nested models by fixing one of the model parameters. See Appendix E for a more detailed explanation and the results.

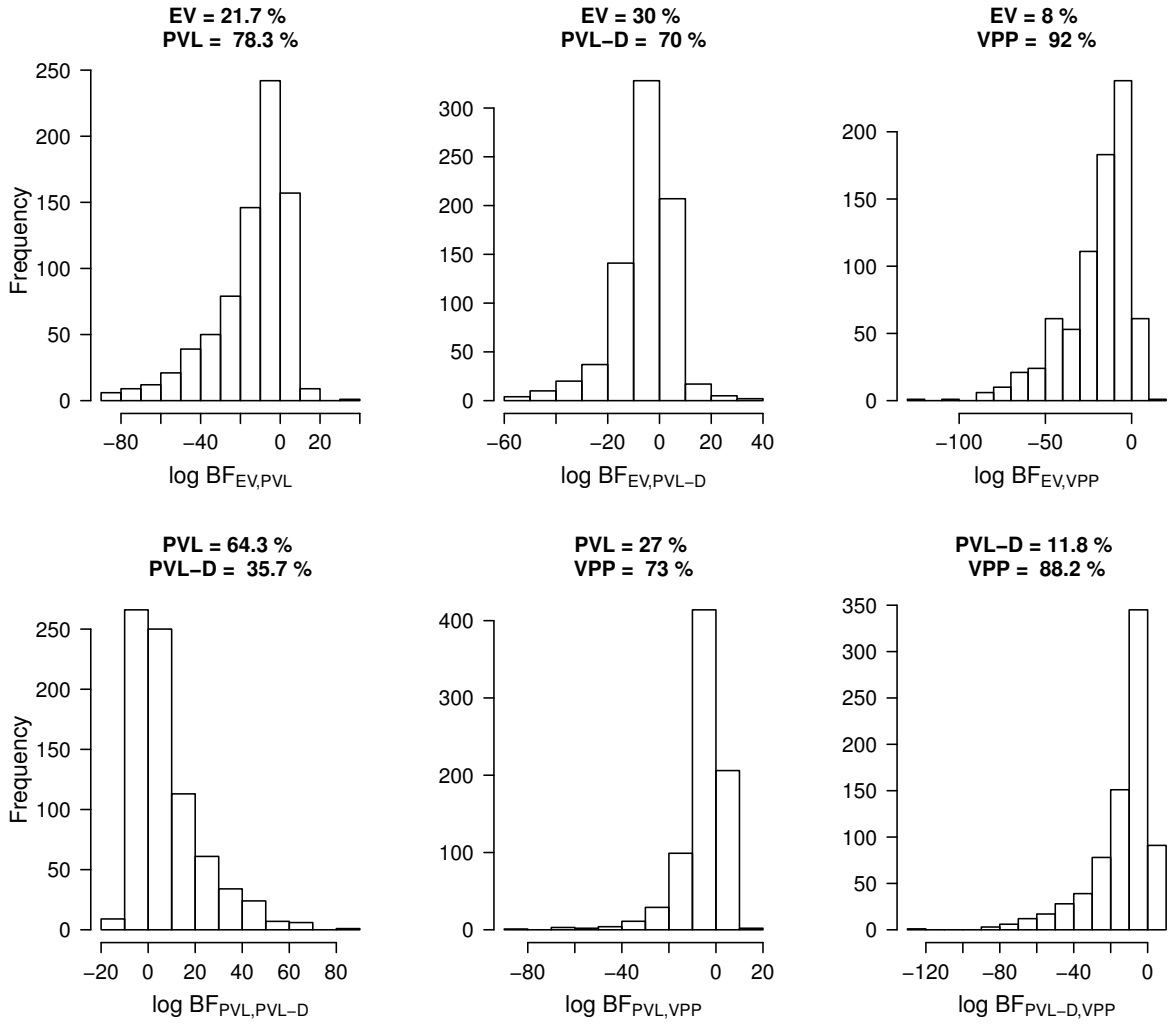

Figure 7.2: Histograms of the log(BF) for pairwise comparison of four RL models applied to the IGT data from each of 771 participants. A positive $\log(BF_{12})$ indicates that the data are more likely to occur under the first model than under the second model, whereas a negative $\log(BF_{12})$ indicates that the data are more likely to occur under the second model. Note that a log(BF) of 20 corresponds to a BF of almost 500 million, and that Jeffreys (1961) considers as extreme evidence a Bayes factor larger than 100 (i.e., $\log(BF) > 4.6$). The header of each histogram presents the percentage of participants for whom the data are more likely to occur under the corresponding model.

Table 7.5: *Median posterior model probabilities (MPMP; note that these need not sum to 1), and percentage of participants for whom the corresponding model has the largest posterior model probability. Grey shaded cells refer to the best model.*

	MPMP	%
EV	.00	7
PVL	.04	25
PVL-Delta	.00	9
VPP	.64	59

are more likely to occur under the second model than under the first model. The header of each histogram presents the percentage of participants for whom the data are more likely to occur under model $\mathcal{M}_1$ than model $\mathcal{M}_2$.

The top left panel of Figure 7.2, for example, shows that the data of 78.3% of the participants are more likely to occur under the PVL model than under the EV model. From this panel, it is also evident that the degree of evidence in favor of the PVL model is in general very strong, but also differs greatly across participants: 31% of the participants have a log Bayes factor in favor of the PVL model in between -10 and 0, 29% of the participants have a log Bayes factor in favor of the PVL model in between -30 and -10, and 18% of the participants have a log Bayes factor in favor of the PVL model in between -80 and -30. Note that, first, the smaller the log Bayes