Safe models for risky decisions

Steingröver, H.M.

Link to publication

Creative Commons License (see https://creativecommons.org/use-remix/cc-licenses):
Other

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
In everyday life, we often have to decide between options that differ in their immediate and long-term consequences. Would you, for example, opt for a delicious piece of cake or rather eat a healthy apple? To investigate how people make risky decisions, this thesis focuses on the Iowa gambling task (IGT) and scrutinizes assumptions about the performance of healthy participants on the IGT. This thesis also challenges the trustworthiness of conclusions typically obtained from fitting reinforcement-learning models to IGT data. I argue that the risk of drawing premature conclusions from behavioral analyses and computational modeling can be minimized if researchers follow a number of crucial steps. These steps concern behavioral data analyses, model selection, model fitting, and assessment of absolute model fit. In particular, I advocate Bayesian techniques involving Bayesian repeated measures ANOVA for behavioral data analyses, the Bayes factor for model selection, the Bayesian hierarchical framework for model fitting, and posterior predictives to assess the absolute account of the models for the data at hand. Discussing a large variety of models and methods to compare the models, this dissertation illustrates that research efforts about risky decision making greatly advanced during the last years. On the other hand, this dissertation also illustrates the major challenges by pointing to problems with respect to behavioral analyses and cognitive modeling. Pursuing these suggestions will hopefully yield more reliable measures of risky decision making and a better understanding of the underlying psychological processes.

helensteingroever.com
Safe Models for Risky Decisions

Helen Maria Steingröver
SAFE MODELS FOR RISKY DECISIONS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op vrijdag 21 april 2017, te 10.00 uur

door Helen Maria Steingröver

geboren te Münster, Duitsland
Promotiecommissie:

<table>
<thead>
<tr>
<th>Rol</th>
<th>Naam</th>
<th>Universiteit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promotor</td>
<td>Prof. dr. E.-J. Wagenmakers</td>
<td>University of Amsterdam</td>
</tr>
<tr>
<td>Overige leden</td>
<td>Dr. R.P.P.P. Grasman</td>
<td>University of Amsterdam</td>
</tr>
<tr>
<td></td>
<td>Prof. dr. H.M. Huizenga</td>
<td>University of Amsterdam</td>
</tr>
<tr>
<td></td>
<td>Prof. dr. M.D. Lee</td>
<td>University of California, Irvine</td>
</tr>
<tr>
<td></td>
<td>Prof. dr. T.J. Pleskac</td>
<td>Max Planck Institute for Human Development, Berlin</td>
</tr>
<tr>
<td></td>
<td>Prof. dr. H.L.J. van der Maas</td>
<td>University of Amsterdam</td>
</tr>
<tr>
<td></td>
<td>Dr. I. Visser</td>
<td>University of Amsterdam</td>
</tr>
</tbody>
</table>

Faculteit: Faculteit der Maatschappij– en Gedragswetenschappen
Für meine Eltern
# Contents

1 Introduction .......................... 1
   1.1 Chapter Outline .......................... 6

2 Performance of Healthy Participants on the Iowa Gambling Task .......................... 9
   2.1 The Iowa Gambling Task .......................... 11
   2.2 Literature Reviews and Data Analyses .......................... 15
   2.3 Discussion .......................... 26

3 A Comparison of Reinforcement-Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning .......................... 31
   3.1 The IGT and Three Reinforcement-Learning Models .......................... 33
   3.2 Comparison of the EV, PVL, and EV-PU Models .......................... 38
   3.3 Discussion .......................... 50

4 Validating the PVL-Delta Model for the Iowa Gambling Task .......................... 55
   4.1 The Iowa Gambling Task and the PVL-Delta Model .......................... 56
   4.2 Parameter Space Partitioning .......................... 59
   4.3 Test of Selective Influence .......................... 69
   4.4 Discussion .......................... 76

5 Absolute Performance of Reinforcement-Learning Models for the Iowa Gambling Task .......................... 79
   5.1 The Iowa Gambling Task and Three Reinforcement-Learning Models .......................... 81
   5.2 Performance of the EV, PVL, and PVL-Delta Models .......................... 87
   5.3 General Discussion .......................... 99

6 $\hat{w} = .2, \hat{a} = .8, \hat{c} = .6$: So What? On the Meaning of Parameter Estimates from Reinforcement-Learning Models .......................... 103
   6.1 Clarification of Our Initial Goal .......................... 104
   6.2 Post Hoc Absolute Fit Method Versus Simulation Method .......................... 108
   6.3 Statistical Adequacy Versus Psychological Relevance .......................... 110
   6.4 Conclusion .......................... 110

7 Bayes Factors for Reinforcement-Learning Models of the Iowa Gambling Task .......................... 111
   7.1 The Iowa Gambling Task .......................... 113
   7.2 The EV, PVL, PVL-Delta, and VPP Models .......................... 114
# Contents

7.3 The Bayes Factor ................................................................. 118
7.4 Obtaining Bayes Factors for RL Models Using Importance Sampling 119
7.5 Application to IGT Data from 771 Healthy Participants ................. 120
7.6 Discussion ........................................................................... 125

8 A Tutorial on Bridge Sampling .................................................. 129
8.1 Four Sampling Methods to Approximate the Marginal Likelihood 132
8.2 Case Study: Bridge Sampling for Reinforcement Learning Models 149
8.3 Discussion ........................................................................... 157

9 Bayesian Techniques for Analyzing Group Differences in the Iowa Gambling Task: A Case Study of Intuitive and Deliberate Decision Makers ..................................................... 159
9.1 The IGT and PVL-Delta Model ................................................ 161
9.2 Proposed Methodology for Comparing Groups on the IGT .......... 164
9.3 Case Study: Intuitive versus Deliberate Decision Making .......... 166
9.4 Discussion ........................................................................... 172

10 Using Bayesian Regression to Incorporate Covariates into Hierarchical Cognitive Models .................................................. 175
10.1 Regression Framework for Relating Cognitive Model Parameters to Covariates ..................................................... 179
10.2 Simulation Study ................................................................... 182
10.3 Discussion ........................................................................... 190

11 Summary and Future Directions .............................................. 193
11.1 Summary of Results ............................................................. 193
11.2 Future Directions ................................................................ 197
11.3 Concluding Remarks ........................................................... 200

Appendices................................................................................. 203

A Performance and Awareness in the Iowa Gambling Task .......... 205

B Data from 617 Healthy Participants Performing the Iowa Gambling Task: A “Many Labs” Collaboration ..................................................... 209
B.1 Overview ............................................................................. 210
B.2 Methods .............................................................................. 210
B.3 Dataset Description ............................................................. 211
B.4 Reuse Potential ................................................................. 214
B.5 Supporting Text 1 ............................................................... 215

C Appendix to Chapter 4: “Validating the PVL-Delta Model for the Iowa Gambling Task” ..................................................... 219

D Appendix to Chapter 5: “Absolute Performance of Reinforcement-Learning Models for the Iowa Gambling Task” ..................................................... 225
D.1 Recipe for Obtaining Choice Probabilities According to the Post Hoc Absolute Fit Method ..................................................... 225
D.2 Recipe for Obtaining Choice Probabilities According to the Simulation Method ..................................................... 226
In everyday life, we often have to decide between options that differ in their immediate and long-term consequences. You may be tempted to eat a delicious piece of cake, but on the other hand, you might be afraid of gaining weight and getting diabetes, and consider eating an apple instead. Or imagine you are on a motor trip. You may choose to drive fast to experience excitement and adrenaline, but on the other hand, the danger of risking a heavy accident might convince you that it is more reasonable to drive slowly. Or yet another example, imagine you have to choose between one of the following two job opportunities: The first one comes with little challenges and opportunities for development, but is sufficiently paid and for an unlimited period; the alternative concerns a challenging job opportunity with new tasks and responsibilities, but that is limited to half a year. Which job opportunity would you prefer?

Risky decisions with unclear outcomes are the focus of much psychological research. To investigate how risky decisions are made in a controlled, experimental context, a large variety of tasks has been proposed. Popular examples are multi-armed bandit tasks or gambling tasks where participants are repeatedly asked to choose between at least two options (e.g., gambles) in such a way that they optimize their long-term outcomes. In this dissertation, I focus on a particular gambling task proposed in 1994 by [Bechara, Damasio, Damasio, and Anderson], the Iowa gambling task (IGT; for alternative tasks, see the Balloon Analogue Risk Task, [Lejuez et al., 2002]; Wallsten, Pleskac, & Lejuez, 2005; the Columbia card task, Figner, Mackinlay, Wilkening, & Weber, 2009; and the Soochow gambling task, Lin, Chiu, & Huang, 2009).

In the IGT, participants are presented with four decks of cards; each card yields a reward and, occasionally, also a loss (see Figure 1.1 for an example screen shot of the IGT). Participants are told to repeatedly choose cards from the decks to optimize their long-term outcomes. Unbeknownst to the participants, two decks contain risky cards that have high immediate rewards on every trial but negative long-term outcomes because of occasionally very high losses, whereas the other two decks contain safe cards that have small immediate rewards on every trial and occasionally low losses, therefore resulting in positive long-term outcomes. For an overview of the payoff scheme, Table 1.1 presents the wins and losses associated with 10 example choices from each deck.

Participants can succeed on the IGT only when they learn to forego high immediate rewards and prefer the safe options (i.e., decks C and D) over the risky options. [Bechara et al., 1994] claim that this choice behavior is characteristic for healthy participants and that they base their choices on the long-term outcomes of the decks. Such characteristic choice behavior of two typical healthy participants of [Bechara et al., 1994] is presented in Figure 1.2. It is evident that the two participants “initially sampled all decks ( . . . ) but eventually switched to more and more selections.
1. Introduction

Figure 1.1: Example screen shot of the IGT showing that a choice from the leftmost deck resulted in a reward of $100 and a loss of $0.

from the good decks C and D, with only occasional returns to decks A and B” (Bechara et al., 1994 p. 12).

In contrast to the successful IGT performance of healthy participants, Bechara et al. (1994) claim that patients with lesions to the ventromedial prefrontal cortex (vmPFC) – a part of the brain that is located at the frontal lobe at the bottom of the cerebral hemispheres – perform deficiently on the IGT. These vmPFC patients often make decisions that are irresponsible, risky, and go against their own interests. When confronted with a decision problem where the gains of immediate reward need to be weighted against the risks of long-term loss, vmPFC patients tend to focus on the immediate reward and disregard possible negative future outcomes of their choices, such as separation from family and friends, and loss of reputation and job (Bechara & Damasio, 2002; Dunn, Dalgleish, & Lawrence, 2006). The most famous vmPFC patient, Phineas Gage, survived an accident in which an iron rod lesioned his frontal cortex, causing a profound change in

Table 1.1: Wins and losses associated with 10 example choices from each deck of the IGT. Decks A and B are labeled as bad decks because they result in negative long-term outcomes, whereas the good decks (i.e., decks C and D) result in positive long-term outcomes. The last row shows, for each deck, the net outcome after 10 choices.

<table>
<thead>
<tr>
<th>Choice</th>
<th>Bad decks</th>
<th>Good decks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deck A</td>
<td>Deck B</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>100, –150</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>100, –350</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>100, –300</td>
<td>100, –1250</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>100, –250</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>100, –200</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Net outcome</td>
<td>–250</td>
<td>–250</td>
</tr>
</tbody>
</table>
Figure 1.2: Deck selection profiles of two typical participants in Bechara et al. (1994)’s control group. These two participants first explore all decks, then gradually switch to the good decks, and only occasionally return to the bad decks.

Figure 1.3: Gage’s skull with an iron rod lesioning his ventromedial prefrontal cortex (downloaded October 2, 2016, from https://neurophilosophy.wordpress.com/2006/12/04/the-incredible-case-of-phineas-gage/).
his decision-making abilities and social behavior (Figure 1.3). Prior to the accident, Gage was a hard-working, efficient, and responsible railroad construction foreman. However, after his accident he could not resume work because his behavior was impulsive, impatient, and antisocial. These symptoms are characteristic for vmPFC patients in general (Boes et al., 2011).

Due to the symptoms of patients with lesions to the vmPFC—in particular, their insensitivity to future consequences of their choices—they are assumed to prefer the risky options on the IGT (i.e., decks A and B). According to Bechara and his colleagues, the result of this “myopia for the future” is that the choice behavior of vmPFC patients is controlled primarily by immediate prospects, positive or negative (Bechara et al., 1994; Bechara, Tranel, & Damasio, 2000): “these subjects are unresponsive to future consequences, whatever they are, and are thus more controlled by immediate prospects” (Bechara et al., 1994, p. 14).

IGT performance is typically recorded for a clinical group and a control group. Subsequently, performance of the groups is compared using behavioral data analyses usually focusing on the overall proportion of choices from the good decks. Such analyses can tell us which choices participants made (e.g., that vmPFC patients made more risky choices than healthy participants), and how their preferences change as the task proceeds or as a reaction to experienced losses. However, behavioral data analyses cannot tell us whether, for example, the performance deficit of vmPFC patients was caused by a focus on immediate gains, or poor memory, making it impossible for them to remember previously experienced payoffs associated with the different decks. In this dissertation, I argue that to overcome this shortcoming of behavioral data analyses, cognitive models should be used. These models make assumptions about the relevant psychological processes such as motivation (e.g., perceiving rewards as important as losses), memory (e.g., the ability to correctly remember past outcomes of the decks), and response consistency (e.g., the tendency to base decisions on one’s expectations about the decks), and how these processes interact to produce overt choice behavior.

Popular cognitive models for the IGT are so-called reinforcement learning (RL) models (Sutton & Barto, 1998). These models explain how autonomous agents learn to shape their decisions through trial and error, and all the experiences they had with each option up to the point of the decision. RL models are especially suitable for IGT data because of the repeated nature of the IGT that raises the question as to how much people explore the different options before they exploit the most profitable ones, and what drives this exploration-exploitation tradeoff. As will become apparent in this dissertation, there exist a large variety of RL models that have been applied to IGT data from various clinical groups. A prominent example is the study of Yechiam, Busmeyer, Stout, and Bechara (2005) using the Expectancy Valence model (EV; Busmeyer & Stout, 2002)—the first RL model that has been proposed for the IGT. This study presents the parameters of the EV model for 10 different clinical groups relative to the parameters of their respective control group of healthy participants (Figure 1.4). The purpose of this analysis was to understand differences between the 10 clinical groups in terms of the model parameters that reflect the psychological processes involved in risky decision making. According to Yechiam et al. (2005) Figure 1.4 suggests that there are two clusters of clinical populations with EV model parameters that clearly differ from the ones of their respective control group. The first cluster is located in the top right part of the figure. These groups are characterized by a high attention to gains and high attention to recent outcomes. The second cluster located in the left part of the figure contains clinical populations with a high attention to losses. The authors conclude that “these findings demonstrate the potential contribution of cognitive models in building bridges between neuroscience and behavior” (p. 973).

When relying on these models for clinical purposes, it is important that conclusions from the model parameters are meaningful. In this dissertation, I elaborate on aspects that have to be
taken into account to avoid premature conclusions. These aspects concern three areas: (1) model selection; (2) model fitting; and (3) model account for the data. I argue that all three areas, in particular the first two ones, can greatly profit from the use of Bayesian statistics instead of classical statistics (also known as frequentist statistics). Bayesian statistics is the practical and principled way to reason with uncertainty, and has become increasingly popular in mathematical psychology (Andrews & Baguley, 2013; Poirier, 2006).

In order to compare models (e.g., to find a suitable model for IGT data or to investigate whether two groups differ in specific model parameters), I advocate the gold standard of model selection in Bayesian statistics—the Bayes factor—and I illustrate several methods to obtain the Bayes factor for RL models. A crucial advantage of the Bayes factor is that it can be used not only to identify the model that is best supported by the data, but also to quantify the relative evidence that the data provide for two competing models.

In order to fit RL models to IGT data, I argue that the state-of-the-art method, that is, Bayesian hierarchical modeling, should be used. Bayesian hierarchical modeling incorporates both differences and commonalities between and within the participants of one group, and thereby yields more informative and accurate parameter inferences (Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011; Horn, Pachur, & Mata, 2015; Lejarraga, Pachur, Frey, & Hertwig, 2016; Navarro, Griffiths, Steyvers, & Lee, 2006; Rouder & Liu, 2005; Rouder, Lu, Speckman, Sun, & Jiang, 2005; Rouder, Lu, Morey, Sun, & Speckman, 2008; Scheibehenne & Pachur, 2015; Shiffrin, Lee, Kim, & Wagenmakers, 2008; Wetzels, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2010).

Finally, in order to assess the account of a model for the data, I argue that posterior predictives should be used. A major advantage of posterior predictives is that they also take the uncertainty of the parameter estimates into account (i.e., they make use of the entire posterior distribution) whereas frequentist methods commonly only use a point estimate (i.e., the maximum likelihood

Figure 1.4: Difference in the EV model parameters of 10 different clinical groups and their respective control group of healthy participants. Figure taken from Yechiam et al. (2005).
1. Introduction

In the remainder of the introduction, I will give an overview and a brief description of the problems to be addressed in the coming chapters.

1.1 Chapter Outline

Chapter 2 gives an overview of the performance of healthy participants on the IGT, and tests three key assumptions about the IGT performance of healthy participants: (1) healthy participants learn to prefer the good options over the bad options; (2) healthy participants show homogeneous choice behavior; and (3) healthy participants first explore the different options and then exploit the most profitable ones (see Figure 1.2). For this goal we use two extensive literature reviews and analysis of eight data sets. Our findings show that all three assumptions may be invalid and therefore question the prevailing interpretation of IGT data.

In Chapter 3 we use parameter space partitioning (PSP) to compare three popular RL models—the EV (Busemeyer & Stout, 2002) and Prospect Valence Learning (PVL; Ahn, Busemeyer, Wagenmakers, & Stout, 2008) model, and a combination of these models, the EV-PU model. The PSP method investigates which choice patterns the models generate across their entire parameter space to yield an indication of the data-fitting potential of the models. Ideally, choice patterns that are often observed in experiments should also occupy a large part of the parameter space of the model under consideration because in such a situation, the model is most likely to provide a good account for the data. PSP thus focuses on the flexibility of a model. In addition, PSP is a global model comparison technique because it is independent of a specific data set. Our results suggest that there is a large discrepancy between the choice patterns that are central to the models and the choice patterns that are often observed in experiments. Overall, our results suggest that the search of an appropriate IGT model has not yet come to an end.

Chapter 4 illustrates three important methods for model validation on the example of the PVL-Delta model—yet another combination of the EV and PVL models. The validation methods cover parameter recovery, parameter space partitioning, and test of specific influence. To assess parameter recovery we fit a model in an idealized scenario (i.e., the fitted model has also generated the data; data generation is based on representative parameter values and a representative number of trials and participants), and investigate whether the parameter estimates coincide with the data-generating parameter values. If the data-generating values cannot even be identified in an idealized scenario, this suggests that parameter estimates obtained from fitting real data may not be reliable indicators of the underlying psychological processes. The test of specific influence, on the other hand, assesses whether experimental manipulations that were intended to affect specific 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. Our results suggest that, despite a few shortcomings, the PVL-Delta model seems to be a better IGT model than the models discussed in Chapter 3.

In Chapter 5 we claim that parameters obtained from fitting an RL model to IGT data are often interpreted without sufficient assessment of the model’s account for the data in absolute terms. In order to avoid premature conclusions, we propose a minimum threshold of adequacy for two tests: the post hoc absolute fit test and the simulation method test. The post hoc absolute fit test assesses a model’s ability to fit an observed choice pattern when provided with information on the observed choices and payoffs. The simulation method test, on the other hand, assesses a model’s ability to generate the observed choice pattern with parameter values obtained from
model fitting. The crucial difference between the two tests is that the first test is guided by information on the observed choices and payoffs, whereas the second test makes predictions using new, unobserved payoff sequences. These tests are illustrated using two stylized data sets and five published data sets. Our results highlight that a model’s ability to fit a particular choice pattern does not guarantee that the model can also generate that same choice pattern. Future applications of RL models should carefully assess absolute model performance to avoid premature conclusions about the psychological processes that drive performance on the IGT.

The next chapter, Chapter 6, is a rejoinder to Konstantinidis, Speekenbrink, Stout, Ahn, and Shanks (2014)’s reply on Chapter 5. In this chapter, we clarify our initial goal, that is, to illustrate why assessment of absolute model performance is necessary to avoid premature conclusions about the psychological processes that drive performance on the IGT. In addition, we elaborate on the advantages and drawbacks of both the post hoc absolute fit method and the simulation method. Finally, we highlight the distinction between statistical aspects of model adequacy and psychological relevance of parameter estimates.

Chapter 7 argues that, by only considering the number of free parameters, the BIC post hoc fit criterion—the most popular method to compare RL models—does not correctly discount model complexity. A coherent and complete discounting of complexity is provided by the Bayes factor. We use Bayes factors that are obtained with a Monte Carlo method, known as importance sampling, in order to compare four RL models of the IGT. (i.e., the EV, PVL, PVL-Delta, and Value-Plus-Perseveration, VPP, models). The method is illustrated using a data pool of 771 healthy 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.

The next chapter, Chapter 8, focuses on bridge sampling, a Monte Carlo method for obtaining Bayes factors more general than importance sampling. We show that bridge sampling can be applied to a Bayesian hierarchical implementation of RL models in a straightforward manner. Such easy extension to hierarchical models presents a crucial advantage compared to importance sampling introduced in Chapter 7. For educational purposes, we first explain the bridge sampling method using a simple Beta-Binomial model. In the second part of Chapter 8, we apply bridge sampling to both an individual-level and hierarchical implementation of the EV model, and compare our results of the individual-level implementation to the ones reported in Chapter 7. Our results suggest that bridge sampling is an attractive method for mathematical psychologists who typically aim to approximate the marginal likelihood for a limited set of (complex, high-dimensional) models.

Chapter 9 proposes a suite of three complementary model-based methods for assessing the cognitive variables and processes underlying IGT performance: (1) Bayesian hierarchical parameter estimation; (2) Bayes factor model comparison; and (3) Bayesian latent-mixture modeling. To illustrate these Bayesian analysis techniques, we test the extent to which differences in decision style (i.e., intuitive, affective vs. deliberate, planned) explain differences in IGT performance. The Bayes factor model comparison is related to the one presented in Chapter 8 but here we do not compare different types of RL models; instead we focus on one specific RL model—the PVL-Delta model—and our model comparison concerns the question whether intuitive and deliberate decision makers differ on the parameters of this model. To quantify the evidence that the data provide for this hypothesis we use Bayes factors obtained by yet another method—the product space method. Our results challenge the notion that individual differences in intuitive and deliberative decision styles have a very broad impact on decision making, and that intuitive processes in healthy adults play a central role for IGT performance.