factor, the stronger the evidence in favor of the PVL model; second, a log(BF) of -20, for example, corresponds to a Bayes factor of almost 500 million in favor of the PVL model; finally, Jeffreys (1961) considers as extreme evidence a Bayes factor larger than 100 (i.e., $\log(\text{BF}) > 4.6$). Thus, more than 47% of the participants provide extreme evidence in favor of the PVL model. In the case of the participants that provide evidence in favor of the EV model (i.e., those participants that have a positive log Bayes factor), the differences across participants are much smaller (i.e., the largest Bayes factor is below 40), and most of these participants (i.e., 20%) have a log Bayes factor between 0 and 10, and none has a log Bayes factor in favor of the EV model larger than 40.

Taking all panels together, Figure 7.2 suggests that, among the four RL models, the data provide the most evidence for the VPP model, and the least evidence for the EV model. In addition, the data provide more evidence for the PVL model than for the PVL-Delta model. It is also evident that the data of many participants provide extreme evidence for a particular model.

The findings from Figure 7.2 are corroborated by Table 7.5. The second column of Table 7.5 shows the median posterior model probabilities, and the third column shows the percentage of participants for whom the corresponding model has the largest posterior model probability. It is evident that the VPP model is supported the most; that is, the data from 59% of the participants provide the most evidence for the VPP model. The PVL model is favored by the second largest proportion of the participants (i.e., 25%). Only a minority of participants provides the strongest evidence for the EV and PVL-Delta model (i.e., 7% and 9%, respectively) suggesting that the data provide weak evidence for the EV and PVL-Delta models. These findings are confirmed by the median posterior model probabilities presented in the second column of Table 7.5; in general, the VPP model is supported most, followed by the PVL model.

The distributions of individual posterior model probabilities are visualized in Figure 7.3, which presents violin plots of the 771 posterior model probabilities for each of the four RL models. The dots indicate the median posterior model probability (cf. second column of Table 7.5), and the boxes indicate the interquartile range (i.e., the distance between the .25 and .75 quantiles). From Figure 7.3, it is evident that in the case of the EV, PVL, and PVL-Delta models, the individual

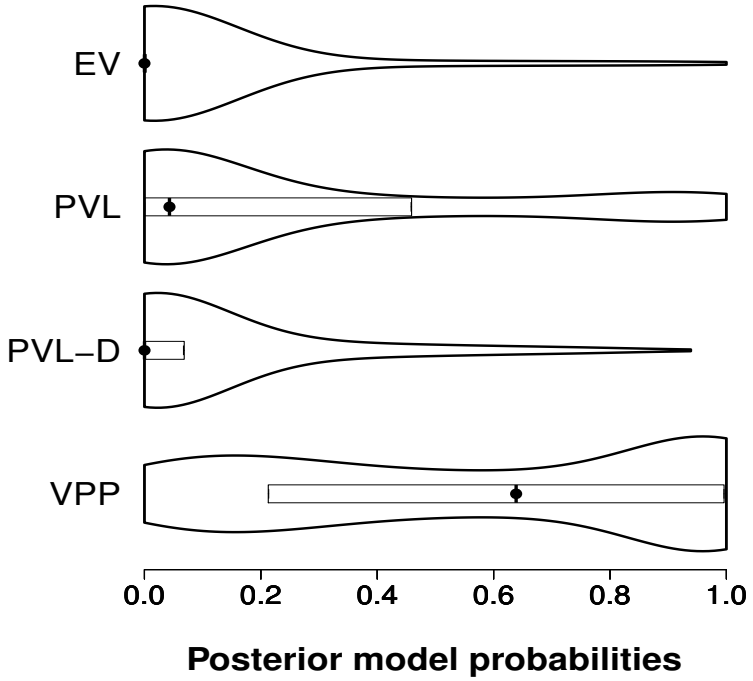

Figure 7.3: Distribution of the posterior model probabilities of 771 participants derived with importance sampling. Each violin plot shows the distribution of posterior model probabilities for one model. The dots indicate the median posterior model probability (cf. second column of Table 7.5), and the boxes indicate the interquartile range (i.e., the distance between the .25 and .75 quantiles).

posterior model probabilities follow a right skewed distribution suggesting that the data of most participants provide little evidence for these models, but that there are also a few participants who provide medium to strong evidence for either the EV, PVL, or PVL-Delta model. It is also evident that the tail of the distribution in the case of the EV and PVL-Delta models is thinner than in the case of the PVL model. This suggests that there are more participants who provide strong evidence for the PVL model then for the EV and PVL-Delta models. In the case of the VPP model, the distribution of the posterior model probabilities is bimodal with the right mode being more pronounced than the left mode. This distribution suggests that the evidence for the VPP model differs greatly across participants, but that most participants provide compelling evidence in favor of the VPP model.

7.6 Discussion

In this article, we illustrated how Bayes factor analyses can be performed to compare four RL models of the Iowa gambling task using a large data pool with 771 participants from 11 studies. The overall results provided strong evidence for the VPP model, moderate evidence for the PVL model, but weak evidence for the EV and PVL-Delta models.

The possibility to obtain Bayes factors for RL models is an important contribution to the

ongoing model comparison efforts in the RL literature. The most popular method to compare RL models for the IGT is the BIC post hoc fit criterion (i.e., Ahn et al., 2008; Busemeyer & Stout, 2002; Fridberg et al., 2010; Worthy, Hawthorne, & Otto, 2013; Yechiam & Busemeyer, 2005; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008). This method implements the tradeoff between goodness-of-fit and parsimony. However, the current implementation of the BIC post hoc fit criterion takes only one dimension of model complexity into account, that is, the number of free parameters. The Bayes factor, on the other hand, implicitly and automatically considers three dimensions of model complexity: (1) the number of free parameters, (2) the functional form of the model, and (3) the extension of the parameter space (e.g., Busemeyer et al., in press; I. J. Myung & Pitt, 1997). Thus, the Bayes factor offers the opportunity for a more comprehensive and sophisticated implementation of the BIC post hoc fit criterion. In fact, a model comparison study using the BIC post hoc fit criterion for the same models and data pool as used in this article (see Appendix E), shows that, for the models and data pool under consideration, the BIC post hoc fit criterion does not offer a good approximation of the Bayes factor; according to the BIC post hoc criterion the PVL model should be preferred over the VPP, PVL-Delta, and EV models. This analysis nicely illustrates the critique of the BIC that it prefers simple models that underfit the data (Burnham & Anderson, 2002). In this particular case, among the four models, the VPP model receives the strongest punishment for having relatively many parameters; however our Bayes factor analysis reveals that for this specific model comparison exercise, the number of free parameters alone is a limited and possibly misleading index of model complexity.