The final chapter of this thesis, Chapter 10, reuses the data from Chapter 9 and proposes a more elegant solution on how to compare model parameters across two groups avoiding a multi-step procedure. In particular, in Chapter 9 we first submitted participants’ responses to a decision style
questionnaire to a principal component analysis, and then classified participants as intuitive or deliberate depending on their factor scores; subsequently, we compared the model parameters of the two groups. In Chapter 10 we explain that this procedure can be erroneous. As a solution, we propose a Bayesian regression framework that can be used to extend existing RL models, but in general also other cognitive models. This framework allows researchers to quantify the evidential support for relationships between covariates (e.g., decision style) and model parameters using Bayes factors.
Chapter 2

Performance of Healthy Participants on the Iowa Gambling Task

This chapter has been published as:


Performance of healthy participants on the Iowa gambling task.


Abstract

The Iowa gambling task (IGT; Bechara et al., 1994) is often used to assess decision-making deficits in clinical populations. The interpretation of the results hinges on three key assumptions: (1) healthy participants learn to prefer the good options over the bad options; (2) healthy participants show homogeneous choice behavior; and (3) healthy participants first explore the different options and then exploit the most profitable ones. Here we test these assumptions using two extensive literature reviews and analysis of eight data sets. The results show that all three assumptions may be invalid, that is (1) healthy participants often prefer decks with infrequent losses; (2) healthy participants show idiosyncratic choice behavior; and (3) healthy participants do not show a systematic decrease in the number of switches across trials. Our findings question the prevailing interpretation of IGT data and suggest that, in future applications of the IGT, key assumptions about performance of healthy participants warrant close scrutiny.

Patients with lesions to the ventromedial prefrontal cortex (vmPFC) suffer from severe decision-making deficits; even though their cognitive, memory, and problem-solving abilities remain intact, they are often unable to consider the consequences of their actions and to learn from their mistakes (Bechara et al., 1994; Bechara, Damasio, & Lee, 1999; Bechara et al., 2000; Yechiam et al., 2005). The most famous vmPFC patient, Phineas Gage, survived an accident in which an iron rod lesioned his frontal cortex, causing a profound change in his decision-making abilities and social behavior. Prior to the accident, Gage was a hard-working, efficient, and

1The final publication is available at http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=2012-25099-001
2. Performance of Healthy Participants on the Iowa Gambling Task

In order to study the decision-making deficits of vmPFC patients, Bechara et al. (1994) developed the Iowa gambling task (IGT). The purpose of this task is to mimic real-life decision-making in an experimental context. In the IGT, participants are presented with four decks of cards; each card yields a reward and, occasionally, also a loss. Participants are told to repeatedly choose cards to optimize their long-term outcomes. Unbeknownst to the participants, two decks contain risky cards that have high immediate rewards but negative long-term outcomes, whereas the other two decks contain safe cards that have small immediate rewards but positive long-term outcomes. Participants can succeed on the IGT only when they learn to forego high immediate rewards and prefer the safe options over the risky options. This choice behavior is assumed to be characteristic for healthy participants and is taken as evidence that they base their choices on the long-term outcomes of the decks (Bechara et al., 1994). In contrast, vmPFC patients presumably prefer the risky options because they are insensitive to the future consequences of their choices. The result of this “myopia for the future” is that the choice behavior of vmPFC patients is controlled primarily by immediate prospects, positive or negative (Bechara et al., 1994, 2000).

Over the last two decades, the IGT has become what is arguably the most popular neuropsychological paradigm to measure decision-making deficits in clinical populations (Toplak, Sorge, Benoit, West, & Stanovich, 2010). To illustrate, a search for “Iowa gambling task” yields about 178,000 hits on Google, and about 3,640 hits on Google Scholar. Prior to 2005, the IGT had already been used in almost 100 neurological and psychiatric studies (Bowman, Evans, & Turnbull, 2005). Studies that use the IGT have involved a broad variety of clinical populations such as patients with vmPFC lesions (Bechara, Damasio, Tranel, & Anderson, 1998; Bechara et al., 1999, 2000), pathological gambling (Cavedini, Riboldi, Keller, D’Annucci, & Bellodi, 2002), obsessive-compulsive disorder (Cavedini, Riboldi, D’Annucci, et al., 2002), psychopathic tendencies (Blair, Colledge, & Mitchell, 2001), schizophrenia (Bark, Dieckmann, Bogerts, & Northoff, 2005; Martino, Bucay, Butman, & Allegrig, 2007), cocaine users (Stout, Busemeyer, Lin, Grant, & Bonson, 2004), traffic offenders (Lev, Hershkovitz, & Yechiam, 2008), and inmates (Yechiam, Kanz, et al., 2008). The traditional way to compare performance of a clinical group to that of a control group is based either on the overall proportion of choices from the good decks, or on a difference score between the overall proportion of choices from the good and bad decks. Note that these procedures collapse choice proportions over the two good decks and over the two bad decks, leading to a loss of potentially diagnostic information (Chiu & Lin, 2007; Dunn et al., 2006; Lin, Chiu, Lee, & Hsieh, 2007), a point to which we will return later.

In addition to the broad application of the IGT in neuropsychological studies, Bechara (2007) promote to use the IGT in clinical assessment of patients with decision-making deficits. Bechara therefore developed IGT software including a professional manual. The software yields different normative scores, such as the total net gain and the total number of cards chosen from each deck. These normative scores can be compared to those of a demographically corrected sample or U.S.
Table 2.1: Payoff scheme of the traditional IGT as developed by Bechara et al. (1994).

<table>
<thead>
<tr>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad deck with frequent losses</td>
<td>Bad deck with infrequent losses</td>
<td>Good deck with frequent losses</td>
<td>Good deck with infrequent losses</td>
</tr>
<tr>
<td>Reward/trial</td>
<td>100</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Number of losses/10 cards</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Loss/10 cards</td>
<td>−1250</td>
<td>−1250</td>
<td>−250</td>
</tr>
<tr>
<td>Net outcome/10 cards</td>
<td>−250</td>
<td>−250</td>
<td>250</td>
</tr>
</tbody>
</table>

census-matched sample. According to Bechara, the IGT is “ideal to assess patients who exhibit poor decision-making skills in the presence of otherwise normal or unaffected intelligence because of head injury or insult or any other condition thought to impact the function of the prefrontal cortex.”

The validity of IGT scores as a diagnostic instrument hinges on at least three assumptions about the performance of healthy participants on the IGT. First, there is the explicit assumption that healthy participants learn to prefer the good options over the bad options (Bechara et al., 1994); Second, there is the implicit assumption that choice performance for healthy participants is homogeneous—a clinical test in which healthy participants have widely diverging scores is of limited use; Third, there is the explicit assumption that healthy participants first explore the different options, and then settle down and exploit the most profitable ones (i.e., the exploration-exploitation tradeoff; Bechara et al., 1994).

The goal of this article is to assess the validity of the above assumptions using two extensive reviews of the literature and detailed analysis of eight experimental data sets. To anticipate the conclusion, our results call into question each of the above assumptions: First, the choices of healthy participants are primarily driven by the frequency of losses rather than the long-term outcomes of the decks (i.e., the frequency-of-losses effect); Second, the performance of healthy participants is characterized by a considerable variability, both across groups and across participants within groups; Third, healthy participants fail to show a pronounced exploration-exploitation tradeoff.

The outline of this article is as follows. The first section explains the IGT, outlines the assumptions about the IGT performance of healthy participants, and reviews current criticism on the IGT. The second section tests each of the three key assumptions above by reviewing findings from 39 groups of healthy participants and by analyzing data from eight IGT studies. Section 3 summarizes our findings and discusses their ramifications.

2.1 The Iowa Gambling Task

Description of the Iowa Gambling Task

The purpose of the IGT is to measure decision-making deficits of clinical populations in an experimental setting. The IGT mimics real-life decision-making in the sense that participants are required to integrate rewards and losses, weight benefits and risks associated with each possible option, remember that information and use it in subsequent decisions. In the traditional IGT, participants are initially given $2000 facsimile money and are presented with four decks of cards.

2. Performance of Healthy Participants on the Iowa Gambling Task

Figure 2.1: Example screen shot of the IGT.

Participants are told to successively choose cards from the four decks to maximize their long-term outcome (Bechara et al. 1994; Bechara, Damasio, Tranel, & Damasio 1997).

In the IGT, each card is associated with a specific amount of monetary reward and potentially also of monetary loss. The difference between the four decks lies in the payoff scheme: Two decks (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 A and B are therefore called bad or disadvantageous decks. The other two decks (decks C and D) are associated with lower immediate, constant rewards, but with even lower unpredictable, occasional losses and thus result in positive long-term outcomes. Decks C and D are therefore called good or advantageous decks. 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). The payoff scheme of the traditional IGT as developed by Bechara et al. (1994) is presented in Table 2.1. Each time participants choose a card, they receive feedback on the reward and the loss (if any) associated with that card, and the running tally. Figure 2.1 shows an example screen shot of the IGT. Participants are not told how long the task takes, but typically it contains 100 trials. The task aims to determine whether participants learn through trial and error to prefer the good, safe decks (i.e., decks C and D) over the bad, risky decks (i.e., decks A and B).

IGT performance is typically recorded for a clinical group and a control group. Different outcome measures have been reported, but most studies compute either the overall proportion of choices from the good decks or a difference score between the overall proportion of choices from the good and bad decks. An overall proportion of choices from the good decks larger than .50 or a positive difference score are considered as non-impaired performance (Bechara et al. 1998, 1999; Bowman & Turnbull 2003). The disadvantage of these outcome measures is that they do not reflect the change in the deck preferences across trials. Therefore, many studies report these outcome measures also in blocks of 10 to 20 trials. These trial-dependent outcome measures are often entered in a block x group analysis of variance (ANOVA). Note that these procedures collapse choice proportions over the two good decks and over the two bad decks (Chiu & Lin 2007; Dunn et al. 2006; Lin et al. 2007).
Assumptions Underlying Performance on the Iowa Gambling Task

Bechara et al. (1994) developed the IGT to measure decision-making deficits of patients with vmPFC lesions. According to Bechara et al. (1994, 2000), these brain-damaged patients perform poorly on the IGT because they are insensitive to the future consequences of their actions (“myopia for the future”). Performance of vmPFC patients is therefore assumed to be guided by immediate prospects, positive or negative, rather than the accumulation of long-term positive outcomes (Bechara et al., 1994, 2000). Such myopic choice behavior results in a preference for the bad, risky decks, that is, those decks that yield constant high immediate rewards but even higher unpredictable losses (i.e., decks A and B); hence, this choice behavior leads to long-term negative outcomes. Healthy participants, on the other hand, are assumed to prefer the good, safe decks (i.e., decks C and D) because they correctly infer that these decks yield positive long-term outcomes, despite rather small immediate rewards (Bechara et al., 1994). Figure 2.2 displays the mean proportion of choices from each deck in Bechara et al. (1994)’s control group of healthy participants (N = 44). This figure illustrates that the control group showed a strong overall preference for the good decks, and thus clearly supports the assumption of Bechara et al. (1994) that healthy participants prefer the good decks due to their positive long-term outcomes. It is important to note that this group of healthy participants chose both good decks and both bad decks about equally often, indicating that these participants do not prefer the decks with infrequent losses (decks B and D) over the decks with frequent losses (decks A and C).

In order to support their assumptions about the characteristic choice behavior of healthy participants across trials, Bechara et al. (1994) presented deck selection profiles of two typical control participants (cf. Figure 2.3). These profiles illustrate that these control participants “initially sampled all decks (. . .), but eventually switched to more and more selections from the good decks C and D, with only occasional returns to decks A and B” (Bechara et al., 1994, p. 12). In the case of Bechara et al. (1994)’s two typical control participants, the last 50 trials featured only five and six choices from the bad decks (cf. Figure 2.3).
The difference in the performance of healthy participants and vmPFC patients has been frequently interpreted as support of the somatic marker hypothesis. This hypothesis states that emotion-related signals generated from the body, so-called somatic markers, regulate decision-making. Especially in complex, uncertain situations, somatic markers are assumed to mark the possible options with an emotional signal related to the goodness or badness of the associated outcome. Based on these emotional signals, healthy participants anticipate certain bodily states associated with each response option (Bechara et al., 1999; Bechara & Damasio, 2002; A. R. Damasio, Everitt, & Bishop, 1996). These expectations are assumed to guide decisions by restricting the decision-making space to options associated with positive affective states. Crucially, vmPFC patients are found to be unable to rely on somatic markers when facing multiple response options (Bechara et al., 1999; Bechara & Damasio, 2002). This inability is assumed to lead to poor decision-making (but see Dunn et al., 2006; Maia & McClelland, 2004 for critiques).

In sum, the basic assumption underlying the IGT is that patients with vmPFC lesions prefer the bad decks because of their “myopia for the future”, whereas healthy participants base their choices on the long-term outcomes and therefore gradually learn to prefer both good decks.
Current Criticism on the Iowa Gambling Task

Despite its frequent use, the IGT has been confronted by a growing body of criticism. It is increasingly apparent that even healthy participants perform poorly because they fail to develop a preference for both good options, contradicting the assumptions of Bechara et al. (1994). More specifically, many studies have shown that healthy participants prefer the decks with infrequent losses (decks B and D) indicating that healthy participants base their choices on the frequency of losses rather than on the long-term outcomes (Caroselli, Hiscock, Scheibel, & Ingram, 2006; Dunn et al., 2006; MacPherson, Phillips, & Della Sala, 2002; Lin et al., 2007; Wilder, Weinberger, & Goldberg, 1998; Yechiam & Busemeyer, 2005). Caroselli et al. (2006) even concluded that “the card selection preferences of undergraduates were more similar to those of Bechara et al. (1994)’s seven patients with frontal lobe damage than to those of their 44 normal controls” (p. 208).

Studies applying the Soochow gambling task (SGT)—a variation of the IGT in which both good options yield high-frequent losses and both bad options yield low-frequent losses—show that healthy participants prefer the bad options over the good options, indicating again that healthy participants base their choices on the frequency of losses (Ahn et al., 2008; Chiu et al., 2008). The frequency-of-losses effect has also been found in a different version of the IGT in which Chiu and Lin (2007) established a higher contrast between rewards and losses on each trial by increasing their magnitude while keeping the traditional long-term outcomes (i.e., −250 for 10 cards from the bad decks, and +250 for 10 cards from the good decks). Chiu and Lin (2007) conclude that “the IGT contains some redundant procedures, confounding features, and problems in interpretation” (p. 9) and that “these problems should be refined to make the IGT a truly useful assessment tool” (p. 9). Further evidence for the frequency-of-losses effect in healthy participants has been reported by Huizenga, Crone, and Jansen (2007) who showed that the dominant IGT strategy in the considered age groups ranging from 6 to 25 years is to focus on the frequency of losses, a dominance that increases with age.

It is important to note that for many published IGT studies, the selection proportions for the two good decks and the two bad decks have been collapsed. This presentation method obscures the impact of the frequency-of-losses effect (Chiu & Lin, 2007; Dunn et al., 2006; Lin et al., 2007). In this article we aim to overcome this complication by reviewing the literature that does contain information from each individual deck, by re-analyzing raw data of previous IGT studies that we received upon request, and by collecting data from a relatively large sample of healthy participants.

2.2 Literature Reviews and Data Analyses

In this section we assess the three key assumptions about the performance of healthy participants on the IGT. Our analyses below provide evidence for three findings that run counter to the key IGT assumptions. First, the primary factor that drives deck selections is frequency-of-losses instead of an appreciation of long-term outcomes. Second, IGT performance is characterized by a considerable variability, both across groups and across participants within groups. Third, healthy participants fail to progress from an initial stage of exploration to a later a stage of exploitation.

Methods

In order to test the assumptions of Bechara et al. (1994) about the choice behavior of healthy participants on the IGT, we undertook three steps: We conducted two literature reviews, analyzed seven published data sets, and analyzed a large data set collected by ourselves (N = 162).
2. Performance of Healthy Participants on the Iowa Gambling Task

Literature review I: Choices from each deck separately

Our first literature review focuses on the mean proportion of choices from each of the four decks. The aim of this literature review was to quantify the size of the frequency-of-losses effect in healthy participants, defined here as participants that do not have any neurological impairments and that are typically used as control participants. We tried to locate any study that reported the mean proportion of choices from each deck separately. These studies were located by searching for IGT studies published before July 1, 2011 in ISI Web of Knowledge and Google Scholar (search term: “Iowa gambling task”), and by searching the references of relevant articles.

We restricted this literature review to studies that presented the mean proportion of choices from each of the four decks in the text or tables, or provided figures that enabled us to extract this information. For maximum compatibility, we only included IGT studies that used a procedure and payoff scheme close to the ones described by Bechara et al. (1994). This means that we only included studies that used facsimile money, that yielded a net outcome of $-250$ for the bad decks and $+250$ for the good decks in every 10 card unit, that used two high-frequent losses decks (i.e., decks A and C) and two low-frequent losses decks (i.e., decks B and D), and that used the same reward structure within the good decks and within the bad decks. Our first literature review involves 17 studies covering a total of 479 healthy participants.

Literature review II: Choices from the good and bad decks

Our second literature review focuses on the mean proportion of choices from the good and bad decks. The aim of this literature review was to quantify the variability between groups of healthy participants. We therefore searched for IGT studies reporting the mean proportion of choices from the good and bad decks. These studies were located by searching for IGT studies published before July 1, 2011 in ISI Web of Knowledge and Google Scholar (search term: “Iowa gambling task”).

We restricted this literature review to studies that presented the mean proportion of choices from the good and bad decks in the text or tables, or provided figures that enabled us to extract this information. The inclusion criteria are the same as described for the first literature review, such as only IGT studies that used a procedure and payoff scheme close to the ones described by Bechara et al. (1994). Our second literature review involves 39 groups with a total of 1427 healthy participants. Because of the vast number of published IGT studies this review is not exhaustive; however, we considered it large enough to address our research question. To the best of our knowledge, our review of healthy participants’ performance on the IGT is the most comprehensive to date.

Published data sets

We emailed all corresponding authors of the studies included in at least one of the literature reviews and requested the data of the healthy participants. Despite three attempts to contact the authors, we only received seven data sets that contained the information we needed. Table 2.2 gives an overview of all groups of healthy participants included in the literature reviews and specifies in which literature review each group is included and whether the data have been obtained or not. Data are labeled as not available if the data could no longer be located, if the complete format of the data was not available (e.g., only available collapsed over the two good decks and the two bad decks), or if the authors did not send us the data by the time we completed this manuscript. The data provided by Wood, Busemeyer, Koling, Cox, and Davis (2005) ($N = 153$) were collapsed into a single group (age: $M = 45.25, SD = 27.21$).
## 2.2. Literature Reviews and Data Analyses

### Table 2.2: Overview of the groups of healthy participants included in the literature reviews.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample size</th>
<th>Age&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Literature review&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Data received?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bark et al. (2005)</td>
<td>26</td>
<td>$M = 29.8, SD = 9.4$</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>Bechara et al. (1994)</td>
<td>44</td>
<td>20-79</td>
<td>1, 2</td>
<td>No answer</td>
</tr>
<tr>
<td>Bechara et al. (1998)</td>
<td>21</td>
<td>24-68</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Bechara et al. (1999)</td>
<td>13</td>
<td>22-58</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Blair et al. (2001)</td>
<td>23</td>
<td>10.5-13.9</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Caroselli et al. (2006)</td>
<td>141</td>
<td>$M = 21.7, SD = 4.6$</td>
<td>1, 2</td>
<td>No answer</td>
</tr>
<tr>
<td>Cavedini, Riboldi, D’Annucci, et al. (2002)</td>
<td>34</td>
<td>$M = 29.5, SD = 8.9$</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Cavedini, Riboldi, Keller, et al. (2002)</td>
<td>40</td>
<td>$M = 30.3, SD = 9.6$</td>
<td>2</td>
<td>No answer</td>
</tr>
<tr>
<td>Fernie and Tunney (2006)</td>
<td>20</td>
<td>Undergraduates</td>
<td>1, 2</td>
<td>Received</td>
</tr>
<tr>
<td>Fridberg et al. (2010)</td>
<td>15</td>
<td>$M = 29.6, SD = 7.6$</td>
<td>1, 2</td>
<td>Received</td>
</tr>
<tr>
<td>S. A. Johnson, Yechiam, Murphy, et al. (2006)</td>
<td>14</td>
<td>$M = 15.9, SD = 2.4$</td>
<td>2</td>
<td>No answer</td>
</tr>
<tr>
<td>Kester et al. (2006)</td>
<td>25</td>
<td>$M = 17.1, SD = 1.8$</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>Kjone et al. (2010)</td>
<td>20</td>
<td>$M = 33.9, SD = 11.2$</td>
<td>2</td>
<td>Received</td>
</tr>
<tr>
<td>Lehto and Elorinne (2003)</td>
<td>51</td>
<td>7-11</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Lehto and Elorinne (2003)</td>
<td>40</td>
<td>19-53</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Lev et al. (2008)</td>
<td>36</td>
<td>$M = 36.9, SD = 12.2$</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Martino et al. (2007)</td>
<td>15</td>
<td>$M = 35.0, SD = 10.9$</td>
<td>1, 2</td>
<td>No answer</td>
</tr>
<tr>
<td>North and O’Carroll (2001)</td>
<td>20</td>
<td>$M = 30.8, SD = 1.9$</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>O’Carroll and Papps (2003)</td>
<td>11</td>
<td>$M = 20.0, SD = 3.1$</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>35</td>
<td>6th grade (~11 years)</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>7th grade (~12 years)</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>57</td>
<td>8th grade (~13 years)</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>59</td>
<td>9th grade (~14 years)</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>10th grade (~15 years)</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>11th grade (~16 years)</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>12th grade (~17 years)</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>17-23</td>
<td>2</td>
<td>Not available</td>
</tr>
<tr>
<td>Petry (2001)</td>
<td>21</td>
<td>$M = 36.1, SD = 11.5$</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>Premkumar et al. (2008)</td>
<td>25</td>
<td>$M = 35.4, SD = 11.9$</td>
<td>2</td>
<td>Received</td>
</tr>
<tr>
<td>Ritter, Meador-Woodruff, and Dalack (2004)</td>
<td>15</td>
<td>$M = 47.1, SD = 10.2$</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>Rodríguez-Sánchez et al. (2005)</td>
<td>22</td>
<td>$M = 26.1, SD = 6.5$</td>
<td>1, 2</td>
<td>Received</td>
</tr>
<tr>
<td>Sevy et al. (2007)</td>
<td>20</td>
<td>$M = 33.0, SD = 10.0$</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>Shurman, Horan, and Nuechterlein (2005)</td>
<td>10</td>
<td>$M = 32.1, SD = 4.5$</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>Tomb, Hauser, Deldin, and Caramazza (2002)</td>
<td>10</td>
<td>Undergraduates</td>
<td>1, 2</td>
<td>Not available</td>
</tr>
<tr>
<td>Toplak, Jain, and Tannock (2005)</td>
<td>34</td>
<td>$M = 15.4, SD = 1.5$</td>
<td>1, 2</td>
<td>Received</td>
</tr>
<tr>
<td>Wilder et al. (1998)</td>
<td>30</td>
<td>18-52</td>
<td>1, 2</td>
<td>No answer</td>
</tr>
<tr>
<td>Wood et al. (2005)</td>
<td>88</td>
<td>$M = 22.1, SD = 4.5$</td>
<td>2</td>
<td>Received</td>
</tr>
<tr>
<td>Wood et al. (2005)</td>
<td>67</td>
<td>$M = 77.3, SD = 4.6$</td>
<td>2</td>
<td>Received</td>
</tr>
</tbody>
</table>

<sup>a</sup> Presented as mean ($M$) and standard deviation ($SD$), age range, profession, or as school grade with the approximate age in brackets.

<sup>b</sup> 1 indicates inclusion in literature review I (choices from each deck separately) and 2 indicates inclusion in literature review II (choices from the good and bad decks).
Own data set

We conducted an experiment that featured 162 healthy participants (82 female) performing a 100-trial IGT. We used a computerized version of the IGT with facsimile money and the payoff scheme as developed by Bechara et al. (1994). The mean age of all participants was 25.56 years \( (SD = 4.86) \). Female participants had a mean age of 25.52 years \( (SD = 5.32) \), and male participants of 25.60 years \( (SD = 4.37) \). All participants reported an unremarkable neurological history and were paid for their participation.

Evidence Against Assumption I: Frequency-of-Losses Effect

Literature review I: Choices from each deck separately

In order to assess whether the choices of healthy participants are guided primarily by the long-term outcomes of the decks, we conducted a literature review on the overall proportion of choices from each of the four decks. Only 17 studies with a total of 479 participants met our inclusion criteria; Table 2.3 shows the mean proportion of choices from each deck for each of the studies separately. Out of these 17 studies, 13 studies report that deck B (i.e., the bad deck with infrequent losses) is chosen more often or about as often as deck C or D (i.e., the two good decks). The remaining four studies show that healthy participants tend to prefer the good decks. One of these four studies is the study of Bechara et al. (1994); note that none of the remaining three studies yielded an overall preference for the good decks that is stronger than the one reported by Bechara et al. (1994).

The frequency-of-losses effect is evident from a consideration of the weighted mean proportion of choices from the different decks (i.e., deck A: .19, deck B: .28, deck C: .24, deck D: .29)—participants prefer the decks with infrequent losses over the decks with frequent losses. These findings suggest that the choices of healthy participants are primarily influenced by the frequency of losses rather than the expected valence of the decks.

Experimental data

The seven received data sets and our own data set enabled us to consider the mean proportion of choices from each deck as a function of trial number, as displayed in Figure 2.4. This figure shows that only the data set of Premkumar et al. (2008) supports the importance of long-term outcomes for deck selection frequencies; in this data set, healthy participants gradually develop a preference for both good decks (i.e., decks C and D) over both bad decks. The remaining data sets instead support the frequency-of-losses effect. Some data sets show a pronounced frequency-of-losses effect with a clear preference for both decks with infrequent losses (i.e., decks B and D) over both decks with frequent losses (i.e., Fridberg et al. 2010, Rodriguez-Sánchez et al. 2005, our own data set). Other data sets show a less pronounced frequency-of-losses effect, indicating that, at the end of the IGT, participants choose about equally often from decks B, C, and D, while clearly avoiding deck A (i.e., Fernie & Tunney 2006, Kjome et al. 2010, Wood et al. 2005). In the data set of Toplak et al. (2005), participants show a slight preference for decks with infrequent losses over the decks with frequent losses during the entire sequence of trials, however, at the end of the IGT participants choose about equally often from all decks. Table 2.4 summarizes the information of Figure 2.4 by presenting the overall mean proportion of choices from each deck. The weighted mean proportion of choices from each deck of the seven received data sets and our own data set underscore the frequency-of-losses effect (i.e., deck A: .17, deck B: .31, deck C: .21, deck D: .31) and thus corroborate the findings of the first literature review (Table 2.3).
2.2. Literature Reviews and Data Analyses

Figure 2.4: Mean proportion of choices from each deck within five blocks in each data set that was sent to us upon request, and in our own data set. Each block contains 20 trials, except the last block of Fridberg et al. (2010)’s data set (15 trials). Note that Kjome et al. (2010), Premkumar et al. (2008), and Wood et al. (2005) did not use exactly the same payoff scheme as reported by Bechara et al. (1994); these studies nevertheless met our inclusion criteria since they maintained the traditional net outcomes and the traditional frequency of losses structure, and the reward structure stayed the same within the good decks and within the bad decks.
2. Performance of Healthy Participants on the Iowa Gambling Task

Table 2.3: Literature review I: Mean proportion of choices from each deck separately of healthy participants. The weighted mean proportion of choices of the 17 studies support the frequency-of-losses effect.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample size</th>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bad deck</td>
<td>Bad deck</td>
<td>Good deck</td>
<td>Good deck</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with frequent losses</td>
<td>with infrequent losses</td>
<td>with frequent losses</td>
<td>with infrequent losses</td>
</tr>
<tr>
<td>Bark et al. (2005)</td>
<td>26</td>
<td>.23</td>
<td>.29</td>
<td>.24</td>
<td>.24</td>
</tr>
<tr>
<td>Bechara et al. (1994)</td>
<td>44</td>
<td>.14</td>
<td>.16</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td>Caroselli et al. (2006)</td>
<td>141</td>
<td>.22</td>
<td>.35</td>
<td>.20</td>
<td>.23</td>
</tr>
<tr>
<td>Fernie and Tunney (2006)</td>
<td>20</td>
<td>.20</td>
<td>.33</td>
<td>.22</td>
<td>.25</td>
</tr>
<tr>
<td>Fridberg et al. (2010)</td>
<td>15</td>
<td>.13</td>
<td>.30</td>
<td>.14</td>
<td>.43</td>
</tr>
<tr>
<td>Kester et al. (2006)</td>
<td>25</td>
<td>.21</td>
<td>.25</td>
<td>.26</td>
<td>.28</td>
</tr>
<tr>
<td>Martino et al. (2007)</td>
<td>15</td>
<td>.15</td>
<td>.27</td>
<td>.21</td>
<td>.37</td>
</tr>
<tr>
<td>North and O’Carroll (2001)</td>
<td>20</td>
<td>.10</td>
<td>.20</td>
<td>.36</td>
<td>.34</td>
</tr>
<tr>
<td>O’Carroll and Papps (2003)</td>
<td>11</td>
<td>.15</td>
<td>.28</td>
<td>.20</td>
<td>.37</td>
</tr>
<tr>
<td>Ritter et al. (2004)</td>
<td>15</td>
<td>.18</td>
<td>.25</td>
<td>.24</td>
<td>.33</td>
</tr>
<tr>
<td>Rodriguez-S. et al. (2005)</td>
<td>22</td>
<td>.16</td>
<td>.30</td>
<td>.20</td>
<td>.34</td>
</tr>
<tr>
<td>Sevy et al. (2007)</td>
<td>20</td>
<td>.19</td>
<td>.31</td>
<td>.23</td>
<td>.27</td>
</tr>
<tr>
<td>Shurman et al. (2005)</td>
<td>10</td>
<td>.16</td>
<td>.18</td>
<td>.34</td>
<td>.32</td>
</tr>
<tr>
<td>Tomb et al. (2002)</td>
<td>10</td>
<td>.15</td>
<td>.19</td>
<td>.34</td>
<td>.32</td>
</tr>
<tr>
<td>Toplak et al. (2005)</td>
<td>34</td>
<td>.23</td>
<td>.30</td>
<td>.20</td>
<td>.27</td>
</tr>
<tr>
<td>Wilder et al. (1998)</td>
<td>30</td>
<td>.20</td>
<td>.27</td>
<td>.24</td>
<td>.29</td>
</tr>
<tr>
<td>Total sample size, mean (^a)</td>
<td>479</td>
<td>.19</td>
<td>.28</td>
<td>.24</td>
<td>.29</td>
</tr>
</tbody>
</table>

Note. All studies are based on a 100 trial-IGT except Fridberg et al. (2010) (95-trial IGT).

\(^a\) Weighted by the sample size of each study.

Evidence Against Assumption II: High Variability in Performance Between and Within Groups

Literature review II: Choices from the good and bad decks

In order to assess whether the choices of healthy participants are relatively consistent between groups, we conducted a literature review on the mean proportion of choices from the good decks (i.e., decks C and D) and the bad decks (i.e., decks A and B). Table 2.5 presents the mean proportion of choices from good and bad decks for all groups included in our second literature review, and Figure 2.5 presents the histogram of the corresponding mean proportion of choices from the good decks. From the table and the figure, it is evident that the majority of the studies reported a weak overall preference for the good decks (i.e., between 50% and 60% choices from the good decks), and that some studies even reported an overall preference for the bad decks. Few studies reported a pronounced overall preference for the good decks, and none reported a preference for the good decks that is as strong as the one reported by Bechara et al. (1994).

Our second literature review reveals substantial between-group variability in deck preference. In addition, the review underscores that the performance of healthy participants is generally much worse than initially reported by Bechara et al. (1994). It is especially striking that the majority of the studies reports a weak overall preference for the good decks. The strong evidence for the frequency-of-losses effect, discussed in the last section, suggests that a weak overall preference for...
Table 2.4: Mean proportion of choices from each deck separately of healthy participants in each data set that was sent to us upon request, and in our own data set.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample size</th>
<th>Deck A Bad deck with frequent losses</th>
<th>Deck B Bad deck with infrequent losses</th>
<th>Deck C Good deck with frequent losses</th>
<th>Deck D Good deck with infrequent losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fernie and Tunney (2006)</td>
<td>20</td>
<td>.20</td>
<td>.33</td>
<td>.22</td>
<td>.25</td>
</tr>
<tr>
<td>Fridberg et al. (2010)</td>
<td>15</td>
<td>.13</td>
<td>.30</td>
<td>.14</td>
<td>.43</td>
</tr>
<tr>
<td>Kjome et al. (2010)*</td>
<td>19</td>
<td>.17</td>
<td>.26</td>
<td>.22</td>
<td>.35</td>
</tr>
<tr>
<td>Premkumar et al. (2008)*</td>
<td>25</td>
<td>.16</td>
<td>.21</td>
<td>.27</td>
<td>.36</td>
</tr>
<tr>
<td>Rodriguez-S. et al. (2005)</td>
<td>19</td>
<td>.16</td>
<td>.30</td>
<td>.20</td>
<td>.34</td>
</tr>
<tr>
<td>Toplak et al. (2005)</td>
<td>34</td>
<td>.23</td>
<td>.30</td>
<td>.20</td>
<td>.27</td>
</tr>
<tr>
<td>Wood et al. (2005)*</td>
<td>153</td>
<td>.17</td>
<td>.29</td>
<td>.23</td>
<td>.31</td>
</tr>
<tr>
<td>Own data set</td>
<td>162</td>
<td>.15</td>
<td>.35</td>
<td>.20</td>
<td>.30</td>
</tr>
</tbody>
</table>

Total sample size, mean \(^b\) 447 .17 .31 .21 .31

Note. All studies are based on a 100-trial IGT except Fridberg et al. (2010) (95-trial IGT). Studies marked by * did not use exactly the same payoff scheme as reported by Bechara et al. (1994); these studies nevertheless met our inclusion criteria since they maintained the traditional net outcomes and the traditional frequency of losses structure, and the reward structure stayed the same within the good decks and within the bad decks.

\(^a\) Sample sizes might be lower than in Table 2, 3, and 5 due to incomplete received data sets.

\(^b\) Weighted by the sample size of each study.

Figure 2.5: Literature review II: Histogram of the mean proportion of choices from good decks (i.e., decks C and D) of 39 groups of healthy participants. The majority of the groups shows a weak overall preference for the good decks while some groups show an overall preference for the bad decks and a few groups show a pronounced overall preference for the good decks, as reported by Bechara et al. (1994).
Table 2.5: Literature review II: Mean proportion of choices from good (decks C and D) and bad decks (decks A and B) of healthy participants.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Sample size</th>
<th>Age *</th>
<th>Good decks</th>
<th>Bad decks</th>
<th>Good - bad decks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caroselli et al. (2006)</td>
<td>141</td>
<td>$M = 21.7, SD = 4.6$</td>
<td>.43</td>
<td>.57</td>
<td>-.14</td>
</tr>
<tr>
<td>Fernie and Tunney (2006)</td>
<td>20</td>
<td>Undergraduates</td>
<td>.47</td>
<td>.53</td>
<td>-.06</td>
</tr>
<tr>
<td>Toplak et al. (2005)</td>
<td>34</td>
<td>$M = 15.4, SD = 1.5$</td>
<td>.47</td>
<td>.53</td>
<td>-.06</td>
</tr>
<tr>
<td>Bark et al. (2005)</td>
<td>26</td>
<td>$M = 29.8, SD = 9.4$</td>
<td>.48</td>
<td>.52</td>
<td>-.04</td>
</tr>
<tr>
<td>Sevy et al. (2005)</td>
<td>20</td>
<td>$M = 33.0, SD = 10.0$</td>
<td>.50</td>
<td>.50</td>
<td>.00</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>7th grade (~12 years)</td>
<td>.53</td>
<td>.47</td>
<td>.06</td>
</tr>
<tr>
<td>Wilder et al. (1998)</td>
<td>30</td>
<td>18-52</td>
<td>.53</td>
<td>.47</td>
<td>.06</td>
</tr>
<tr>
<td>Kester et al. (2006)</td>
<td>25</td>
<td>$M = 17.1, SD = 1.8$</td>
<td>.54</td>
<td>.46</td>
<td>.08</td>
</tr>
<tr>
<td>Lehto and Elorinne (2003)</td>
<td>51</td>
<td>7-11</td>
<td>.54</td>
<td>.46</td>
<td>.08</td>
</tr>
<tr>
<td>Rodriguez-Sánchez et al. (2005)</td>
<td>22</td>
<td>$M = 26.1, SD = 6.5$</td>
<td>.54</td>
<td>.46</td>
<td>.08</td>
</tr>
<tr>
<td>Wood et al. (2005)</td>
<td>88</td>
<td>$M = 22.1, SD = 4.5$</td>
<td>.54</td>
<td>.46</td>
<td>.08</td>
</tr>
<tr>
<td>Cavedini, Riboldi, D'Annucci, et al. (2002)</td>
<td>34</td>
<td>$M = 29.5, SD = 8.9$</td>
<td>.55</td>
<td>.45</td>
<td>.10</td>
</tr>
<tr>
<td>Cavedini, Riboldi, Keller, et al. (2002)</td>
<td>40</td>
<td>$M = 30.3, SD = 9.6$</td>
<td>.55</td>
<td>.45</td>
<td>.10</td>
</tr>
<tr>
<td>Wood et al. (2005)</td>
<td>67</td>
<td>$M = 77.3, SD = 4.6$</td>
<td>.55</td>
<td>.45</td>
<td>.10</td>
</tr>
<tr>
<td>Lehto and Elorinne (2003)</td>
<td>40</td>
<td>19-53</td>
<td>.56</td>
<td>.44</td>
<td>.12</td>
</tr>
<tr>
<td>Lev et al. (2008)</td>
<td>36</td>
<td>$M = 36.9, SD = 12.2$</td>
<td>.56</td>
<td>.44</td>
<td>.12</td>
</tr>
<tr>
<td>Fridberg et al. (2010)</td>
<td>15</td>
<td>$M = 29.6, SD = 7.6$</td>
<td>.57</td>
<td>.43</td>
<td>.14</td>
</tr>
<tr>
<td>Kjome et al. (2010)</td>
<td>20</td>
<td>$M = 33.9, SD = 11.2$</td>
<td>.57</td>
<td>.43</td>
<td>.14</td>
</tr>
<tr>
<td>O’Carroll and Papps (2003)</td>
<td>11</td>
<td>$M = 20.0, SD = 3.1$</td>
<td>.57</td>
<td>.43</td>
<td>.14</td>
</tr>
<tr>
<td>Ritter et al. (2004)</td>
<td>15</td>
<td>$M = 47.1, SD = 10.2$</td>
<td>.57</td>
<td>.43</td>
<td>.14</td>
</tr>
<tr>
<td>S. A. Johnson et al. (2006)</td>
<td>14</td>
<td>$M = 15.9, SD = 2.4$</td>
<td>.58</td>
<td>.42</td>
<td>.16</td>
</tr>
<tr>
<td>Martino et al. (2007)</td>
<td>15</td>
<td>$M = 35.0, SD = 10.9$</td>
<td>.58</td>
<td>.42</td>
<td>.16</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>35</td>
<td>6th grade (~11 years)</td>
<td>.58</td>
<td>.42</td>
<td>.16</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>57</td>
<td>8th grade (~13 years)</td>
<td>.59</td>
<td>.41</td>
<td>.18</td>
</tr>
<tr>
<td>Petry (2001)</td>
<td>21</td>
<td>$M = 36.1, SD = 11.5$</td>
<td>.59</td>
<td>.41</td>
<td>.18</td>
</tr>
<tr>
<td>Blair et al. (2001)</td>
<td>23</td>
<td>10.5-13.9</td>
<td>.60</td>
<td>.40</td>
<td>.20</td>
</tr>
<tr>
<td>Bechara et al. (1998)</td>
<td>21</td>
<td>24-68</td>
<td>.62</td>
<td>.38</td>
<td>.24</td>
</tr>
<tr>
<td>Bechara et al. (1999)</td>
<td>13</td>
<td>22-58</td>
<td>.63</td>
<td>.37</td>
<td>.26</td>
</tr>
<tr>
<td>Premkumar et al. (2008)</td>
<td>25</td>
<td>$M = 35.4, SD = 11.9$</td>
<td>.63</td>
<td>.37</td>
<td>.26</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>59</td>
<td>9th grade (~14 years)</td>
<td>.64</td>
<td>.36</td>
<td>.28</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>10th grade (~15 years)</td>
<td>.64</td>
<td>.36</td>
<td>.28</td>
</tr>
<tr>
<td>Shurman et al. (2005)</td>
<td>10</td>
<td>$M = 32.1, SD = 4.5$</td>
<td>.66</td>
<td>.34</td>
<td>.32</td>
</tr>
<tr>
<td>Tomb et al. (2002)</td>
<td>10</td>
<td>Undergraduates</td>
<td>.66</td>
<td>.34</td>
<td>.32</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>11th grade (~16 years)</td>
<td>.67</td>
<td>.33</td>
<td>.34</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>17-23</td>
<td>.67</td>
<td>.33</td>
<td>.34</td>
</tr>
<tr>
<td>Overman et al. (2004)</td>
<td>60</td>
<td>12th grade (~17 years)</td>
<td>.68</td>
<td>.32</td>
<td>.36</td>
</tr>
<tr>
<td>Bechara et al. (1994)</td>
<td>44</td>
<td>20-79</td>
<td>.70</td>
<td>.30</td>
<td>.40</td>
</tr>
<tr>
<td>North and O’Carroll (2001)</td>
<td>20</td>
<td>$M = 30.8, SD = 1.9$</td>
<td>.70</td>
<td>.30</td>
<td>.40</td>
</tr>
</tbody>
</table>

*Note. All studies are based on a 100-trial IGT except Fridberg et al. (2010) (95-trial IGT) and Overman et al. (2004) (200-trial IGT).

*Presented as mean (M) and standard deviation (SD), age range, profession, or as school grade with the approximate age in brackets.
2.2. Literature Reviews and Data Analyses

Figure 2.6: Choice behavior of healthy participants in Fridberg et al. (2010). The left panel shows the mean proportion of choices from good and bad decks within five blocks. Each block contains 20 trials, except the last block (15 trials). The right panel shows the mean proportion of choices from each deck within five blocks. This data set illustrates how collapsing the choices from the good decks and from the bad decks might hide the frequency-of-losses effect.

The good decks may only be due to a summing effect. For example, the data of Fridberg et al. (2010) show that healthy participants prefer the good decks over the bad decks (i.e., .57 vs. .43, see Table 2.5). This finding is underscored by the left panel of Figure 2.6 that presents the mean proportion of good and bad choices as a function of trial number (cf. Figure 1 in Fridberg et al., 2010). This panel shows that the proportion of good choices increases as the number of trials increases, and leads the authors to conclude that “Controls subsequently learned to select from the advantageous decks” (Fridberg et al., 2010, p. 31). However, this conclusion is mainly due to a summing effect, and hides the frequency-of-losses effect, as can be seen in our first literature review where the mean proportion of choices from each deck were presented separately (i.e., deck A: .13, deck B: .30, deck C: .14, deck D: .43; Table 2.3). The right panel of Figure 2.6 (cf. Figure 2 in Fridberg et al., 2010) displays the mean proportion of choices from each deck as a function of trial number. This figure shows that the decks with infrequent losses (i.e., decks B and D) are at any point during the IGT preferred over the decks with frequent losses (i.e., decks A and C). The reason why the control participants appear to learn to prefer the good decks is that the choices from deck D strongly increase whereas the choices from the other three decks slowly decrease; nevertheless, deck B is still chosen much more often than deck A or C at any point during the IGT. This example illustrates how collapsing the choices from the good decks and from the bad decks might hide the frequency-of-losses effect and highlights the importance of analyzing the choice behavior of each deck separately.

While reviewing the literature, we also noticed that some studies report a high within-group variability in the performance of healthy participants on the IGT. We found six studies that explicitly reported the number of healthy participants that showed impaired performance (i.e., overall preference for the bad decks). Table 2.6 lists these six studies and the corresponding number and percentage of healthy participants with impaired performance. This table points out that a
Table 2.6: Number and percentage of healthy participants with impaired performance (> 50% choices from the bad decks) as specified in six studies.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Participants with impaired performance [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adinoff et al. (2003)</td>
<td>4 [27.0]</td>
</tr>
<tr>
<td>Bechara et al. (1999)</td>
<td>3 [23.0]</td>
</tr>
<tr>
<td>Bechara et al. (2001)</td>
<td>13 [32.5]</td>
</tr>
<tr>
<td>Bechara and Damasio (2002)</td>
<td>18 [37.0]</td>
</tr>
<tr>
<td>Bowman and Turnbull (2003)</td>
<td>2 [12.0]</td>
</tr>
<tr>
<td>Lehto and Elorinne (2003)</td>
<td>3 [7.5]</td>
</tr>
</tbody>
</table>

These participants were classified as impaired because they performed within the range of vmPFC patients.

These participants even made ≥ 70% choices from the bad decks.

substantial number of healthy participants perform poorly on the IGT, suggesting that there is high within-group variability in the performance of healthy participants in these six studies. Moreover, the table suggests that a substantial number of participants violate the assumption that healthy participants prefer the good decks over the bad decks.

In sum, the performance of healthy participants on the IGT is characterized by high variability between groups: In some studies healthy participants seem to behave according to the assumptions of Bechara et al. (1994) by expressing a strong preference for the good decks; however, in most studies the preference for the good decks is relatively weak, and in some studies healthy participants express a preference for the bad decks, in complete contradiction to the assumptions that underlie the IGT. In addition, a few data sets suggest that there may also be substantial variability across participants within the same study, a finding we will return to in more detail in the next section.