Despite the many advantages of the Bayes factor, it should be stressed that the Bayes factor only considers relative model adequacy. When the Bayes factor expresses an extreme preference for model $\mathcal{M}_1$ over model $\mathcal{M}_2$, this does not mean that model $\mathcal{M}_1$ fits the data at all well (Vandekerckhove et al., 2015). In addition, results from a model comparison based on Bayes factors depend on the specifics of the data set under consideration. Thus, in order to obtain a more complete account of the relative and absolute adequacy of the models under consideration, a Bayes factor analysis needs to be conducted in combination with other analyses featuring both relative and absolute indices of model adequacy (e.g., Ahn et al., 2008; Steingroever et al., 2014; Worthy, Pang, & Byrne, 2013; Yechiam & Busemeyer, 2005).

For example, crucial additional dimensions that a model comparison study should consider are parameter recovery, test of parameter consistency, test of generalizability, and test of specific influence (e.g., Ahn et al., 2008, 2011; Steingroever, Wetzels, & Wagenmakers, 2013b; Wetzels, Vandekerckhove, et al., 2010; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008). We briefly discuss these alternative methods in turn. First, among the most fundamental requirements of model adequacy is accurate parameter recovery. Accurate parameter recovery means that when fitting a model to a synthetic data set that was generated by that same model, the parameter estimates converge to the true data-generating values (e.g., Ahn et al., 2011, 2014; Steingroever, Wetzels, & Wagenmakers, 2013b; Wetzels, Vandekerckhove, et al., 2010). Poor parameter recovery suggests that a model is not identifiable because there are several parameter combinations that are equally likely to have generated the data. If the true data-generating parameters cannot be identified in an idealized scenario (i.e., where the fitted model has also generated the data) this suggests that parameter estimates obtained from fitting real data may not be reliable indicators of the underlying psychological processes. Thus, if researchers wish to draw meaningful conclusions about the psychological processes underlying IGT performance, they should avoid a model with poor parameter recovery even if this model outperforms its competitors on other model comparison tests.

Second, the test of parameter consistency compares the correlations between model parameters estimated in two different tasks that are intended to measure the same psychological processes

(i.e., Yechiam & Busemeyer, 2008). Good parameter consistency suggests that the psychological processes driving the performance on both tasks are invariant across tasks, and that the model captures these psychological processes. Thus, good parameter consistency confirms the validity of the model parameters. Third, the test of generalizability also requires data from two different, but related tasks. However, this method assesses a model's predictions for the second task based on parameters estimated in the first 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). Thus, good performance on the test of generalizability suggests that the model can be used to make accurate predictions about the behavior of a decision maker in other situations confirming the validity of the model parameters. Finally, the test of specific influence assesses whether the model parameters indeed correspond to distinct psychological processes (i.e., Steingroever, Wetzels, & Wagenmakers, 2013b; Wetzels, Vandekerckhove, et al., 2010). In particular, it assesses whether experimental manipulations that were intended to affect the model parameters are also reflected by the parameter estimates. If participants are, for example, distracted during the IGT by means of a filler task, this manipulation should be reflected by the parameter capturing the memory process involved in IGT performance. To conclude, a good performance on the above methods suggests that the estimated model parameters can be used to draw meaningful conclusions about the psychological processes underlying performance on the IGT.