Experimental data

The extent of within-group variability in performance of healthy participants can be understood best by an individual-participants analysis of deck selection profiles. Figure 2.7 shows the deck selection profiles of five participants in our own data set. This figure illustrates that choice behavior differs dramatically between the five healthy participants: The first participant keeps switching between all decks and does not develop a preference for any deck; the second participant seems to develop a preference for the bad decks with occasional switches to deck C; the third and fourth participant both develop a preference for decks B and D, but the third participant tends to stay longer at the same deck than the fourth participant; finally, the fifth participant develops a preference for deck B and only rarely switches to decks C or D. These individual deck selection profiles suggest that each participant displays a highly idiosyncratic choice behavior. The deck selection profiles of all available data sets further underscore the substantial within-group variability, as can be seen in the online appendix containing a total of 394 individual deck selection profiles from six different studies.

The extent of within-group variability can be further assessed by determining the number of participants that perform in the deficient range i.e., > 50% choices from the bad decks. Table 2.7 presents the number and percentage of healthy participants with impaired performance in each data set that was sent to us upon request, and in our own data set. From this table, it is apparent that a...
Figure 2.7: Deck selection profiles of five participants in our own data set express the high within-group variability in the performance of healthy participants. The filled circles indicate the occurrence of rewards and losses together; the empty circles indicate the occurrence of only rewards.
Table 2.7: Number and percentage of healthy participants with impaired performance (>50% choices from the bad decks) in each data set that was sent to us upon request, and in our own data set.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Participants with impaired performance [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fernie and Tunney (2006)</td>
<td>13 [65.0]</td>
</tr>
<tr>
<td>Fridberg et al. (2010)</td>
<td>5 [33.3]</td>
</tr>
<tr>
<td>Kjome et al. (2010)*</td>
<td>6 [31.6]</td>
</tr>
<tr>
<td>Premkumar et al. (2008)*</td>
<td>5 [20.0]</td>
</tr>
<tr>
<td>Rodriguez-S. et al. (2005)</td>
<td>9 [47.4]</td>
</tr>
<tr>
<td>Toplak et al. (2005)</td>
<td>20 [58.8]</td>
</tr>
<tr>
<td>Wood et al. (2005)*</td>
<td>55 [35.9]</td>
</tr>
<tr>
<td>Own data set</td>
<td>79 [48.8]</td>
</tr>
</tbody>
</table>

Note. All studies are based on a 100-trial IGT except Fridberg et al. (2010) (95-trial IGT). Studies marked by * did not use exactly the same payoff scheme as reported by Bechara et al. (1994); these studies nevertheless met our inclusion criteria since they maintained the traditional net outcomes and the traditional frequency of losses structure, and the reward structure stayed the same within the good decks and within the bad decks.

substantial number of healthy participants—in some studies, even the majority of the participants—perform deficiently on the IGT. This table corroborates the findings from Table 2.6 and Figure 2.7: The performance of healthy participants is characterized by high within-group variability, and a substantial number of participants violate the assumption that healthy participants prefer the good decks over the bad decks.

Evidence Against Assumption III: Absence of a Pronounced Exploration-Exploitation Tradeoff

Bechara et al. (1994) assume that healthy participants first explore all decks, then switch to the good decks, and only occasionally return to the bad decks. In a reinforcement-learning context (Sutton & Barto, 1998), this means that healthy participants progress from an initial stage of exploration to a later stage of exploitation. This assumption implies that the number of switches decreases as the number of trials increases. However, the available data sets fail to support this assumption. Figure 2.8 shows the mean proportion of switches as a function of trial number. The data sets of Premkumar et al. (2008) and Wood et al. (2005) show a slight decrease in the mean proportion of switches across trials; however, this decrease is not as pronounced as assumed by Bechara et al. (1994) who showed that, in the case of two typical healthy participants, the last 10 trials featured only one and no switch (cf. Figure 2.3). In the remaining data sets, the mean proportion of switches remains surprisingly stable. Therefore, Figure 2.8 supports the conclusion that healthy participants fail to progress from an initial stage of exploration to a later stage of exploitation. In a reinforcement-learning context (Sutton & Barto, 1998), this suggests that participants failed to learn about the relative value of the choice options that they have available.

2.3 Discussion

This article focused on performance of healthy participants on the IGT. Two literature reviews and the analysis of eight data sets together challenge the assumptions of Bechara et al. (1994).

As in the previous section, the data set of Rodríguez-Sánchez et al. (2005) and Toplak et al. (2005) could not be included in this analysis because we only obtained these data in bins of 20 trials.
2.3. Discussion

Figure 2.8: Mean proportion of switches within 10 blocks in each data set that was sent to us upon request, and in our own data set. Each block contains 10 trials, except the first block in all data sets (9 trials) and the last block in Fridberg et al. (2010)'s data set (5 trials). The mean proportion of switches remains stable across trials in each data set indicating the absence of a pronounced exploration-exploitation tradeoff. Note that Kjome et al. (2010), Premkumar et al. (2008), and Wood et al. (2005) did not use exactly the same payoff scheme as reported by Bechara et al. (1994); these studies nevertheless met our inclusion criteria since they maintained the traditional net outcomes and the traditional frequency of losses structure, and the reward structure stayed the same within the good and bad decks.
about the IGT choice behavior of healthy participants. Our findings reveal the presence of a frequency-of-losses effect in healthy participants, as already indicated by many studies (Ahn et al., 2008; Caroselli et al., 2006; Chiu & Lin, 2007; Chiu et al., 2008; Dunn et al., 2006; Huizenga et al., 2007; MacPherson et al., 2002; Lin et al., 2007; Wilder et al., 1998; Yechiam & Busemeyer, 2005). In addition, we showed that performance of healthy participants is characterized by considerable variability, both across groups and across participants within the same study, and that healthy participants fail to progress from an initial stage of exploration to a later stage of exploitation. These findings clearly contradict the common belief that “most healthy participants sample cards from each deck, and after about 40 or 50 selections are fairly good at sticking to the good decks”.

This evidence against the key IGT assumptions suggests that IGT scores may not serve as a valid measure of decision-making deficits in clinical populations, for two different reasons (cf. Dunn et al., 2006). First, the high variability in healthy participants’ performance complicates the prevailing interpretation of IGT data. Specifically, conclusions about deficient choice performance in clinical populations depend strongly on the performance of the control group. For instance, the absence of a difference between a clinical and control group may be due to intact decision-making ability in the clinical group, or it may be due to poor performance of the control group (Dunn et al., 2006).

Second, our findings question the ecological validity of IGT scores. Since a substantial number of healthy participants perform poorly on the IGT, but likely do not show any decision-making deficits in real life, it is unclear to what extent IGT scores measure everyday, real-life decision making (Dunn et al., 2006).

Our analysis demonstrated that, in order to reveal the frequency-of-losses effect and accurately describe the learning process, it is important that future studies analyze the choice behavior as a function of the trial number from each deck separately. One way to capture the learning process is to apply trend analyses (e.g., a linear or quadratic contrast) on the choice pattern from each deck (Dunn et al., 2006; Premkumar et al., 2008).

Our results showed that many healthy participants perform poorly on the IGT, begging the question as to why they do not learn to prefer both good decks over both bad decks. Altogether, it seems that many participants find it particularly difficult to figure out that deck B is a bad deck—after all, deck B yields high immediate, constant rewards. One explanation for why healthy participants do not learn to prefer both good decks over deck B is related to the payoff scheme as developed by Bechara et al. (1994). As first pointed out by Lin et al. (2007), cards from deck C never yield a net loss, but cards from deck A do. This goes against Bechara et al. (1994)’s idea of designing one good deck and one bad deck with high-frequent losses and suggests that the good decks and bad decks are pseudo-balanced. Decks B, C, and D are more similar than expected because they yield either no or very few net losses, whereas deck A yields frequent and big net losses and is thus the only deck that clearly differs from the others with respect to the net outcomes. Another explanation is that the payoff scheme encourages participants to only focus on the immediate losses; the losses vary but the immediate rewards are completely predictable as they are constant across trials for each deck. The predictability of immediate rewards also contradicts Bechara et al. (1994)’s intention of “uncertainty of reward and punishment” (p. 8).

Another explanation for the surprisingly poor performance of healthy participants is a lack of motivation. Indications for a lack of motivation can be seen in the individual deck selection profiles, as shown in Figure 2.7 and in the online appendix. Studying these profiles, it seems impossible to identify a learning process for the majority of the participants. Many participants keep switching between the decks, irrespective of the occurrence of losses and of the long-term outcomes of the decks. However, Fernie and Tunney (2006) have shown that participants do not perform better

---

when they are more motivated to do well (i.e., by using real money vs. facsimile money); instead, participants did improve when they were given more information about the task (i.e., some decks are worse than others, most money is made when the poor options are avoided). This suggests that it is the difficulty of the IGT rather than the lack of motivation that contributes to a poor performance of healthy participants.

One final explanation for the poor performance of healthy participants is that 100 trials are not enough to learn much about the nature of the decks. In support of this conjecture, Vandekerckhove, et al. (2010) showed that healthy participants do learn to prefer the good decks over the bad decks, but that they require at least 100 trials to do so. The reason for the slow learning process might be the infrequent occurrence of losses in decks B and D (i.e., once in 10 cards), providing participants with too little information to learn quickly that deck B should be avoided.

It is still unclear what accounts for the relatively poor performance of clinical populations on the IGT. It may be that clinical populations are even more susceptible to the frequency-of-losses effect than healthy participants are. It may also be that clinical populations are primarily driven by the immediate prospects of the decks, as originally proposed by Bechara et al. (1994). In order to understand the factors that determine the choices on the IGT, one possibility is to fit a multivariate normal mixture model to clinical samples and samples of healthy participants (cf. Huizenga et al., 2007). This analysis allows one to identify groups that use distinct IGT decision strategies, such as focusing on the frequency of losses, focusing on immediate rewards, or focusing on long-term outcomes. Such an analysis does not explain the learning process and therefore requires a relatively long IGT. Another possibility to determine the factors driving performance on the IGT is to apply a formal model that explicitly describes how the learning process unfolds over time (e.g., Busemeyer & Stout, 2002).

In sum, we showed that the IGT is characterized by several anomalous phenomena that go against the assumptions of the prevailing theory (Bechara et al., 1994). Specifically, performance of healthy participants is primarily influenced by the frequency of losses rather than the long-term outcomes; moreover, performance of healthy participants is highly idiosyncratic, and often lacks a tendency to first explore decks and then exploit those that are most attractive. These anomalies question the extent to which IGT scores measure real-life decision-making deficits and complicate the prevailing interpretation of IGT data. Our findings suggest that the IGT needs to be reevaluated before being accepted as convincing tool to measure decision-making deficits in clinical populations.

**Acknowledgements**

We thank Fernie and Tunney (2006), Fridberg et al. (2010), Kjome et al. (2010), Premkumar et al. (2008), Rodríguez-Sánchez et al. (2005), Toplak et al. (2005), and Wood et al. (2005) for providing the data used in this article. AH and JN are supported by the Federal Ministry of Education and Research (BMBF), Germany, FKZ: 01EO1001, and HS is supported by the Netherlands Organisation for Scientific Research (NWO: 404-10-086).
Chapter 3

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

This chapter has been published as:
A comparison of reinforcement-learning models for the Iowa gambling task using parameter space partitioning.
The Journal of Problem Solving, 5, Article 2.

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 reinforcement-learning 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 (i.e., 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.

The Iowa gambling task (IGT; Bechara et al., 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,
3. A Comparison of Reinforcement-Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning

such as patients with pathological gambling disorder (Cavedini, Riboldi, Keller, et al., 2002), attention-deficit-hyperactivity disorder (Agay, Yechiam, Carmel, & Levkovitz, 2010; Toplak et al., 2005), obsessive-compulsive disorder (Cavedini, Riboldi, D’Annuzzi, et al., 2002), psychopathic tendencies (Blair et al., 2001), bipolar disorder (Brambilla et al., 2012), and schizophrenia (Martino et al., 2007; Premkumar et al. 2008). In addition, the IGT has been applied to cocaine addicts (Stout et al., 2004), chronic cannabis users (Fridberg et al., 2010), traffic offenders (Lev et al., 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; W. 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; S. A. Johnson et al., 2006; Kester et al., 2006; Kjome et al., 2010; Lane, Yechiam, & Busemeyer, 2006; Lev et al., 2008; Lobjoie, Yechiam, Sorocco, Vincent, & Collins, 2006; Peathfield, Parkinson, & Intriligator, 2012; Premkumar et al., 2008; Sevy et al., 2006; 2007; Stout et al., 2004; van den Bos, Homberg, Gijsbers, den Heijer, & Cuppen, 2009; Wood et al., 2005; Yechiam et al., 2005; Yechiam, Hayden, et al., 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 challenged by the Prospect Valence Learning model (PVL; Ahn et al., 2008, 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
3.1. 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, 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 labeled as decks C and D, whereas the bad, risky decks are labeled as decks A and B. Table 3.1 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 (i.e., decks A and C) and two decks yield infrequent losses (i.e., decks B and D).

Table 3.1: Payoff scheme of the traditional IGT as developed by Bechara et al. (1994).

<table>
<thead>
<tr>
<th></th>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with frequent</td>
<td>with infrequent</td>
<td>Good deck</td>
<td>Good deck</td>
</tr>
<tr>
<td>Bad deck</td>
<td>losses</td>
<td>losses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reward/trial</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Number of losses/10 cards</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Loss/10 cards</td>
<td>−1250</td>
<td>−1250</td>
<td>−250</td>
<td>−250</td>
</tr>
<tr>
<td>Net outcome/10 cards</td>
<td>−250</td>
<td>−250</td>
<td>250</td>
<td>250</td>
</tr>
</tbody>
</table>

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 3.2 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...
3. A Comparison of Reinforcement-Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning

Table 3.2: Formalization of the three reinforcement-learning models.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Model(s)</th>
<th>Model equation</th>
<th>Free parameter(s)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility function</td>
<td>EV</td>
<td>$u_k(t) = (1 - w) \cdot W(t) + w \cdot L(t)$</td>
<td>$w$: Loss aversion parameter</td>
<td>[0, 1]</td>
</tr>
<tr>
<td></td>
<td>PVL, EV-PU</td>
<td>$u_k(t) = \begin{cases} X(t)^A &amp; \text{if } X(t) \geq 0 \ -w \cdot</td>
<td>X(t)</td>
<td>^A &amp; \text{if } X(t) &lt; 0 \end{cases}$</td>
</tr>
<tr>
<td>Learning rule</td>
<td>EV, EV-PU</td>
<td>$EV_k(t) = EV_k(t - 1) + a \cdot (u_k(t) - EV_k(t - 1))$</td>
<td>$a$: Recency parameter</td>
<td>[0, 1]</td>
</tr>
<tr>
<td></td>
<td>PVL</td>
<td>$EV_k(t) = a \cdot EV_k(t - 1) + \delta_k(t) \cdot u_k(t)$</td>
<td>$a$: Recency parameter</td>
<td>[0, 1]</td>
</tr>
<tr>
<td>Choice rule</td>
<td>all</td>
<td>$P[S_k(t + 1)] = \frac{e^{\theta(t)EV_k(t)}}{\sum_{j=1}^{4} e^{\theta(t)EV_j(t)}}$</td>
<td>$c$: Consistency parameter</td>
<td>[-5, 5]</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>EV, EV-PU</td>
<td>$\theta(t) = (t/10)^c$</td>
<td>$c$: Consistency parameter</td>
<td>[0, 5]</td>
</tr>
<tr>
<td></td>
<td>PVL</td>
<td>$\theta(t) = 3^c - 1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

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).

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 $w$, i.e., $w < .5$, 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 $w$, i.e., $w > .5$, 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
3.1. The IGT and Three Reinforcement-Learning Models

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 \(a\) 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 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.
3. A Comparison of Reinforcement-Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning

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. 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. 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, i.e., 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.

---

2 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.

3 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.
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 (i.e., decks C and D) over the risky decks (i.e., 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 (i.e., decks B and D) over the decks with frequent losses (i.e., 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 (i.e., 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. 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 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. 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 et al. (2006), Chiu & Lin (2007), Chiu et al. (2008), Dunn et al. (2006), MacPherson et al. (2002), Lin et al. (2007), Wilder et al. (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 (2009) and Yechiam and Ert (2007) for newly developed methods to compare RL models applied to experiments with two and three choice options, respectively.
3. A Comparison of Reinforcement-Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning

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, Wetzels, Horstmann, 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.

3.2 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 3.1. 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 3.1.

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 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, i.e., \( \{C, D\} \succ \{A, B\} \), (2) preference for bad decks over good decks, i.e., \( \{A, B\} \succ \{C, D\} \), (3) preference for decks with infrequent losses over decks with frequent losses, i.e., \( \{B, D\} \succ \{A, C\} \), (4) preference for decks with frequent losses over decks with infrequent losses, i.e., \( \{A, C\} \succ \{B, D\} \), 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, Wetzels, Horstmann, et al., 2013) that a good model for
3.2. Comparison of the EV, PVL, and EV-PU Models

Figure 3.1: 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.

The IGT should only generate the first three types of choice patterns, as presented in the left panel of Figure 3.1.

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”, i.e., \( \{C, D\} \succ \{A, B\} \). 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 \cite{Steingroever, Wetzels, Horstmann, et al., 2013}. 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 3.1.

\footnote{Note that we did not use the data of Rodriguez-Sánchez et al. \cite{2005}, and Toplak et al. \cite{2010} because we have received their data only in bins of several trials.}
3. A Comparison of Reinforcement-Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning

Table 3.3: Proportions of choice patterns generated by each model.

<table>
<thead>
<tr>
<th>Choice pattern</th>
<th>Proportion of all choice patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EV</td>
</tr>
<tr>
<td>Good &gt; bad decks</td>
<td>{C, D} &gt; {A, B}</td>
</tr>
<tr>
<td>Bad &gt; good decks</td>
<td>{A, B} &gt; {C, D}</td>
</tr>
<tr>
<td>Infr. &gt; frequent losses</td>
<td>{B, D} &gt; {A, C}</td>
</tr>
<tr>
<td>Frequent &gt; infr. losses</td>
<td>{A, C} &gt; {B, D}</td>
</tr>
<tr>
<td>Remaining</td>
<td></td>
</tr>
</tbody>
</table>

by dividing by 100 (cf. Ahn et al., 2011). The codes for the PSP studies are available on www.helensteingroever.com.

Results

Broad definition of choice patterns

Table 3.3 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.

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, i.e., 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, Wetzels, Horstmann, 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).
3.2. Comparison of the EV, PVL, and EV-PU Models

Table 3.4: 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, Wetzels, Horstmann, et al.] (2013).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(C + D) ≥ .65</td>
<td>54</td>
<td>.10 [.05]</td>
<td>.14 [.05]</td>
<td>.36 [.17]</td>
<td>.40 [.14]</td>
<td>.35 [.08, .52] (0.00, 0.96)</td>
</tr>
<tr>
<td>(A + B) ≥ .65</td>
<td>18</td>
<td>.25 [.07]</td>
<td>.52 [.11]</td>
<td>.11 [.05]</td>
<td>.12 [.06]</td>
<td>.43 [.31, .58] (0.10, 0.86)</td>
</tr>
<tr>
<td>(B + D) ≥ .65</td>
<td>139</td>
<td>.12 [.05]</td>
<td>.37 [.12]</td>
<td>.13 [.05]</td>
<td>.39 [.12]</td>
<td>.47 [.28, .66] (0.02, 1.00)</td>
</tr>
</tbody>
</table>

Restricted definition of choice patterns

Table 3.4 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, Wetzels, Horstmann, 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 a RL model for the IGT is able to produce such choice patterns. Table 3.4 also contains for each group the mean proportion of switches during the last 50 trials and statistics quantifying the distribution of switches in each group (i.e., the interquartile range and the minimum and maximum proportion of switches during the last 50 trials). This information is visualized by the boxplots shown in the first column of Figure 3.2. 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 minimum and maximum proportion 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 3.3, 3.7, and 3.10 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, i.e., at least 65% choices from the two most preferred decks, can be found in the online appendix.

The mean proportions of choices shown in Table 3.4 were used to investigate whether the models can also generate these pronounced deck preferences. This means that we searched the results of the PSP studies for those parameter combinations that yielded pronounced deck preference as presented in Table 3.4 (i.e., a pronounced preference for the good decks (C ≥ .36, D ≥ .40), a pronounced preference for the bad decks (A ≥ .25, B ≥ .52), and a pronounced preference for the decks with infrequent losses (B ≥ .37, D ≥ .39)). Table 3.5 presents for each model the proportion

---

*The deck selection profiles can be downloaded here: [https://dl.dropbox.com/u/12798592/DeckSelectionProfilesPSP.zip](https://dl.dropbox.com/u/12798592/DeckSelectionProfilesPSP.zip)*
Table 3.5: 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 proportion of switches during the last 50 trials are presented.

<table>
<thead>
<tr>
<th>Model</th>
<th>Choice pattern</th>
<th>Proportion of all choice patterns</th>
<th>Parameter combinations</th>
<th>Switches during the last 50 trials [25%, 75% quantile] (min, max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EV</td>
<td>$C \geq .36, D \geq .40$</td>
<td>.00027</td>
<td>$A \geq .25, B \geq .52$</td>
<td>.00013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$w \ [sd]$</td>
<td>.45 [.04]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$a \ [sd]$</td>
<td>.04 [.03]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$c \ [sd]$</td>
<td>none</td>
</tr>
<tr>
<td>PVL</td>
<td>$C \geq .36, D \geq .40$</td>
<td>.010</td>
<td>$A \geq .25, B \geq .52$</td>
<td>.0000035</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$w \ [sd]$</td>
<td>.66 [.22]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$a \ [sd]$</td>
<td>.92 [.07]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$c \ [sd]$</td>
<td>.46 [.76]</td>
</tr>
<tr>
<td>EV-PU</td>
<td>$C \geq .36, D \geq .40$</td>
<td>.0019</td>
<td>$A \geq .25, B \geq .52$</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$w \ [sd]$</td>
<td>.66 [.21]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$a \ [sd]$</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$c \ [sd]$</td>
<td>none</td>
</tr>
</tbody>
</table>

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 proportion of switches during the last 50 trials.

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. 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

---

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.
3.2. Comparison of the EV, PVL, and EV-PU Models

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 proportion of switches during the last 50 trials predicted by each model. The last column of Table 3.5 shows for each model and each choice pattern separately, the proportion 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 proportion of switches during the last 50 trials. This information is visualized by the last three columns of Figure 3.2. 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 3.4 and 3.5). 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 of the 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 3.2, Table 3.4 and 3.5).

Figure 3.2 and Table 3.5 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 3.4). 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 proportion of switches during the last 50 trials, it best approximates the observed mean proportion of choices among the three models (Figure 3.2, Tables 3.4 and 3.5). 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 good decks, strongly resemble the deck selection profiles of healthy participants (Figures 3.7 and 3.8).

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 proportion of switches lies in interquartile ranges of the observed proportions of switches (cf. Figure 3.2, Tables 3.4 and 3.5). This suggests that for very rare parameter combinations the PVL model meets both empirical regularities: the rank order of deck preferences and the proportion of switches. This is illustrated by Figures 3.9 and 3.11 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 3.9). This is supported by Figure 3.2 and Tables 3.4 and 3.5 showing that, for choice patterns featuring a pronounced preference from the bad decks, the largest generated proportion 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
3.2. Comparison of the EV, PVL, and EV-PU Models

Figure 3.3: 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 3.4: Deck selection profiles of four synthetic participants showing a preference for the good decks (generated by the EV model; \( w = 0.42, a = 0.03, c = 5.00 \)).
3. A Comparison of Reinforcement-Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning

Figure 3.5: 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 3.6: Deck selection profiles of four synthetic participants showing a preference for the good decks (generated by the EV-PU model; $A = 0.81, w = 0.34, a = 0.24, c = 5.00$).
3.2. Comparison of the EV, PVL, and EV-PU Models

Figure 3.7: Deck selection profiles of four healthy participants showing a preference for the bad decks.

Figure 3.8: Deck selection profiles of four synthetic participants showing a preference for the bad decks (generated by the EV model; $w = 0.00, a = 0.25, c = 1.61$).
3. A Comparison of Reinforcement-Learning Models for the Iowa Gambling Task Using Parameter Space Partitioning

Figure 3.9: 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 3.10: Deck selection profiles of four healthy participants showing a preference for the decks with infrequent losses.
3.2. Comparison of the EV, PVL, and EV-PU Models

Figure 3.11: Deck selection profiles of four synthetic participants showing a preference for the decks with infrequent losses (generated by the PVL model; $A = 0.07, w = 0.25, a = 0.88, c = 0.34$).

Figure 3.12: Deck selection profiles of four synthetic participants showing a preference for the decks with infrequent losses (generated by the EV-PU model; $A = 0.03, w = 0.08, a = 0.36, c = 4.66$).
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 proportion of switches. This is illustrated by Figures 3.6 and 3.12 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, i.e., a pronounced preference for good decks \((C + D) \geq 0.65\), a pronounced preference for bad decks \((A + B) \geq 0.65\), or a pronounced preference for decks with infrequent losses \((B + D) \geq 0.65\) (Table 3.4). 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 3.5). 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.

### 3.3 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 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,
i.e., 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, researches can decide which model is most appropriate. For instance, our findings suggest that the EV model should not be used if the data 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 (i.e., decks B and D) over decks with frequent losses (i.e., decks 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 using 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 conceded 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 non-human-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, Wetzels, Horstmann, 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 et al., 2005; Toplak et al., 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, Wetzels, Horstmann, 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; W. 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; S. A. 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; Prenkunmar et al., 2008; Sevy et al., 2006, 2007; Stout et al., 2004; van den Bos et al., 2000; 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, i.e., mean proportions of choices from each deck and mean proportion 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 et al. (2005) for providing the data used in this article. This research was supported by a Netherlands Organisation for Scientific Research (NWO) grant to HS (404-10-086).
Chapter 4

Validating the PVL-Delta Model for the Iowa Gambling Task

This chapter has been published as:
Validating the PVL-Delta model for the Iowa gambling task.
Frontiers in Decision Neuroscience, 4, 898.¹

Abstract

Decision-making deficits in clinical populations are often assessed with the Iowa gambling task (IGT). Performance on this task is driven by latent psychological processes, the assessment of which requires an analysis using cognitive models. Two popular examples of such models are the Expectancy Valence (EV) and Prospect Valence Learning (PVL) models. These models have recently been subjected to sophisticated procedures of model checking, spawning a hybrid version of the EV and PVL models—the PVL-Delta model. In order to test the validity of the PVL-Delta model we present a parameter space partitioning (PSP) study and a test of selective influence. The PSP study allows one to assess the choice patterns that the PVL-Delta model generates across its entire parameter space. The PSP study revealed that the model accounts for empirical choice patterns featuring a preference for the good decks or the decks with infrequent losses; however, the model fails to account for empirical choice patterns featuring a preference for the bad decks. The test of selective influence investigates the effectiveness of experimental manipulations designed to target only a single model parameter. This test showed that the manipulations were successful for all but one parameters. To conclude, despite a few shortcomings, the PVL-Delta model seems to be a better IGT model than the popular EV and PVL models.

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 to isolate and identify the psychological processes that drive performance on the IGT, behavioral analyses of IGT data are insufficient. A promising alternative analysis approach is to use cognitive

¹The final publication is available at http://dx.doi.org/10.3389/fpsyg.2013.00898.
process models. The IGT imposes high demands on these models because it is a complex task producing various types of choice patterns that a good model should be able to generate (Steingroever, Wetzels, Horstmann, et al., 2013; Steingroever, Wetzels, & Wagenmakers, 2013a). In addition, the models should also account for individual differences and for participants’ switch behavior on the task (e.g., Steingroever, Wetzels, & Wagenmakers, 2013a; Zhao & Costello, 2007). Despite the high demands, some plausible and elegant IGT models have been proposed. Two of the most frequently used representatives include the Expectancy Valence model (EV; see Steingroever, Wetzels, & Wagenmakers, 2013a for references), and the Prospect Valence Learning model (PVL; see Steingroever, Wetzels, & Wagenmakers, 2013a for references and a detailed description of the models). The parameters of these models correspond to distinct psychological processes such as motivation, learning/memory, and response consistency (Busemeyer, Stout, & Finn, 2003).

Since the development of the EV model in 2002, reinforcement-learning (RL) models for IGT data have been subjected to sophisticated procedures of model checking (e.g., Ahn et al., 2008; Busemeyer & Stout, 2002; Fridberg et al., 2010; Steingroever, Wetzels, & Wagenmakers, 2013a; Yechiam & Busemeyer, 2005; Yechiam & Ert, 2007; Yechiam & Busemeyer, 2008). These model comparison efforts spawned a hybrid version of the EV and PVL models—the PVL-Delta model (Ahn et al., 2008; Fridberg et al., 2010; Steingroever, Wetzels, & Wagenmakers, 2014; see next section for a detailed description of the PVL-Delta model and recent model comparison efforts). This model seems to be promising for IGT data because it can generate a variety of empirical choice patterns better than its competitors (Steingroever et al., 2014).

Whereas previous procedures of model checking focused mostly on relative comparisons of different RL models for IGT data, no efforts have been carried out to validate the PVL-Delta model (i.e., assess its adequacy in isolation). Here, we focus on two different ways of validating the PVL-Delta model: First, we conduct a parameter space partitioning (PSP) study that systematically assesses which choice patterns the PVL-Delta model generates across its entire parameter space. Thus, with this first validity check we aim to answer the question: Can the PVL-Delta model generate typical empirical choice patterns over a wide range of parameter settings? Second, we conduct a test of selective influence that investigates the effectiveness of experimental manipulations designed to target only one of the model parameters. Thus, with this second validity check we aim to answer the question: Do the parameters of the PVL-Delta model indeed correspond to the proposed psychological processes?

The outline of this article is as follows. In the first section, we explain the IGT, outline the PVL-Delta model, and review previous efforts to compare RL models for IGT data. In the second and third section, we present the PSP study and the test of selective influence. In the last section, we summarize our findings and discuss their ramifications. To anticipate our results, our PSP study shows that the PVL-Delta model can account for empirical choice patterns featuring a preference for the good decks or the decks with infrequent losses; however, the model fails to account for empirical choice patterns featuring a preference for the bad decks. Our test of selective influence shows that the manipulations were successful for all but one parameters.

4.1 The Iowa Gambling Task and the PVL-Delta Model

The Iowa Gambling Task

In this section we describe the IGT (see also Steingroever, Wetzels, & Wagenmakers, 2013a; 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. Participants are instructed to choose
4.1. The Iowa Gambling Task and the PVL-Delta Model

Table 4.1: Payoff scheme of the traditional IGT as developed by Bechara et al. (1994).

<table>
<thead>
<tr>
<th></th>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deck</td>
<td>Bad</td>
<td>Bad</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>Loss trials</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Loss/10 cards</td>
<td>−1250</td>
<td>−1250</td>
<td>−250</td>
<td>−250</td>
</tr>
<tr>
<td>Net outcome/10 cards</td>
<td>−250</td>
<td>−250</td>
<td>250</td>
<td>250</td>
</tr>
</tbody>
</table>

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 C and D, whereas the bad, risky decks are labeled A and B. Table 4.1 presents the traditional payoff scheme as developed by Bechara et al. (1994). This table illustrates that decks A and B yield high immediate, constant rewards, but even higher unpredictable, occasional losses: hence, the long-term net outcome is negative. Decks C and D, on the other hand, yield low immediate, constant rewards, but even lower unpredictable, occasional losses: hence, the long-term net outcome is positive. In addition to the different payoff magnitudes, the decks also differ in the frequency of losses: Two decks yield frequent losses (decks A and C) and two decks yield infrequent losses (decks B and D).

The PVL-Delta Model

In this section, we describe the PVL-Delta model in detail. The model formalizes participants’ performance on the IGT through the interaction of four model parameters that represent distinct psychological processes (Ahn et al., 2008; Fridberg et al., 2010; Steingroever et al., 2014).

The first model assumption is that after choosing a card from deck $k \in \{1, 2, 3, 4\}$ on trial $t$, participants evaluate the net outcome associated with the just-chosen card by means of a non-linear utility function from prospect theory (Tversky & Kahneman, 1992)—the Prospect Utility function:

$$u_k(t) = \begin{cases} 
X(t)^A & \text{if } X(t) \geq 0 \\
-w \cdot |X(t)|^A & \text{if } X(t) < 0.
\end{cases}$$

(4.1)

Here $X(t)$ represents the net outcome on trial $t$, that is, the sum of the experienced reward and loss (i.e., $X(t) = W(t) - |L(t)|$). The Prospect Utility function contains the first two model parameters—the shape parameter $A \in [0, 1]$, that determines the shape of the utility function, and the loss aversion parameter $w \in [0, 5]$. 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 negative net outcomes than positive net outcomes on the subjective utility, whereas a value of $w$ approaching one indicates identical impact of negative
net outcomes and positive net outcomes. As $w$ approaches zero, the model predicts that negative net outcomes will be neglected.

The PVL-Delta model further assumes that, after having formed the utility of the just chosen deck through Equation 4.1, decision makers update their expected utility of the just chosen deck, while keeping the expected utilities of the remaining decks unchanged. This updating process is described by the Delta learning rule:

$$E_{vk}(t) = E_{vk}(t-1) + a \cdot (u_k(t) - E_{vk}(t-1)).$$ (4.2)

The Delta learning rule states that the expected utility of the chosen deck $k$ is adjusted upward if the experienced utility $u_k(t)$ is higher than expected. 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 third model parameter—the updating parameter $a \in [0, 1]$. 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.

In the next step, the model assumes that the expected utilities of each deck guide participants’ choices on the next trial $t+1$. This assumption is formalized by the softmax choice rule, also known as the ratio-of-strength choice rule. The PVL-Delta model uses this rule to compute the probability of choosing each deck on each trial (Luce 1959, Equation 4.3). This rule contains a sensitivity parameter $\theta$ that indexes the extent to which trial-by-trial choices match the expected deck utilities. Values of $\theta$ close to zero indicate random choice behavior (i.e., strong exploration), whereas large values of $\theta$ indicate choice behavior that is strongly determined by the expected utilities (i.e., strong exploitation).

$$P[S_k(t+1)] = \frac{e^{\theta E_{vk}(t)}}{\sum_{j=1}^{4} e^{\theta E_{vj}(t)}}$$ (4.3)

The PVL-Delta model assumes a trial-independent sensitivity parameter $\theta$, which depends on the final model parameter: the response consistency $c \in [0, 5]$ (Equation 4.4). Small values of $c$ cause a random choice pattern, whereas large values of $c$ cause a deterministic choice pattern.

$$\theta = 3^c - 1$$ (4.4)

In sum, the PVL-Delta 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 updating parameter $a$, which determines the memory for past expectancies, and (4) the response consistency parameter $c$, which determines the amount of exploitation versus exploration.

**Previous Comparisons of RL Models**

This section reviews previous model comparison studies. These studies compared the EV model, PVL model, and alternative RL models using a large variety of methods, for instance: the post hoc fit criterion (i.e., Ahn et al. 2008; Busemeyer & Stout 2002; Fridberg et al. 2010; Yechiam & Busemeyer 2005; Yechiam & Ert 2007; Yechiam & Busemeyer 2008; 3 the simulation method (i.e., Ahn et al. 2008; Fridberg et al. 2010; Steingroever et al. 2014; Worthy, Hawthorne, & Otto...)

---

2We initialized the expectancies of each deck $k$ to zero, $E_{vk}(0) = 0$.

3The post hoc fit criterion is also known as the one-step-ahead prediction method.
4.2 Parameter Space Partitioning

Method

We performed a PSP study to evaluate the flexibility of the PVL-Delta model (Pitt et al., 2006, Pitt et al., 2008; see also Steingroever, Wetzels, & Wagenmakers, 2013a) who performed a PSP study of the EV model, PVL model, and another hybrid model: the EV model with Prospect Utility function). The PSP method systematically assesses the choice patterns predicted by the PVL-Delta model across its entire parameter space. 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

4 Note that the PSP study of Steingroever, Wetzels, and Wagenmakers (2013a) did not focus on the PVL-Delta model, but on the EV model, the PVL model, and another hybrid model: the EV model with the Prospect Utility function.

5 The Bernoulli 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.
possible, but never observed. Instead, one should prefer a less flexible, parsimonious model that–ideally–only generates choice patterns that are also frequently observed in experiments (Pitt et al., 2006, 2008).

Note that PSP is a global method (i.e., the full range of parameter values is considered), whereas the other methods that were used to compare RL models are local (i.e., assessment at a particular point in the model’s parameter space; for instance, post hoc fit criterion, simulation method, tests of generalizability, and tests of parameter consistency). The advantage of global methods is that they enable one to assess the full range of choice patterns a model can generate, whereas the results of local methods always depend on the idiosyncrasies of any single data set (Pitt et al., 2006, 2008).

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 (see also Steingroever, Wetzels, & Wagenmakers, 2013a): For each parameter of the PVL-Delta model, we chose 60 values that were equally spaced over the corresponding parameter range. Each combination of these parameter values was used to generate data for 100 synthetic participants completing a 100-trial IGT. For all analyses in this paper, we scaled the traditional payoffs of the IGT as presented in Table 4.1 by dividing by 100 (cf. Ahn et al., 2011).

The generated data were used to analyze which choice patterns the PVL-Delta model can generate across its entire parameter space. Such analysis naturally requires a definition of choice patterns. Here we used two different definitions—the “broad definition of choice patterns” and the “restricted definition of choice patterns”. These definitions are the same as used by Steingroever, Wetzels, and Wagenmakers (2013a).

Broad definition of choice patterns

The “broad definition of choice patterns” is intended to provide a general idea of which choice patterns the PVL-Delta model can generate. Following Steingroever, Wetzels, and Wagenmakers (2013a), we defined five possible choice patterns: (1) Preference for the good decks over bad decks (i.e., \( \{C, D\} \succ \{A, B\} \)), (2) preference for the bad decks over good decks (i.e., \( \{A, B\} \succ \{C, D\} \)), (3) preference for the decks with infrequent losses over decks with frequent losses (i.e., \( \{B, D\} \succ \{A, C\} \)), (4) preference for the decks with frequent losses over decks with infrequent losses (i.e., \( \{A, C\} \succ \{B, D\} \)), and (5) remaining choice patterns. For each parameter combination, we computed the proportion of choices from each deck averaged across all 100 trials and all 100 repeated data generations. These average choice proportions were then sorted to determine the generated rank order of deck preferences for each parameter combination. Finally, we computed the proportion of the entire parameter space occupied by each of the defined choice patterns. 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, Wetzels, Horstmann, et al., 2013) that a good model for IGT data should only generate the first three types of choice patterns.

Restricted definition of choice patterns

Note that the broad 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 does not matter whether the PVL-Delta model generated, for example, a very strong or a very weak preference for the good decks over the bad decks. Both generated choice patterns are classified as the choice pattern “good decks over bad
Table 4.2: 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 (see Table 4 in Steingroever, Wetzels, & Wagenmakers, 2013a). Healthy participants are selected from the Steingroever, Wetzels, Horstmann, et al. (2013) data pool (N = 394).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(C + D) ≥ .65</td>
<td>54</td>
<td>.10 [.05]</td>
<td>.14 [.05]</td>
<td>.36 [.17]</td>
<td>.40 [.14]</td>
<td>.35 [.08, .52]</td>
</tr>
<tr>
<td>(A + B) ≥ .65</td>
<td>18</td>
<td>.25 [.07]</td>
<td>.52 [.11]</td>
<td>.11 [.05]</td>
<td>.12 [.06]</td>
<td>.43 [.31, .58]</td>
</tr>
<tr>
<td>(B + D) ≥ .65</td>
<td>139</td>
<td>.12 [.05]</td>
<td>.37 [.12]</td>
<td>.13 [.05]</td>
<td>.39 [.12]</td>
<td>.47 [.28, .66]</td>
</tr>
</tbody>
</table>

decks” (i.e., \{C, D\} > \{A, B\}). To go beyond this coarse classification, we also analyzed the model’s behavior when confronted with pronounced deck preferences. To get an indication of pronounced deck preferences shown by healthy participants on the IGT, we used Steingroever, Wetzels, and Wagenmakers (2013a)’s definition of pronounced deck preferences: Specifically, Steingroever, Wetzels, and Wagenmakers (2013a) searched their IGT data pool (N = 394; Steingroever, Wetzels, Horstmann, et al., 2013) for healthy participants that chose at least 65% cards from either the good decks (i.e., (C + D) ≥ .65), the bad decks (i.e., (A + B) ≥ .65), or the decks with infrequent losses (i.e., (B + D) ≥ .65). By using the .65-criterion, Steingroever, Wetzels, and Wagenmakers (2013a) included healthy participants with pronounced deck preferences and excluded healthy participants with random choice behaviors. For each of these three groups, Steingroever, Wetzels, and Wagenmakers (2013a) computed the mean proportions of choices from each deck (as shown in Table 4.2). For instance, participants classified to the group “pronounced preference for the good decks” chose on average 36 cards from deck C and 40 cards from deck D. Note that 53.6% of all participants in the Steingroever, Wetzels, Horstmann, et al. (2013) data pool showed a pronounced deck preference by making at least 65% choices from the two most preferred decks. This empirical popularity of pronounced deck preferences underscores how important it is that a RL model for the IGT is able to produce such choice patterns.

Table 4.2 thus provides an indication of pronounced deck preferences shown by healthy participants on the IGT. We used the mean proportion of choices from these three constructed groups for our second, restricted definition of choice patterns. Specifically, we define a pronounced preference for the good decks as at least 36 and 40 choices from decks C and D, respectively; we define a pronounced preference for the bad decks as at least 25 and 52 choices from decks A and B, respectively; and we define a pronounced preference for the decks with infrequent losses as at least 37 and 39 choices from decks B and D, respectively. Based on our simulations, we then determined the proportion of the parameter space of the PVL-Delta model that produced choice patterns that satisfy this second, restricted definition.

Switch behavior

Finally, a good RL model for the IGT should also capture the switches participants make on the IGT (Zhao & Costello, 2007). Steingroever, Wetzels, and Wagenmakers (2013a) therefore determined the mean proportion of switches during the last 50 trials for the three groups of healthy participants.
showing pronounced decks preferences (revisited here in the last column of Table 4.2). The table contains for each of the three groups of healthy participants with pronounced choice patterns the mean proportion of switches during the last 50 trials and statistics quantifying the distribution of switch proportions (i.e., the interquartile range and the minimum and maximum switch proportions during the last 50 trials). This information is visualized by the boxplots shown in the left column of Figure 4.1. From Table 4.2 and Figure 4.1 it is evident that, in general, in all three groups participants switch frequently. However, the interquartile ranges and the minimum and maximum proportion of switches during the last 50 trials also indicate that there is large variability in the proportion of switches, such that the switch behavior of healthy participants varies 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 4.2, 4.4, and 4.6 (see also Figures 3, 7, 10 in Steingroever, Wetzels, & Wagenmakers, 2013a) which show the trial-by-trial choices (i.e., choice profiles) of representative healthy participants with a pronounced preference for the good decks, bad decks, and decks with infrequent losses, respectively.

We investigated whether the PVL-Delta model captures the empirical switch behavior by comparing the empirical and generated mean proportions of switches during the last 50 trials. Specifically, the generated mean proportions of switches were obtained by determining the mean proportions of switches during the last 50 trials for all parameter combinations that produced pronounced deck preferences. The code for the PSP study is available on www.helensteingroever.com

Results

Broad definition of choice patterns

Table 4.3 presents the proportion of the parameter space of the PVL-Delta model occupied by each of the five different types of choice patterns. From this table, it is evident that the PVL-Delta model can generate all five different types of choice patterns. However, if we consider its partitioned parameter space more closely, we detect substantial differences between the popularity of the different choice patterns: The choice pattern “good decks over bad decks” is the most central to the model’s overall performance, as this choice pattern occupies the largest part of the model’s parameter space. The second and third largest part of its parameter space are occupied by the choice patterns “remaining” and “infrequent losses over frequent losses”. It is thus evident that choice patterns that are typically shown by healthy participants –the choice patterns “good decks over bad decks” and “infrequent losses over frequent losses” (Steingroever, Wetzels, Horstmann, et al., 2013)– occupy a major part of the model’s parameter space.

Table 4.3 also shows that the choice pattern “bad decks over good decks” is only generated over a minor part of the model’s parameter space. We have therefore grounds to conclude that this choice pattern is uncharacteristic of the PVL-Delta model, and is thus almost irrelevant to its overall performance (Pitt et al., 2006). This finding is important because the choice pattern “bad decks over good decks” is considered characteristic for participants with decision-making deficits (e.g., patients with lesions to the ventromedial prefrontal cortex; Bechara et al., 1994, 1997). These patients are thought to display decision-making deficits on the IGT because their inability to foresee the long-term consequences of their choice behavior leads them to only focus on the immediate rewards.

See Steingroever, Wetzels, and Wagenmakers (2013a) for the deck selection profiles of all healthy participants that showed a pronounced deck preference (i.e., at least 60% choices from the two most preferred decks).
4.2. Parameter Space Partitioning

Table 4.3: Proportions of choice patterns generated by the PVL-Delta model.