Finally, it should be stressed that a Bayes factor analysis should be accompanied by a robustness analysis (e.g., Kass & Raftery, 1995; Matzke, Nieuwenhuis, et al., 2015; Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011). The aim of such an analysis is to investigate the extent to which the conclusions are altered by the choice of different priors for the model parameters. The Bayes factor is sensitive to the prior distributions because the prior is part of the marginal likelihood computation (Equation 7.3). In particular, the marginal likelihood is the likelihood of the data averaged across the entire parameter space, with the prior acting as averaging weights. In Appendix E, we present a robustness analysis which reveals that although different prior choices affect the exact value of the Bayes factors the qualitative conclusions are unaffected by the prior choice.

An interesting and challenging future direction is to derive Bayes factors for a hierarchical implementation of RL models. This allows a comparison of the models on the group-level instead of on the level of individual participants. Unfortunately, importance sampling for hierarchical models is not straightforward. The main reason for this is the increase in the number of estimated parameters. In the case of the individual-level implementation the dimension of the integral that we approximated with importance sampling equals the number of parameters (i.e., 3 in the case of the EV model up to 8 in the case of the VPP model). However, in the case of the hierarchical implementation, the dimension equals $(2 + n) \times k$, with n the number of participants and k the number of parameters.[7] For example, an experiment with 30 participants requires a hierarchical EV and VPP model with $32 \times 3 = 96$ and $32 \times 8 = 256$ parameters, respectively. Thus, a major challenge is to find a method that accurately approximates the marginal likelihood of the data of the entire group given a specific model. It is likely that future effort to approximate such high-dimensional integrals will involve more sophisticated sampling methods such as transdimensional Markov chain Monte Carlo (e.g., Green, 2003; Sisson, 2005). Finally, it should be kept in mind that one drawback of hierarchical models is that they do not account for the possibility that there can be different

[7] We need to add two to the number of participants to also incorporate the group-level parameters.

subgroups of participants—an issue that can be solved with mixture modeling (e.g., Huizenga et al., 2007; Konstantinidis et al., 2014).

Our finding that the majority of the 771 healthy participants provided the strongest evidence for the VPP model suggests that the perseveration process –a process that is included in the VPP model, but not in the three other RL models– plays an important role in risky decision-making. However, as mentioned above, our Bayes factor analysis should not be considered in isolation of other model comparison methods. There are several findings that have raised some skepticism about the VPP model. In particular, Ahn et al. (2014) showed that among the EV, PVL, PVL-Delta, and VPP models, the VPP model had the worst simulation and parameter recovery performance. In addition, Ahn et al. (2014) showed that the posterior distributions of several VPP parameters were very broad. This suggests that some parameters of the VPP model might be hardly interpretable and might have little psychological value, and that a more thorough analysis of the validity of the VPP model is required (for validations of the EV, PVL, and PVL-Delta model see Ahn et al., 2011; Steingroever, Wetzels, & Wagenmakers, 2013b; Wetzels, Vandekerckhove, et al., 2010).

To conclude, Bayes factor analyses can be performed to compare RL models on the level of individual participants. Also demonstrated is that importance sampling offers an efficient way to obtain individual-participant Bayes factors for RL models, and that, for the models and data under consideration, the BIC post hoc fit criterion does not offer a good approximation of the Bayes factor. Our data provide strong evidence for the VPP model, moderate evidence for the PVL model, but weak evidence for the EV and PVL-Delta models. Future efforts should be made to more thoroughly validate the VPP model and to derive Bayes factors for hierarchical models.

Acknowledgements

We thank Busemeyer and Stout (2002), Fridberg et al. (2010), Annette Horstmann, Kjome et al. (2010), Maia and McClelland (2004), Thorsten Pachur, Premkumar et al. (2008), Wood et al. (2005), and Worthy, Pang, and Byrne (2013) for providing the data used in this article. This publication was supported by an ERC grant from the European Research Council and by a Netherlands Organisation for Scientific Research (NWO) grant to HS (404-10-086); the Stan program was supported by the U.S. Department of Education for developing models for education research.