<table>
<thead>
<tr>
<th>Choice pattern</th>
<th>Proportion of all choice patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good ≻ bad decks</td>
<td>{C, D} ≻ {A, B} \ .596</td>
</tr>
<tr>
<td>Bad ≻ good decks</td>
<td>{A, B} ≻ {C, D} \ .006</td>
</tr>
<tr>
<td>Infr. ≻ frequent losses</td>
<td>{B, D} ≻ {A, C} \ .118</td>
</tr>
<tr>
<td>Frequent ≻ infr. losses</td>
<td>{A, C} ≻ {B, D} \ .005</td>
</tr>
<tr>
<td>Remaining</td>
<td>\ .274</td>
</tr>
</tbody>
</table>

Restricted definition of choice patterns

Table 4.4 presents the proportion of all choice patterns generated by the PVL-Delta model that satisfy the restricted definition of choice patterns. The table also presents the mean and standard deviation of the parameter combinations that generated these pronounced deck preferences. The table shows that only minor parts of the parameter space of the PVL-Delta model are occupied by the three types of pronounced choice patterns, even though these patterns are frequently observed in experiments. For instance, 139 healthy participants from the [Steingroever, Wetzels, Horstmann, et al. (2013)] data pool (35.3%) show a pronounced preference for the decks with infrequent losses (i.e., \((B + D) ≥ .65\)). However, the PVL-Delta model only generates this choice pattern over 1.6% of its parameter space.

Switch behavior

In addition to the generated choice proportions, we also determined the generated proportion of switches during the last 50 trials for all parameter combination that satisfy the restricted definition of choice patterns (Columns 2 – 6 of Table 4.4). We averaged these generated switch proportions separately for each of the three types of pronounced deck preferences (last column of Table 4.4).

Table 4.4: Proportion of choice patterns generated by the PVL-Delta model that satisfy the restricted definition of choice patterns. Note that this definition is only based on the mean proportion of choices of the two strongest preferred decks (first column). For the selected choice patterns, the corresponding mean and standard deviation of the model parameters, and the mean proportion of switches during the last 50 trials are presented.

<table>
<thead>
<tr>
<th>Choice pattern</th>
<th>Proportion of all choice patterns</th>
<th>A [sd]</th>
<th>w [sd]</th>
<th>a [sd]</th>
<th>c [sd]</th>
<th>Switches during the last 50 trials [25%, 75% quantile] (min, max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C ≥ .36, D ≥ .40 )</td>
<td>\ .0084</td>
<td>\ .66 \ [.21]</td>
<td>\ .62 \ [.38]</td>
<td>\ .30 \ [.21]</td>
<td>\ .349 \ [.98]</td>
<td>\ .0571 \ [.0014, .0558] \ (.00, .5724)</td>
</tr>
<tr>
<td>( A ≥ .25, B ≥ .52 )</td>
<td>\ .0000028</td>
<td>\ .92 \ [.09]</td>
<td>\ .02 \ [.03]</td>
<td>\ .06 \ [.03]</td>
<td>\ .307 \ [.40]</td>
<td>\ .0043 \ [.0003, .0055] \ (.00, .0210)</td>
</tr>
<tr>
<td>( B ≥ .37, D ≥ .39 )</td>
<td>\ .0162</td>
<td>\ .27 \ [.24]</td>
<td>\ .34 \ [.39]</td>
<td>\ .43 \ [.27]</td>
<td>\ .280 \ [.88]</td>
<td>\ .0705 \ [.0034, .0918] \ (.00, .6450)</td>
</tr>
</tbody>
</table>

63
4. Validating the PVL-Delta Model for the Iowa Gambling Task

The table also contains statistics quantifying the distribution of the generated switch proportions, that is, the interquartile range and the minimum and maximum proportion of switches during the last 50 trials. This information is visualized by the right column of Figure 4.1.

When comparing the generated and observed mean proportion of switches during the last 50 trials given pronounced deck preferences, it is apparent that the PVL-Delta model underestimates the observed switch proportions, that is, the generated mean proportion of switches equals or falls below .07 for all generated pronounced choice patterns, whereas the observed mean proportion of switches equals or exceeds .35 for all observed pronounced choice patterns (Tables 4.2 and 4.4). In addition, for all three types of pronounced choice patterns, the interquartile range of the observed proportion of switches exceeds the interquartile range of the model-generated proportion of switches (Figure 4.1, Table 4.2, and 4.4). However, the largest generated switch proportion given a pronounced preference for the good decks and the decks with infrequent losses, respectively, lie within the corresponding interquartile ranges of the observed switch proportions. This suggests that for a few parameter combinations, the PVL-Delta model meets both empirical regularities—pronounced deck preferences and a tendency to switch frequently.

To illustrate the differences and commonalities between the data and the predictions, we plot in Figures 4.2-4.7 observed and generated choice profiles. Figures 4.2, 4.4, and 4.6 show the choice profiles of representative healthy participants with a pronounced preference for the good decks, bad decks, and decks with infrequent losses, respectively. Figures 4.3, 4.5, and 4.7 show the choice profiles that were generated with those parameter combinations that resulted in a pronounced preference for the good decks, bad decks, and decks with infrequent losses, respectively, and the maximum number of switches during the last 50 trials. From the figures it is evident that there are large discrepancies between the observed and generated choice profiles in the case of the pronounced preference for the bad decks: The PVL-Delta model generates a few switches in the beginning of the IGT and then exploitation of a single deck, even though healthy participants keep switching across the entire IGT. However, the observed and generated choice profiles look very similar in the case of the pronounced preference for the good decks and the decks with infrequent losses.

To conclude, many healthy participants from the Steingroever, Wetzels, Horstmann, et al. (2013) data pool (53.6%) showed pronounced deck preferences, that is, a pronounced preference for the good decks \((C + D) \geq .65\), a pronounced preference for the bad decks \((A + B) \geq .65\), or a pronounced preference for the decks with infrequent losses \((B + D) \geq .65\) (Table 4.2). This empirical popularity of pronounced deck preferences is only partly reflected by the PVL-Delta model; the model produces choice patterns that satisfy the restricted definition of choice patterns only within minor parts of its parameter space (Table 4.3). In addition, healthy participants in general show many switches during the last 50 trials. However, the PVL-Delta model in general predicts that participants who show pronounced deck preferences switch rarely during the last 50 trials; all generated mean proportion of switches during the last 50 trials equal or fall below .07 whereas the observed mean proportions of switches lie around .40. But compared to the popular EV and PVL models (Steingroever, Wetzels, & Wagenmakers, 2013a), the PVL-Delta model performs better: The PVL-Delta model generates higher mean proportions of switches than its two competitors for almost all pronounced choice patterns; the only exception is that the EV model generates a higher mean proportion of switches for the choice pattern featuring a pronounced preference for the bad decks than the PVL-Delta model.

Moreover, healthy participants show large individual differences in the proportion of switches during the last 50 trials, such that their switch behavior varies between no switches at all to switches on every trial. However, the PVL-Delta model tends to generate very few switches, given pronounced deck preferences, and fails to generate large proportion of switches (i.e., switch proportions higher than .65). But compared to the popular EV and PVL models (Steingroever,
Figure 4.1: Boxplots of observed and generated proportions of switches during the last 50 trials, given a pronounced deck preference. Each row presents the results for different pronounced 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 from the Steingroever, Wetzels, Horstmann, et al. (2013) data pool (cf. Table 4.2). The second column presents the switches generated by the PVL-Delta model (cf. Table 4.4).
4. Validating the PVL-Delta Model for the Iowa Gambling Task

Figure 4.2: Deck selection profiles of four healthy participants showing a pronounced preference for the good decks. The filled dots indicate the occurrence of rewards and losses together; the empty dots indicate the occurrence of only rewards.

Figure 4.3: Deck selection profiles of four synthetic participants showing a pronounced preference for the good decks (generated by the PVL-Delta model; $A = 0.88, w = 0.68, a = 0.25, c = 1.27$).
4.2. Parameter Space Partitioning

Figure 4.4: Deck selection profiles of four healthy participants showing a pronounced preference for the bad decks.

Figure 4.5: Deck selection profiles of four synthetic participants showing a pronounced preference for the bad decks (generated by the PVL-Delta model; $A = 1.00, w = 0.08, a = 0.05, c = 2.71$).
4. Validating the PVL-Delta Model for the Iowa Gambling Task

Figure 4.6: Deck selection profiles of four healthy participants showing a pronounced preference for the decks with infrequent losses.

Figure 4.7: Deck selection profiles of four synthetic participants showing a pronounced preference for the decks with infrequent losses (generated by the PVL-Delta model; $A = 0.00, w = 0.00, a = 0.42, c = 1.19$).
the PVL-Delta model again performs better because the EV and PVL model’s failure to generate large proportions of switches, given a pronounced choice pattern, is even stronger: Given a pronounced choice pattern, the EV and PVL models fail to generate switch proportions higher than .35 and .46, respectively. Despite these discrepancies between the empirical and the generated switch behavior, we showed that –given a pronounced preference for the good decks or the decks with infrequent losses and those parameter combinations that yielded the maximum number of switches during the last 50 trials– the PVL-Delta model can produce choice patterns that strongly resemble the empirical choice patterns of healthy participants.

4.3 Test of Selective Influence

In this section we investigate whether the parameters of the PVL-Delta model indeed correspond to distinct psychological processes. We will therefore carry out a test of selective influence for the PVL-Delta model. This means that we fit the model to data collected from the standard IGT, but also from conditions that were designed to affect selectively one of the model parameters. These data were collected by Wetzels, Vandekerckhove, et al. (2010), and their experiment was originally designed as a test of selective influence for the EV model. However, the experimental manipulations that were intended to affect the parameters of the EV model should also be reflected by the parameters of the PVL-Delta model because of the high similarity between the two models.

Method

We fit the PVL-Delta model separately to four data sets reported by Wetzels, Vandekerckhove, et al. (2010). Specifically, Wetzels, Vandekerckhove, et al. (2010) conducted an experiment with a standard condition and three additional conditions that were designed to affect selectively one of the model parameters:

- **In the “standard condition”, 19 participants completed a 150-trial IGT under the standard administration.**
- **In the “rewards condition”, 20 participants completed a 150-trial IGT under the instruction to pay more attention to rewards and to consider losses as less important.** We expected this manipulation to decrease the loss aversion parameter $w$.

- **In the “updating condition”, 19 participants completed a 150-trial IGT under the standard administration. However, each choice was followed by a on-screen presentation of five numbers that the participants had to remember because, after the next choice, participants were asked about the relative position of one of the numbers. We expected this manipulation to increase the updating parameter $a$.**

- **In the “consistency condition”, 16 participants completed a 150-trial IGT under the standard administration. However, they were told that after every 10 trials the payoff schemes for the decks could have changed. We expected this manipulation to decrease the consistency parameter $c$.**

To fit the PVL-Delta model, we used a Bayesian hierarchical approach detailed in the next section. This estimation procedure has been consistently shown to outperform alternatives such as maximum likelihood estimation and Bayesian individual estimation (Ahn et al., 2011; Wetzels, Vandekerckhove, et al., 2010).

To assess whether the chains of all parameters had converged successfully from their starting values to their stationary distributions, we visually inspected the Hamiltonian Monte Carlo (HMC) chains and used 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 indicate inadequate convergence.

To assess model performance in absolute terms, we used two different methods: the post hoc absolute fit method and the simulation method (see also Steingroever et al., 2014). These two methods allow us to assess the model’s ability to fit and generate the choice patterns present in each of the four conditions. Our implementation of both methods relies on visually contrasting–separately for each deck as a function of 15 bins each containing 10 trials–the observed mean choice proportions from the experiment against the mean choice probabilities from the model.

Both methods start by sampling parameter values from the joint posterior distributions over the individual-level parameters (hereafter individual-level joint posteriors). In the case of the post hoc absolute fit method, the model is provided with the sampled parameter values, but also with the actual choices and payoffs of each participant. The post hoc absolute fit method computes the probability of choosing each deck on the next trial based on the information on the observed choices and payoffs up to and including the current trial. The simulation method, on the other hand, is only provided with the sampled parameter values, and relies on generating choices for another sequence of payoffs that could have been observed. In particular, on each trial, the simulation method generates a choice based on the predicted choice probabilities. For both methods and for each participant, we repeated the process of obtaining the predicted choice probabilities 100 times to account for uncertainty in the individual-level joint posteriors (for detailed recipes see Steingroever et al., 2014).

To investigate the effect of the experimental manipulations, we visually compared the posterior distributions of the group-level parameters of all four conditions.

Bayesian hierarchical estimation procedure

To fit the PVL-Delta model to the data of the four experimental conditions, we used a Bayesian hierarchical estimation procedure (see Wetzels, Vandekerckhove, et al. (2010) for the same model specification in the case of the EV model). The Bayesian graphical PVL-Delta model for a hierarchical analysis is shown in Figure 4.8. This figure shows that the graphical model consists of two plates: The inner plate expresses the replications of the choices on \( t = 1, \ldots, T \) trials of the IGT, and the outer plate expresses the replications for \( i = 1, \ldots, N \) participants. For the sake of clarity, we omitted the notation that indexes the deck number \( k \). The quantities \( W_{i,t} \) (rewards of participant \( i \) on trial \( t \)), \( L_{i,t} \) (losses of participant \( i \) on trial \( t \)), and \( Ch_{i,t+1} \) (choice of participant \( i \) on trial \( t + 1 \)) can directly be obtained from the data. The quantities \( u_{i,t}, Ev_{i,t+1}, \) and \( \theta_z \) are deterministic because they can be calculated from Equations 4.1, 4.2, and 4.4. All individual-level parameters \( z_i \), that is, \( \{A_i, w_i, a_i, c_i\} \), are also deterministic because instead of modeling the individual-level parameters directly, we modeled their respective probit transformations \( z_i' \), that is, \( \{A_i', w_i', a_i', c_i'\} \). This means that the parameters \( z_i' \) lie on the probit scale covering the entire real line. The probit transformation is the inverse of the cumulative standard normal distribution function. The parameters \( z_i' \) are assumed to be drawn from group-level normal distributions with mean \( \mu_z \) and standard deviation \( \sigma_z \). Only after the analysis was complete, we transformed the parameters \( \mu_z' \) and \( z_i' \) back to the original scale.

---

8 Note that we used the same payoff schedule as in the corresponding experiment.

9 For completeness, we also produced predicted choices based on the joint posterior of the group-level parameters (hereafter group-level joint posterior); that is, we generated data with 1,000 parameter values that were randomly drawn from the group-level joint posterior. There are slight differences between the two types of posterior predictives, but the general conclusions are the same (see Appendix C for further details).
4.3. Test of Selective Influence

Figure 4.8: Bayesian graphical PVL-Delta model for a hierarchical analysis. $\Phi()$ is the cumulative standard normal distribution function.

The model specification requires a definition of priors for the group-level means and standard deviations. We assigned a normal prior to the group-level means, $\mu_{z'} \sim N(0, 1)$, and a uniform prior to the group-level standard deviations, $\sigma_{z'} \sim U(0, 1.5)$.

We implemented the PVL-Delta model in Stan (Stan Development Team 2014b, 2014c; Hoffman & Gelman 2014). The code to fit the PVL-Delta model in Stan is available on http://www.helensteingroever.com. To confirm that we correctly implemented the PVL-Delta model, we ran several parameter-recovery studies. The results of two such studies are presented in Appendix C.

For each parameter, we ran three HMC chains simultaneously. The fitting procedure consisted of two steps: First, we initialized all chains with randomly generated starting values. We collected 1,000 samples of each chain after having discarded the first 9,000 samples of each chain as burn-in. However, this procedure did not result in successful convergence of the HMC chains of all parameters: for instance, for some parameters, two chains may appear to have converged to their stationary distributions and looked like “hairy caterpillars” that are randomly intermixed, whereas the third chain behaved differently and producing an inferior goodness of fit (GOF). Therefore, in
a second step, we again ran three HCM chains for each parameter, but this time, we initialized all chains with parameter values close to the mean of the HCM chain that produced the best GOF in the first step. However, even this procedure resulted in convergence problems for a few participants (e.g., bimodal posterior distributions). We therefore excluded participants with such convergence issues and repeated the first and second step. This explains why the sample sizes presented in Table 4.5 are slightly smaller than those reported by Wetzels, Vandekerckhove, et al. (2010).

Table 4.5 also presents, for each data set separately, the number of burn-in samples and posterior samples that we collected for each chain. These specifications differ across data sets to ensure that all chains reached convergence. We based our inferences on these posterior samples.

Table 4.5: Sample size of the four data sets and number of burn-in samples and posterior samples that we collected for each chain.

<table>
<thead>
<tr>
<th>Experimental condition</th>
<th>Sample size</th>
<th>Burn-in samples</th>
<th>Posterior samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>17</td>
<td>37,000</td>
<td>3,000</td>
</tr>
<tr>
<td>Rewards</td>
<td>19</td>
<td>30,000</td>
<td>3,000</td>
</tr>
<tr>
<td>Update</td>
<td>16</td>
<td>23,000</td>
<td>1,500</td>
</tr>
<tr>
<td>Consistency</td>
<td>15</td>
<td>18,000</td>
<td>2,000</td>
</tr>
</tbody>
</table>

Results

In this section, we discuss the results of the test of selective influence. We first focus on the behavioral level by describing the choice patterns observed in the four experimental conditions. Second, we focus on the level of the cognitive modeling analyses; we describe tests confirming that the posterior distributions converged successfully from their starting values to their stationary distributions. In addition, we show that the PVL-Delta model results in a satisfactory fit performance and simulation performance for the four conditions. Finally, we visually compare the posterior distributions of the group-level parameters of all four conditions to draw inferences about the effect of the experimental manipulations.

Behavioral data

The mean proportion of choices from each deck within 15 blocks each containing 10 trials as observed in the four experimental conditions reported by Wetzels, Vandekerckhove, et al. (2010) are presented in the first column of Figure 4.9. In the standard condition, participants learned to prefer good deck C over all remaining decks; however, participants failed to learn that deck D is also a good deck.

In the rewards condition (i.e., participants were instructed to pay more attention to rewards and to consider losses as less important), participants learned to prefer bad deck B over all remaining decks. Note that even though bad decks A and B both yield high immediate rewards on every trial, participants did not learn to select deck A more often than good decks C and D. This may suggest that the experimental manipulation was only partly successful.

In the updating condition (i.e., each choice was followed by a on-screen presentation of five numbers that participants had to remember because, after the next choice, they were asked about the relative position of one of the numbers), participants show a very weak learning curve; they only learned to avoid deck A.

In the consistency condition (i.e., participants were told that after every 10 trials the payoff schemes for the decks could have changed), participants –in contrast to the intention of the
Figure 4.9: Observed choice behavior and assessment of absolute model performance. The first column shows the mean proportion of choices from each deck within 15 blocks as observed in the four experimental conditions reported by [Wetzels, Vandekerckhove, et al. (2010)]. Each block contains 10 trials. The second and third column show the fit performance and simulation performance, respectively, for each of the four conditions. Fit performance and simulation performance are based on random draws from the individual-level joint posteriors.
4. Validating the PVL-Delta Model for the Iowa Gambling Task

experimental manipulation– did not evenly explore all decks across the entire 100 trials. Instead participants learned to prefer decks B and C over the remaining decks. It seems that participants prefer bad deck B because it yields high immediate rewards on the majority of the trials; however, participants prefer good deck C because it never yields a net loss and is therefore a safe option.

Convergence checks

Visual inspection of the HMC chains and consideration of the ˆR statistics for all parameters (all parameters had ˆR values below 1.045) suggest that all chains have converged successfully. To illustrate how we visually assessed convergence, we show the chains of one individual-level parameter in Appendix C. From the figure it is evident that the chains have converged successfully from their starting values to their stationary distribution, looking like “hairy caterpillars” that are randomly intermixed.

Absolute model performance

To assess the absolute model performance of the PVL-Delta model with respect to the four experimental conditions, the second and third column of Figure 4.9 show the fit performance and simulation performance, respectively. Fit performance and simulation performance are based on random draws from the individual-level joint posterior. From the second column of the figure it is evident that the PVL-Delta model provides a good fit to the data of all four conditions (i.e., the model makes accurate one-step-ahead predictions when provided with access to the observed sequence of choices and payoffs). In addition, the third column of Figure 4.9 illustrates that the PVL-Delta model adequately generates the choice pattern shown by the standard and update conditions. In the case of the rewards and consistency conditions, the simulation performance of the PVL-Delta model is acceptable; the model correctly predicts the most preferred deck, but fails to account for the rank order of the remaining three decks: In the reward condition, the model predicts that deck D is preferred over decks A and C even though the participants chose these three decks about equally often. In the consistency condition, the model predicts that deck D is preferred over deck C even though the participants showed the reverse pattern. To sum up, the PVL-Delta model captures the global patterns in the data providing an acceptable fit and simulation performance with respect to the four data sets at hand; this allows us to meaningfully compare the group-level parameters of the four conditions.

Test of selective influence

Figure 4.10 presents the posterior distributions of the group-level parameters of all four conditions. It is evident that the experimental manipulation is successfully reflected in the loss aversion parameter and the consistency parameter: First, compared to participants that received the standard instruction, participants who were instructed to focus on rewards (i.e., the rewards condition) had lower values for the loss aversion parameter indicating that they were indeed more reward-seeking. Second, fitting the PVL-Delta model to data of participants that were told that after every 10 trials the payoff schemes for the decks could have changed (i.e., the consistency condition) resulted in a smaller consistency parameter (i.e., a more random choice behavior) than fitting the PVL-Delta model to data of participants that received the standard instructions. However, in the update condition is no clear effect on the updating parameter. Yet, it is evident that the consistency parameter in the update condition is noticeably lower than in the standard condition (i.e., a more random choice behavior); this is consistent with the choice pattern.
4.3. Test of Selective Influence

Figure 4.10: Posterior distributions for the group-level parameters of the PVL-Delta model in the four experimental conditions.
shown by the update condition; participants only learned to avoid deck A, but show a completely
indistinguishable preference for the remaining three decks.

4.4 Discussion

In this article, we conducted two tests to validate the PVL-Delta model: a parameter space
partitioning study and a test of selective influence. Applying PSP to the PVL-Delta model, we
have obtained a deeper understanding of the model’s behavior. 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 model’s overall performance, and the restricted definition allowed us
to assess the model’s data-fitting potential when confronted with data featuring pronounced deck
preferences.

Using the broad definition of choice patterns, the PSP study revealed that the PVL-Delta
model can generate all typical empirical choice patterns. However, the PVL-Delta model generates
the choice pattern featuring a preference for the bad decks only over a minor part of its parameter
space suggesting that this choice pattern is virtually irrelevant to the model’s overall performance.

Using the restricted definition of choice patterns, the PSP study revealed that the PVL-Delta
model can still generate all pronounced empirical choice patterns over a minor part of its parameter
space. But for these pronounced choice patterns, the PVL-Delta model generally underestimates
the empirical switch proportions during the last 50 trials. In particular, given pronounced
preferences for the bad decks, the PVL-Delta model fails to account for the empirical switch
behavior. This failure seems to be caused by the Prospect Utility function of the PVL-Delta
model: In a previous PSP study, Steingroever, Wetzels, & Wagenmakers [2013a] showed that
this failure is also present in the PVL and EV-PU model (i.e., models with the Prospect Utility
function), but not in the EV model (i.e., a model without the Prospect Utility function). However,
in the case of the other two pronounced choice patterns –the choice patterns favoring decks with
high expected value or low loss frequency– we showed that the PVL-Delta model provides a good
account for the empirical switch behavior for some parameter combinations.

The results of the PSP study for the PVL-Delta model and the earlier PSP studies for the
EV and PVL models [Steingroever, Wetzels, & Wagenmakers, 2013a] suggest that the PVL-Delta
model outperforms its two competitors. The EV model fails to generate a pronounced preference for
the decks with infrequent losses; the PVL model is able to generate pronounced decks preferences,
but underestimates the switch proportions even more strongly than the PVL-Delta model. This
superiority of the PVL-Delta model is in line with the posterior predictive checks reported by
Steingroever et al. [2014].

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 the
data-fitting potential of the PVL-Delta model. Thus, if researchers wish to apply the PVL-Delta
model to IGT data, they can decide based on the behavioral results whether it is appropriate to
apply the PVL-Delta model or not.

The PSP results of this paper should be interpreted with care. PSP gives an indication of
how central choice patterns are to the overall performance of the model. However, it is premature
to conclude that the PVL-Delta model cannot generate the choice pattern “bad decks over goods

76
decks” at all, solely because the model generates this choice pattern over a small part of the parameter space. Instead, we can only conclude that this choice pattern is not central to the model’s overall performance.

It should also be noted that the inferences drawn from the PSP study strongly depend on our definitions of choice patterns. The restricted definition of choice patterns was based on IGT performance of healthy participants (Steingroever, Wetzels, & Wagenmakers, 2013a). We could thus detect inconsistencies between the empirical popularity of each pronounced choice pattern in the Steingroever, Wetzels, Horstmann, et al. (2013) data pool and the frequency predicted by the PVL-Delta model. It is troubling that the PVL-Delta model fails to generate a pronounced preference for the bad decks with many switches. But it should be acknowledged that this choice pattern is not central in healthy participants’ IGT performance: in the Steingroever, Wetzels, Horstmann, et al. (2013) data pool, only 5% ($N = 18$) of the healthy participants showed this choice pattern (Table 4.2). Still, this choice pattern is assumed to be characteristic for patients with decision-making deficits (Bechara et al., 1994, 1997), but a better empirical foundation (e.g., a literature review on the IGT performance of clinical groups) is required to accurately judge the gravity of the PVL-Delta model’s failure to generate a pronounced preference for the bad decks with many switches.

The test of selective influence revealed that the experimental manipulations had a noticeable effect on the loss aversion parameter and consistency parameter, but not on the updating parameter. However, it is premature to conclude that the updating parameter does not correspond to memory processes. It may be that the experimental manipulation did not work out properly. In addition, one should bear in mind that every data set is characterized by its own idiosyncrasies. IGT data generally are highly idiosyncratic—possibly because the IGT is a very complex task (Steingroever, Wetzels, Horstmann, et al., 2013). In order to be able to draw more accurate conclusions on whether the parameters represent distinct psychological processes, independent repetitions of the test of selective influence and even different experimental manipulations are necessary.

The results of this article confirm that the PVL-Delta model is an attractive alternative to the popular EV and PVL models. However, the PVL-Delta model is also characterized by a few shortcomings because it underrepresents the choice pattern featuring a preference for the bad decks. Nevertheless, we recommend that researchers use the PVL-Delta model to disentangle psychological processes underlying IGT performance, provided that they rigorously assess absolute model performance before interpreting the model parameters (Steingroever et al., 2014).

Acknowledgement

This publication was supported by the Dutch national program COMMIT, by an Open Access grant from NWO, and by the Netherlands Organisation for Scientific Research (NWO; 404-10-086).
Chapter 5

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

This chapter has been published as:
Absolute performance of reinforcement-learning models for the Iowa gambling task.
Decision, 1, 161–183.

Abstract

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

The final publication is available at http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=2013-44393-001.
The Iowa gambling task (IGT; Bechara et al., 1994) is arguably the most popular neuropsychological paradigm to assess decision-making deficits in clinical populations. Originally, the IGT was developed to assess decision-making deficits of patients with lesions to the ventromedial prefrontal cortex (vmPFC), but in the last two decades the task has been applied to a variety of clinical populations, such as patients with Asperger’s disorder (e.g., S. A. Johnson et al., 2006), attention-deficit-hyperactivity disorder (e.g., Agay et al., 2010; Toplak et al., 2005), bipolar disorder (e.g., Brambilla et al., 2012), obsessive-compulsive disorder (e.g., Cavedini, Riboldi, D’Annucci, et al., 2002), pathological gambling disorder (e.g., Cavedini, Riboldi, Keller, et al., 2002), psychopathic tendencies (e.g., Blair et al., 2001), and schizophrenia (e.g., Martino et al., 2007; Premkumar et al., 2008). In addition, the IGT has been applied to cocaine addicts (e.g., Stout et al., 2004), chronic cannabis users (e.g., Fridberg et al., 2010), heavy drinkers (e.g., Gullo & Stieger, 2011), inmates (e.g., Yechiam, Kanz, et al., 2008), and traffic offenders (e.g., Lev et al., 2008). Impaired performance on the IGT may be caused by several factors, such as only focusing on immediate rewards, avoidance of immediate losses, poor memory for past payoffs, or underweighting of rare events (e.g., Barron & Erev, 2003; Yechiam & Busemeyer, 2005).

In order to isolate and identify the psychological processes that drive performance on the IGT, behavioral analyses of IGT data need to be complemented with cognitive modeling analyses. To further this goal, several reinforcement-learning (RL) models have been proposed, and here we focus on the two most popular exemplars—the Expectancy Valence model (EV; Busemeyer & Stout, 2002) and the Prospect Valence Learning model (PVL; Ahn et al., 2008, 2011)—and the hybrid PVL-Delta model (Ahn et al., 2008; Fridberg et al., 2010; a detailed description of the three models can be found in the section “The Ev, PVL, and PVL-Delta Models”). The parameters of these models correspond to psychological processes such as motivation, learning/memory, and response consistency (Busemeyer et al., 2003); hence, the purpose of fitting these models to empirical data is to allow applied researchers to draw conclusions about the latent psychological processes that drive performance on the IGT. For instance, fit the EV model to data of 10 groups of people suffering from various neuropsychological disorders (e.g., Asperger’s syndrome, vmPFC lesions, chronic cannabis abuse), and mapped these groups according to the differences between their model parameters and those of their control group. The purpose of this analysis was to characterize the decision-making deficits of each clinical group in terms of underlying psychological processes, and to examine whether differences in neuropsychological disorders can be explained by differences in psychological processes underlying decision-making deficits (i.e., differences in the model parameters).

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

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

To anticipate our main result, the post hoc absolute fit method revealed that all models provided
5.1. The Iowa Gambling Task and Three Reinforcement-Learning Models

The Iowa Gambling Task

In this section we describe the IGT (see also Steingroever, Wetzels, & Wagenmakers, 2013a). 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 5.1 presents the traditional payoff scheme as developed by Bechara et al. (1994). This table illustrates that decks A and B yield high immediate, constant

<table>
<thead>
<tr>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad deck with frequent losses</td>
<td>Bad deck with infrequent losses</td>
<td>Good deck with frequent losses</td>
<td>Good deck with infrequent losses</td>
</tr>
<tr>
<td>Reward/trial</td>
<td>100</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Number of losses/10 cards</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Loss/10 cards</td>
<td>−1250</td>
<td>−1250</td>
<td>−250</td>
</tr>
<tr>
<td>Net outcome/10 cards</td>
<td>−250</td>
<td>−250</td>
<td>250</td>
</tr>
</tbody>
</table>

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

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

5.1 The Iowa Gambling Task and Three Reinforcement-Learning Models

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

5.1 The Iowa Gambling Task and Three Reinforcement-Learning Models

The Iowa Gambling Task

In this section we describe the IGT (see also Steingroever, Wetzels, & Wagenmakers, 2013a). 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 5.1 presents the traditional payoff scheme as developed by Bechara et al. (1994). This table illustrates that decks A and B yield high immediate, constant

<table>
<thead>
<tr>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad deck with frequent losses</td>
<td>Bad deck with infrequent losses</td>
<td>Good deck with frequent losses</td>
<td>Good deck with infrequent losses</td>
</tr>
<tr>
<td>Reward/trial</td>
<td>100</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Number of losses/10 cards</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Loss/10 cards</td>
<td>−1250</td>
<td>−1250</td>
<td>−250</td>
</tr>
<tr>
<td>Net outcome/10 cards</td>
<td>−250</td>
<td>−250</td>
<td>250</td>
</tr>
</tbody>
</table>
5. Absolute Performance of Reinforcement-Learning Models for the Iowa Gambling Task

Table 5.2: Formalization of the EV, PVL, and PVL-Delta models.

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

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.

rewards, but even higher unpredictable, occasional losses: hence, the long-term net outcome is negative. Decks C and D, on the other hand, yield low immediate, constant rewards, but even lower unpredictable, occasional losses: hence, the long-term net outcome is positive. In addition to the different payoff magnitudes, the decks also differ in the frequency of losses: Two decks yield frequent losses (decks A and C) and two decks yield infrequent losses (decks B and D).

The EV, PVL, and PVL-Delta Models

In this section, we describe the EV, PVL, and PVL-Delta models (see also [Steingroever, Wetzels, & Wagenmakers] 2013a). Table 5.2 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 EV, PVL, and PVL-Delta models share the assumption that, following each choice, participants evaluate the rewards and losses (if any) associated with the just-chosen card by means of a utility function. These momentary utilities are used to update expectancies about the utilities of all decks. This updating process entails that, on every trial, participants adjust their expected utilities of the decks based on the new utility they just experienced, a process described by a learning rule. In the next step, the models assume that the expected utilities of all decks are used to guide
5.1. The Iowa Gambling Task and Three Reinforcement-Learning Models

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 \( \text{Luce, 1959} \). This rule contains a sensitivity parameter \( \theta(t) \) that indexes the extent to which trial-by-trial choices match the expected deck utilities. Values of \( \theta(t) \) close to zero indicate a random choice behavior (i.e., strong exploration), whereas large values of \( \theta(t) \) indicate a choice behavior that is strongly determined by the expected deck utilities (i.e., strong exploitation). As is customary, for all analyses in this paper, we scaled the traditional payoffs of the IGT as presented in Table 5.1 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. The first model assumption is that after choosing a card from deck \( k \), \( k \in \{1, 2, 3, 4\} \) on trial \( t \), participants compute a weighted mean of the experienced rewards \( W(t) \) and losses \( L(t) \) to obtain the utility of deck \( k \) on trial \( t \), \( u_k(t) \). The weight that participants assign to losses relative to rewards is the attention weight parameter \( w \). A small value of \( w \), that is, \( w < .5 \), is characteristic for decision makers who put more weight on the immediate rewards and can thus be described as reward-seeking, whereas a large value of \( w \), that is, \( w > .5 \), is characteristic for decision makers who put more weight on the immediate losses and can thus be described as loss-averse (Ahn et al., 2008; Busemeyer & Stout, 2002).

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

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

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

The PVL model

The PVL model uses four parameters to formalize its assumptions about participants’ performance on the IGT. 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, \( \delta_k \), 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.

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

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

The PVL-Delta model

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

Methods to Assess Performance of RL Models

In this section, we review methods that previous studies have used to assess performance of RL models. We differentiate between applied studies (i.e., studies that fit an RL model to IGT data to compare model parameters across groups) and model comparison studies (i.e., studies that search for the best performing model among a set of competitor models). First of all, it is evident that many applied studies take model adequacy for granted; only about two thirds of the applied literature assessed model performance at all (Steingroever, Wetzels, & Wagenmakers, 2013a). The
5.1. The Iowa Gambling Task and Three Reinforcement-Learning Models

The standard measure to assess model performance has been the conventional fit index BIC or $G^2$. This index is a relative measure that compares the performance of two models (i.e., the accuracy of one-step-ahead predictions when provided with intermediate feedback on the observed choices and payoffs); the first model is an RL model that aims to explain trial-to-trial dependencies and learning effects; the second model is a baseline model that assumes constant choice probabilities across all trials (equal to the individual’s overall choice proportions from each deck). This method is also called post hoc fit criterion or one-step-ahead prediction method (see for example, Farah et al., 2008; Yechiam et al., 2005; Yechiam, Kanz, et al., 2008). We call this measure post hoc relative fit criterion in the remainder of the article to stress the comparison against a baseline model.

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

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

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

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

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

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

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

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

It may be argued that Yechiam and Busemeyer (2005)’s implementation of the generalizability test is rather a simulation method because the two tasks resemble each other strongly (i.e., both tasks used an implementation of the IGT, but the payoff scheme used in the second task differs by a constant factor of 1.5 from the payoff scheme of the
Data set dependency has also been confirmed by the PSP study of Steingroever, Wetzels, and Wagenmakers (2013a) showing that the EV model, PVL model, and a modified version of the EV model (i.e., the EV model with Prospect Utility function) all fail to generate the entire spectrum of choice patterns that are typically observed in experiments. In particular, the EV model fails to generate a pronounced preference for the decks with infrequent losses (see also Fridberg et al., 2010; Yechiam & Busemeyer, 2005)—a choice pattern that is often observed in healthy participants (e.g., Caroselli et al., 2006; Dunn et al., 2006; MacPherson et al., 2002; Lin et al., 2007; Steingroever, Wetzels, Horstmann, et al., 2013; Wilder et al., 1998; Yechiam & Busemeyer, 2005). Such a dependency of the models’ performance on the observed choice pattern presents a crucial limitation because a good RL model for the IGT should be able to generate choice patterns present in all groups that are typically tested on the IGT.

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

5.2 Performance of the EV, PVL, and PVL-Delta Models

Methods

We fit the EV, PVL, and PVL-Delta models using a Bayesian hierarchical estimation procedure (detailed in the next section) to two data sets that were constructed from our IGT data pool of healthy participants (Steingroever, Wetzels, Horstmann, et al., 2013). For the first data set, we selected 31 healthy participants with a pronounced preference for the good decks (i.e., participants with at least 75% choices from the good decks, $(C + D) \geq .75$); for the second data set, we selected 31 healthy participants with a pronounced preference for the decks with infrequent losses (i.e., participants with at least 75% choices from the decks with infrequent losses, $(B + D) \geq .75$). All participants completed a 100-trial IGT.

We chose these two types of choice patterns because the first type is in line with Bechara et al. (1994)’s assumptions about the performance of healthy participants on the IGT. The second type goes against Bechara et al. (1994)’s assumptions, but it is frequently observed in healthy participants (see for example, Caroselli et al., 2006; Chiu & Lin, 2007; Chiu et al., 2008; Dunn et al., 2006; Fridberg et al., 2010; MacPherson et al., 2002; Lin et al., 2007; Steingroever, Wetzels, Horstmann, et al., 2013; Wilder et al., 1998; Yechiam & Busemeyer, 2005). Since healthy first task). However, the question whether Yechiam and Busemeyer (2005)’s method should be classified as a test of generalizability or as a simulation method does not affect our point that results of previous model comparison studies may depend on the data set analyzed because Ahn et al. (2008) showed that models with Delta-RL rule performed better than models with Decay-RL rule on both the simulation method and test of long-term generalizability.

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

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

87
participants are typically used as a control group, it is important that the models can account for these two types of choice patterns.

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

To visualize the representativeness of our two stylized data sets, Figure 5.1 displays the proportions of choice patterns shown by participants in our IGT data pool considered for this article (total \(N = 359\)). For this figure, we defined five different types of choice patterns based on the deck rank order: (1) Preference for the decks with infrequent losses (i.e., \(\{B,D\} \succ \{A,C\}\)); (2) preference for the good decks (i.e., \(\{C,D\} \succ \{A,B\}\)); (3) preference for the bad decks (i.e., \(\{A,B\} \succ \{C,D\}\)); (4) preference for the decks with frequent losses (i.e., \(\{A,C\} \succ \{B,D\}\)); (5) remaining choice patterns. From the figure it is evident that the choice patterns “preference for the decks with infrequent losses” and “preference for the good decks” are most central in our IGT data pool considered for this article. Even though we chose an arbitrary cutoff value of .75 to construct the two groups, Figure 5.1 suggests that the construction of the two groups is empirically well founded (i.e., 32.6\% \((N = 31)\) and 38.3\% \((N = 57)\) of the participants with a preference for the decks with infrequent losses and for the good decks, respectively, show a pronounced deck preference).

To assess the models’ performance in absolute terms, we used two different methods: the post hoc absolute fit method and the simulation method. These two methods allow us to assess the models’ ability to fit and generate the choice patterns present in the two stylized data sets. Our implementation of both methods relies on visually contrasting –separately for each deck as a function of 10 bins– the observed mean choice proportions from the experiment against the mean choice probabilities from a particular model. For the data sets at hand a visual inspection is
5.2. Performance of the EV, PVL, and PVL-Delta Models

sufficient; a more formal approach is provided by posterior predictive p-values (Gelman, Meng, & Stern, 1996; Meng, 1994 but see Bayarri & Berger, 1999, 2000).

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

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

Our implementation of the post hoc absolute fit method and the simulation method differ from previously used tests in the following ways. First, our post hoc absolute fit method entails an absolute comparison to the data instead of a relative comparison to a baseline model (i.e., the post hoc relative fit; but see Busemeyer & Stout, 2002; Wood et al., 2005, for an absolute presentation featuring only the good decks). Second, our implementation of the simulation method considers the choice probabilities for each deck and trial separately, instead of averaging across all trials (see Ahn et al., 2008; Fridberg et al., 2010; Worthy, Hawthorne, & Otto, 2013). Yechiam and Busemeyer (2005) also used this implementation, but they did not account for uncertainty in the parameter estimates.

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

---

6Note that we used the same payoff schedule as in the corresponding experiment.

7Note that we did not fit the models to the data sets of Fernie and Tunney (2006), Rodriguez-Sánchez et al. (2005), and Toplak et al. (2005) because we received their data only in bins of several trials or because we did not receive information on the payoff of each participant. The data set labeled as “own data set” in Steingroever, Wetzels, Horstmann, et al. (2013) is here labeled as “Horstmann” because Annette Horstmann collected the data.
Bayesian hierarchical estimation procedure

To fit the EV, PVL, and PVL-Delta models to the data, we used a Bayesian hierarchical estimation procedure (see Ahn et al., 2011; Wetzels, Vandekerckhove, et al., 2010, for advantages of the Bayesian hierarchical approach). The Bayesian graphical PVL (and PVL-Delta) model for a hierarchical analysis is shown in Figure 5.2. The Bayesian graphical EV model looks very similar; the only difference is that the EV model has one fewer parameter, that parameter \( w_i \) is immediately drawn from a group-level distribution instead of being obtained from \( w'_i \), and that the sensitivity parameter is trial-dependent (i.e., \( \theta_{t,i} \)). Figure 5.2 shows that the graphical model consists of two plates: The inner plate expresses the replications of the choices on \( t = 1, \ldots, T \) trials of the IGT, and the outer plate expresses the replications for \( i = 1, \ldots, N \) participants. For the sake of clarity, we omitted the notation that indexes the deck number \( k \). The quantities \( W_{i,t} \) (rewards of participant \( i \) on trial \( t \)), \( L_{i,t} \) (losses of participant \( i \) on trial \( t \)), and \( Ch_{i,t+1} \) (choice of participant \( i \) on trial \( t + 1 \)) can directly be obtained from the data, and the quantities \( u_{i,t} \), \( Ev_{i,t+1} \), and \( \theta_i \) can be calculated with the equations presented in Table 5.2. Each individual-level parameter vector \( z_i \), that is \( \{w_i, A_i, a_i, c_i\} \) in the case of the PVL and PVL-Delta models, and \( \{w_i, a_i, c_i\} \) in the case of the EV model, is assumed to be drawn from a group-level beta distribution, Beta(\( \alpha_z, \beta_z \)). Since beta distributions are restricted to the \([0, 1]\) interval, we transformed parameters with different ranges (see Table 5.2) to the \([0, 1]\) interval, and only transformed them back to their correct ranges after the analysis was complete. Beta distributions are typically defined by two shape parameters \( \alpha \) and \( \beta \). Here we reparameterize the two shape parameters in terms of the group-level mean \( \mu_z \) and group-level precision \( \lambda_z \) as follows:

\[
\alpha_z = \mu_z \lambda_z
\]

\[
\beta_z = \lambda_z (1 - \mu_z)
\]

We assigned a uniform prior to the group-level means, \( \mu_z \sim U(0, 1) \), and to the logarithm of the group-level precisions, \( \log(\lambda_z) \sim U(\log(2), \log(600)) \). Setting the lower limit of the prior on \( \log(\lambda_z) \) to \( \log(2) \) prevents the beta group-level distributions from being bimodal (Beta distribution, 2013, January 17). However, to prevent numerical problems in the estimation program we had to increase this lower limit to a maximum of 21 for the most challenging stylized data set and to a maximum of 31 for the most challenging complete data set—a modification that can reduce the variance of the group-level distributions; here this increase had little effect on our inferences as the posterior distributions of the group-level precision parameters were not cut off at their lower limit.

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

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

We initialized all chains with different starting values that were generated from uniform distributions covering a wide range of possible parameter values (i.e., randomly overdispersed starting values). Fitting the PVL and PVL-Delta models with three chains often resulted in convergence difficulties: for instance, two chains may appear to have converged to their stationary distributions and gave the appearance of “hairy caterpillars” that are randomly intermixed, whereas the third chain behaved differently, often seemingly stuck at either the lower or upper parameter bound and consequently producing a much larger deviance (i.e., an inferior goodness of fit). In such situations we decided to run at least five chains simultaneously and to base inferences on three chains with the smallest deviance. However, even this procedure resulted in convergence problems.

\footnote{Our convergence difficulties with the PVL and PVL-Delta models are not unique. \cite{Ahn} made available online two alternative fitting routines for the PVL model, and also reported convergence difficulties for their first code. They propose two solutions for the convergence difficulties: The first solution is the same as we proposed here, that is, basing inferences on chains that have converged successfully. The second solution is to use their second code that uses a different prior specification and model formulation in which the individual-level parameters are assumed to be drawn from truncated normal distributions. Also, \cite{Ahn} needed a large amount of}
for a few participants (e.g., bimodal posterior distributions). We therefore excluded participants with such convergence issues and repeated the fitting procedure. This explains why the sample sizes presented in Table 5.3 are slightly smaller than stated earlier and than those reported by Steingroever, Wetzels, Horstmann, et al. (2013).

Table 5.3 also contains, for each data set separately, the number of samples we discarded as burn-in and the number of posterior samples that we collected for each chain. These specifications differ across data sets to ensure that all chains reached convergence. We based our inferences on these posterior samples.

Table 5.3: Sample size of the two stylized and five complete data sets, number of samples discarded as burn-in, and number of posterior samples collected for each chain.

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

Results

Visual inspection of the MCMC chains and consideration of the $\hat{R}$ statistics for all parameters suggested that all chains converged successfully (i.e., all parameters of the two stylized data sets and complete data sets had $\hat{R}$ values below 1.04 and 1.05, respectively). To illustrate how we assessed convergence visually, Figure 5.3 shows the chains of one individual-level parameter. From the figure it is evident that the chains have converged to their stationary distribution, giving the appearance of “fat hairy caterpillars” that are randomly intermixed.

Ability to fit

Figure 5.4 presents the post hoc absolute fit of the three RL models with respect to the two stylized data sets. The first column presents the observed choice proportions from each deck as a function of 10 bins; the second, third, and fourth column present the mean probabilities of choosing each deck on each trial as calculated with the EV, PVL, and PVL-Delta models, respectively. The participants from the first data set show a pronounced preference for the good decks (first panel of the first row); the participants from the second data set show a pronounced preference for the decks with infrequent losses (first panel of the second row).

---

9 The deck selection profiles of all 61 participants included in the two stylized data sets can be downloaded here: [https://dl.dropbox.com/u/12798592/DeckSelectionProfilesFit.zip](https://dl.dropbox.com/u/12798592/DeckSelectionProfilesFit.zip). In these profiles, filled dots indicate the occurrence of rewards and losses together, whereas unfilled dots indicate the occurrence of rewards only.
5.2. Performance of the EV, PVL, and PVL-Delta Models

Figure 5.3: MCMC chains of the individual-level PVL parameter $c$ of the third participant in the stylized data set featuring a pronounced preference for the good decks. We inspected this type of plot for every parameter to assess convergence visually, in addition to quantifying convergence through the formal diagnostic measure $\hat{R}$.

It is evident from Figure 5.4 that all models provide an acceptable fit to the observed data. All models capture the qualitative choice patterns (i.e., the rank order of the decks) shown by the two stylized data sets. In addition, the models also adequately capture the size of the choice proportions. Nonetheless, the EV model seems to fit the two stylized data sets slightly worse than the PVL and PVL-Delta models: According to the EV model, it takes a few trials at the beginning of the IGT until participants start learning and develop the pronounced deck preferences—a pattern that is inconsistent with the observed data. But altogether, Figure 5.4 suggests that only small qualitative differences exist in the models’ ability to fit the two stylized data sets.

To ascertain that our results generalize to other data sets, Figure 5.5 shows the post hoc absolute fit performance of the three RL models with respect to the five complete data sets. The first column presents the observed choice proportions from each deck as a function of 10 bins; the second, third, and fourth column present the mean probabilities of choosing each deck on each trial as calculated with the EV, PVL, and PVL-Delta models, respectively. At the behavioral level, Figure 5.5 illustrates that only the data set of Premkumar et al. (2008) shows a preference for the good decks. The remaining four data sets show a frequency-of-losses effect (i.e., a preference for the decks with infrequent losses)—an effect that differs in its extent across the four data sets: The data sets of Fridberg et al. (2010) and Horstmann show a pronounced frequency-of-losses effect with a clear preference for both decks with infrequent losses (i.e., decks B and D) over both decks with frequent losses, whereas the remaining two data sets show a less pronounced frequency-of-losses effect, indicating that, at the end of the IGT, participants choose about equally often from decks.
Figure 5.4: Post hoc absolute fit performance of the three RL models with respect to the two stylized data sets. The first column presents the observed mean proportions of choices from each deck within 10 blocks. Each block contains 10 trials. The second, third, and fourth column present the mean probabilities of choosing each deck on each trial as calculated with the EV, PVL, and PVL-Delta model, respectively.

B, C, and D, while clearly avoiding deck A (i.e., [Kjome et al., 2010; Wood et al., 2005]). In general, it is evident that the choice patterns shown by the five complete data sets are less pronounced than those of the two stylized data sets presented in Figure 5.4.

It is evident that Figure 5.5 corroborates the conclusions from Figure 5.4: All models provide an acceptable fit to the data, but the EV model fits the five complete data sets slightly worse than the PVL and PVL-Delta models.

In addition to visually assessing the models’ ability to fit the two stylized and five complete data sets, we also compared the deviance measure provided by WinBUGS of all models and data sets (Table 5.4). The deviance is defined as \( D(\theta) = -2 \log p(y|\theta) \), where \( p(y|\theta) \) is the likelihood of the data \( y \) given the parameters \( \theta \). Thus, the smaller the deviance, the better the fit. In line with Figures 5.4 – 5.5, Table 5.4 shows that the EV model provides the worst fit to the two stylized and five complete data sets. In addition, the PVL model has a smaller deviance (i.e., a better fit) than the PVL-Delta model for five out of seven data sets.

Ability to generate

Figure 5.6 illustrates how the three models perform on the simulation method with respect to the two stylized data sets. The first column presents the observed choice proportions from each deck as a function of 10 bins (i.e., identical to the first column of Figure 5.4); the second, third, and fourth column present the probabilities of choosing each deck on each trial as generated with the
5.2. Performance of the EV, PVL, and PVL-Delta Models

Figure 5.5: Post hoc absolute fit performance of the three RL models with respect to the five complete data sets. The first column presents the observed mean proportions of choices from each deck within 10 blocks. Each block contains 10 trials, except the last block of Fridberg et al. (2010, 5-trials). The second, third, and fourth column present the mean probabilities of choosing each deck on each trial as calculated with the EV, PVL, and PVL-Delta model, respectively.
5. Absolute Performance of Reinforcement-Learning Models for the Iowa Gambling Task

Table 5.4: Deviance measures provided by WinBUGS of all models and data sets. For each data set, we printed in bold the lowest deviance value to identify the model with the best fit.

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

EV, PVL, and PVL-Delta models, respectively. From the figure it is evident that neither the EV nor the PVL model succeeds to generate both observed choice patterns. Specifically, the EV model fails to generate a choice pattern featuring a pronounced preference for the decks with infrequent losses (second stylized data set), whereas the PVL model fails to generate a choice pattern featuring a pronounced preference for the good decks (first stylized data set).

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

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

To ascertain that our results generalize to other data sets, Figure 5.7 shows how the three models perform on the simulation method with respect to the five complete data sets. The first column presents the observed choice proportions from each deck as a function of 10 bins (i.e., identical to the first column of Figure 5.5); the second, third, and fourth column present the probabilities of choosing each deck on each trial as generated with the EV, PVL, and PVL-Delta models, respectively.

It is evident that Figure 5.7 corroborates the conclusions from Figure 5.6. The EV model fails to generate choice patterns featuring a preference for the decks with infrequent losses as shown by four data sets; for these data sets the EV model makes almost random predictions. Interestingly, the EV model makes very similar predictions for the data sets of Fridberg et al. (2010) and Horstmann, and the data sets of Kjome et al. (2010) and Wood et al. (2005), even though there are pronounced differences in the choice patterns at the behavioral level. In the case of the choice pattern featuring
a preference for the good decks as shown by the data set of Premkumar et al. (2008), the EV model correctly predicts a preference for the good decks over the bad decks, but –as in the case of the stylized data set with a preference for the good decks– the EV model underestimates the preference for deck D and overestimates the preference for deck A.

As already suggested by Figure 5.6, Figure 5.7 underscores that the PVL model fails to generate a choice pattern featuring a preference for the good decks (i.e., as present in the data set of Premkumar et al., 2008). However, the PVL model makes acceptable predictions for the four data sets with a frequency-of-losses effect; for all four data sets the PVL model correctly generates that the decks with infrequent losses are preferred over the decks with frequent losses. Yet, it is evident that, for the data sets of Fridberg et al. (2010) and Wood et al. (2005), the PVL model fails to generate that deck D is on average preferred over deck B—a discrepancy between the data and the predictions that was already apparent in Figure 5.6. In addition, both Figures 5.6 and 5.7 illustrate that the PVL model generates learning curves that are relatively flat.

Moreover, as already suggested by Figure 5.6, the PVL-Delta model demonstrates adequate simulation performance: Figure 5.7 illustrates that the PVL-Delta model correctly generates the choice patterns shown by most data sets. Only in the case of the data set of Fridberg et al. (2010) does the PVL-Delta model slightly underestimate the mean choice proportions from deck D and slightly overestimate the choice proportions from deck C.

Overall, the results of this section showed that the simulation method—in contrast to the post hoc absolute fit method—allows for a good discrimination between the models. In addition, a
Figure 5.7: Simulation performance of the three RL models with respect to the two stylized data sets. The first column presents the observed mean proportions of choices from each deck within 10 blocks. Each block contains 10 trials, except the last block of Fridberg et al. (2010, 5-trials). The second, third, and fourth column present the mean probabilities of choosing each deck on each trial as generated with the EV, PVL, and PVL-Delta model, respectively.
5.3 General Discussion

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

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

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

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

It should be noted that, in this article, we did not rule out the possibility that other parameter
5. Absolute Performance of Reinforcement-Learning Models for the Iowa Gambling Task

combinations may result in better simulation performance: to assess simulation performance, we used parameter values that were obtained with a likelihood-based estimation procedure that optimizes the fit for the exact sequences of observed choices and payoffs. Nevertheless, at least in the case of the EV model, we can be certain that no matter which parameter combination we choose the EV model will never generate a frequency of losses effect (see the PSP study of Steingroever, Wetzels, & Wagenmakers, 2013a, and the simulation performance of the EV model reported by Fridberg et al., 2010; Worthy, Hawthorne, & Otto, 2013, and Yechiam & Busemeyer, 2005). However, instead of searching a model’s entire parameter space for those parameter values that produce the best simulation performance, it is conventional to assess simulation performance with parameter values obtained from a likelihood-based estimation procedure because researchers typically base their inferences on these parameter values when they wish to draw conclusions about psychological processes underlying performance on the IGT.

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

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

It stands to reason that model performance cannot be summarized with only one measure. Previous model comparison studies proposed other sophisticated and sound methods to assess model performance; in particular, to investigate whether a given model captures the underlying decision-making processes (e.g., parameter consistency, generalization to another task, test of
5.3. General Discussion

specific influence; see for example Ahn et al. (2008), Wetzels, Vandekerckhove, et al. (2010), Yechiam & Ert (2007), Yechiam & Busemeyer (2008). Even though we support these additional methods, they require data from another task and hence it may be not be realistic to advocate their use in applied work. Thus, we recommend applied researchers to choose a model based on results from previous model comparison studies that used these tests, and then to use the post hoc absolute fit method and the simulation method to assess absolute model performance. If both methods pass a minimum threshold of adequacy, we can be relatively confident that conclusions from model parameters are trustworthy.

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

Acknowledgements

We thank Annette Horstmann, Fridberg et al. (2010), Kjome et al. (2010), Premkumar et al. (2008), and Wood et al. (2005) for providing the data used in this article. We are grateful to Woo-Young Ahn, Jerome Busemeyer, and Eldad Yechiam for their constructive comments on a previous version. This research was supported by a Netherlands Organisation for Scientific Research (NWO) grant to HS (404-10-086).
So What? On the Meaning of Parameter Estimates from Reinforcement-Learning Models

In their comment on Steingroever et al. (SWW; 2014), Konstantinidis et al. (KSSAS; 2014) convincingly argue why a wide range of sophisticated model comparison methods is required to select a good model for the Iowa gambling task (IGT). While we agree with KSSAS on this count, the focus of SWW was not on model comparison. Here we clarify our initial goal, that is, to illustrate why assessment of absolute model performance is necessary to avoid premature conclusions about the psychological processes that drive performance on the IGT. In addition, we elaborate on the advantages and drawbacks of both the post hoc absolute fit method and the simulation method. Finally, we highlight the distinction between statistical aspects of model adequacy and psychological relevance of parameter estimates.

In their comment on Steingroever et al. (2014), Konstantinidis et al. (KSSAS; 2014) convincingly argue that the examination of reinforcement-learning (RL) models for the Iowa gambling task (IGT) is only at its starting point. In particular, KSSAS argue that, in model comparison, a uni-dimensional approach may be ill- advised. According to KSSAS, a more balanced and comprehensive assessment of model adequacy requires one to integrate results obtained from different model comparison methods. KSSAS also state that model comparison methods should take into account model complexity (see for example the generalization criterion; Busemeyer &

1The final publication is available at http://psycnet.apa.org/psycinfo/2015-12661-001/
6. On the Meaning of Parameter Estimates from RL Models

Wang (2000). In addition, KSSAS argue that more model comparison efforts are required to test the psychological meaningfulness of parameter estimates obtained from fitting RL models to IGT data. Altogether, KSSAS provide a focused summary of the current state of research on RL models for the IGT, while making crucial suggestions for future research.

However, KSSAS also express skepticism about the simulation method. According to KSSAS, it is premature to claim that the simulation method offers a more useful method of model discrimination than the post hoc fit method. In particular, KSSAS argue that “the use of parameters estimated with the post hoc fit method could result in a considerable underestimation of a model’s simulation performance” (p. 187). While we agree with this claim in general terms, it is important to note that the parameters obtained from the post hoc method are the ones that applied researchers base their conclusions on, and the validity of these conclusions was the focus of Steingroever et al. (2014).

In the remainder of this comment, we clarify the goal of Steingroever et al. (2014); we elaborate on the advantages and drawbacks of both the post hoc absolute fit method and the simulation method; and we highlight the distinction between statistical aspects of model adequacy and psychological relevance of parameter estimates.

6.1 Clarification of Our Initial Goal

RL models are applied to IGT data whenever researchers are interested in making inferences on the psychological processes that drive IGT performance (see Steingroever, Wetzels, & Wagenmakers, 2013a, for references). Unfortunately, systematic and detailed evaluations of model performance are virtually absent from the applied literature. Such evaluations, however, are necessary to decide whether or not the conclusions from the estimated model parameters are trustworthy. Consequently, estimated model parameters reported in applied studies might not be indicative of the psychological processes they seek to represent. In this context, Steingroever et al. (2014) pointed out that the conclusions from the model parameters are valid only when the applied model provides an adequate account of the observed data. This means that before drawing conclusions from the model parameters, researchers have to assess absolute model performance. It is therefore insufficient to report only relative assessment measures (i.e., the fit of a RL model compared to a baseline model or another RL model). In order to assess absolute model performance, Steingroever et al. (2014) focused on two methods: the post hoc absolute fit method and the simulation method. These methods allow one to assess whether the applied model provides an adequate account for the observed data in absolute terms. Only when this requirement has been fulfilled can the estimated parameters be mapped to the psychological processes of interest to any degree of confidence.

Steingroever et al. (2014) applied the post hoc absolute fit method and the simulation method to seven data sets and three different models. This comprehensive analysis revealed that in many data sets the neglect of absolute model performance can result in misleading conclusions about psychological processes. Thus, Steingroever et al. (2014)’s goal was to emphasize the importance of assessing absolute model performance before interpreting model parameters, and their goal was

---

2The parameter space partitioning study from Steingroever, Wetzels, & Wagenmakers (2013a) suggests that, for example, in the case of the PVL model and the stylized data set showing a pronounced preference for the good decks, other parameter combinations may produce better simulation performance.

3Good performance on other measures, such as 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, 2006; Yechiam & Busemeyer, 2008; see also section “Statistical Adequacy versus Psychological Relevance”) are also important to establish the psychological relevance of parameter estimates.
6.1. Clarification of Our Initial Goal

Figure 6.1: Anscombe’s Quartet. Four different data sets that are identical on summary statistics (i.e., means and variances of the two variables \(x\) and \(y\), the linear regression coefficient, and the Bayes factor, Jeffreys, [1961]). However, the visual presentation of the data sets suggests that they differ strongly, and that the statistical model assuming that there is a linear relationship between the two variables is only valid for the data from the upper left panel. Hence, only in the upper left panel can the model parameter \(\rho\) be interpreted in a meaningful way.

neither to identify a good comparison method for RL models nor to identify a good RL model for the IGT.

A compelling illustration of why it is important to assess absolute model performance prior to the interpretation of model parameters is given by Anscombe’s quartet (Anscombe, [1973], Andraszewicz et al., [2015], presented in Figure 6.1). Each panel shows a different data set consisting of eleven \((x, y)\) pairs of observations. The four data sets are constructed in such a way that the two variables \(x\) and \(y\) have the same means and variances, and that the linear regression coefficient and Bayes factor [Jeffreys, 1961] are identical. The Bayes factor of 23 indicates a strong support for the presence of a linear association in all four data sets. However, the visual presentation of the data sets suggests that the statistical model assuming that there is a linear relationship between the two variables is only valid in the upper left panel. Hence, only in the upper left panel can the
model parameter $\rho$ be interpreted in a meaningful way.

**Current Way of Assessing Model Performance**

As pointed out in the previous section, systematic and detailed evaluations of IGT model performance are virtually absent from the applied literature. The studies that did assess model performance used various model selection criteria, such as Bayesian information criterion (BIC) and $G^2$ (Steingroever et al., 2014). However, the disadvantage of these criteria is that they are relative measures, and thus provide no information on whether a given model is able to provide an adequate account of the data.

To illustrate the risks and drawbacks of the current way of assessing model performance, consider the study of Yechiam et al. (2005). These authors fit the EV model to data of 10 groups of people suffering from various neuropsychological disorders, and mapped these groups according to the differences between their model parameters and those of their control group. The purpose of this analysis was to characterize the decision-making deficits of each clinical group in terms of underlying psychological processes, and to examine whether differences in neuropsychological disorders can be explained by differences in specific psychological processes that underlie performance on the IGT. Specifically, Yechiam et al. (2005) argued that their results suggest the presence of three different clusters. First, there is a cluster of young polydrug abusers and young alcohol abusers with parameter values similar to their respective control groups. Second, there is a cluster of chronic cocaine abusers, chronic cannabis abusers, Huntington’s patients, seniors, and patients with bilateral damage to the ventromedial prefrontal cortices, with parameter values indicating higher attention to gains and greater recency effects than their respective control groups. Finally, there is a cluster of Parkinson’s patients, Asperger’s patients, and patients with lesions in the right somatosensory and insular cortex, with parameter values indicating higher attention to losses than their respective control groups. Yechiam et al. (2005) justified the validity of the estimated model parameters by arguing that the EV model, from which the estimates were obtained, outperforms the baseline model for the majority of the participants (see Table 1 in Yechiam et al., 2005). We agree that it is important to establish the superiority of the EV model compared to the baseline model. However, the fact that the EV model outperforms the baseline model does not guarantee that the fit is adequate; it is possible that the EV model has a better fit than the baseline model, but nonetheless provides a poor fit to the data in absolute terms. And if a model provides a poor fit to the data, this suggests that the parameter estimates are non-interpretable and that they do not link to the psychological processes they seek to represent (cf. Anscombe’s Quartet shown in Figure 6.1). In sum the current procedure to validate the link between model parameters and psychological processes is based solely on assessment of relative model adequacy. Although we agree that relative model adequacy is of great importance, so is the assessment of absolute model adequacy.

**Additional Ways to Assess Model Performance**

In Steingroever et al. (2014) we illustrated why conclusions from model parameters may be misleading when these parameters are interpreted without having confirmed the adequacy of the

---

4This index compares the performance of two models (i.e., the accuracy of one-step-ahead predictions when provided with intermediate feedback on the observed choices and payoffs); the first model is an RL model that aims to explain trial-to-trial dependencies and learning effects; the second model is a baseline model that assumes constant choice probabilities across all trials (equal to the individual’s overall choice proportions from each deck; Busemeyer & Stout, 2002).
model through sensible checks. We proposed two such checks – the post hoc fit and simulation method – as straightforward, general, and easy-to-implement methods to assess absolute model performance. These two methods both assess absolute model performance by considering the data generated by the models relative to the data that were observed. We showed the relevance and feasibility of these methods by applying them to three different RL models and seven data sets. We argued how these methods shed light on which steps are necessary to avoid premature conclusions about the model parameters.

Two important advantages of these methods are that they assess absolute model performance (i.e., not relative to a baseline model or another RL model), and that they do not require data from a second experiment (in contrast to the generalization criterion method proposed by KSSAS). We consider good post hoc fit and simulation performance as minimum requirements that need to be fulfilled before one can claim that conclusions about the model parameters are meaningful. This does not mean that we believe relative model adequacy to be unimportant; both relative and absolute measures of model adequacy are needed to obtain a complete assessment of the extent to which parameter estimates are valid indicators of the associated psychological processes.

To illustrate the importance of assessing absolute model performance using the post hoc fit method and the simulation method, consider the data from Premkumar et al. (2008; left panel of Figure 6.2), and imagine that the authors had conducted a BIC comparison of the EV, PVL and PVL-Delta models, and that the PVL model was the winning model. Current practice dictates that the estimated parameters of the PVL model are then used to draw conclusions about the psychological processes underlying IGT performance. However, the following illustration shows that these conclusions might be invalid, and that two more steps need to be taken following the BIC assessment. In a second step, post hoc absolute fit performance needs to be assessed. This performance is shown in the middle panel of Figure 6.2. A comparison of the left and middle panels suggests that the PVL model makes accurate one step ahead predictions for that data set. Thus, the post hoc fit method can be used to determine whether, in addition to the better BIC score
of the PVL model relative to the EV and PVL-Delta model (i.e., first step), the PVL model has sufficient post hoc fit performance also in absolute terms.

However, these two steps are still not sufficient to conclude that the parameters are meaningful. In a third step, simulation performance needs to be assessed. This performance is shown in the right panel of Figure [6.2]. A comparison of the left and right panels suggests that, when we use the estimated parameters to generate data, the model predicts a qualitatively different choice pattern, that is, a preference for the decks with infrequent losses (i.e., bad deck B and good deck D), even though the participants in the observed data set show an overall preference for the good decks. Thus, even though the PVL model has sufficient relative and absolute post hoc fit performance, the estimated parameters may not be indicative of the psychological processes that drove the participants’ performance on the IGT—a conclusion based on the poor simulation performance of the PVL model.

6.2 Post Hoc Absolute Fit Method Versus Simulation Method

In their comment, KSSAS endorse the post hoc fit method. However, instead of providing arguments supporting the post hoc fit method, they mainly provide arguments against the simulation method. First, KSSAS state that the simulation method is a crude generalization test. We agree with KSSAS on this count; in Steingroever et al. [2014], we introduced the simulation method as the least demanding test of generalizability (compared to different implementations of the generalization criterion; see Ahn et al. 2008; Yechiam & Busemeyer 2005; Yechiam & Ert 2007; Yechiam & Busemeyer 2008).

Second, KSSAS state that the simulation method ignores model complexity. We also agree with KSSAS on this count. However, we never claimed otherwise – our focus was not on model selection, and we proposed neither the simulation method nor the post hoc fit method as a tool for model comparison (see above).

Third, KSSAS state that the simulation method ignores the fact that different models are required to fit different individuals. We also agree with KSSAS on this count. This is why we used a Bayesian hierarchical implementation of the RL models that naturally accounts for similarities and differences between individuals (Shiffrin et al. 2008). While we agree with KSSAS that a mixture model might be even more sensible, it should be noted that this is a valid criticism of all current model comparison approaches in the field of RL models because none of them use a mixture model implementation.

Fourth, KSSAS state that the simulation method, in contrast to the post hoc fit method, can favor models that are unlikely to generalize to new datasets. KSSAS illustrate this claim with a deterministic model that contains only one parameter, say \( \theta \). This parameter identifies a sequence of choices within the set of all possible choice sequences. However, we believe this example is flawed; the following illustration shows that, in contrast to KSSAS’s claim, the post hoc fit method would also favor this model. Here we consider only a two-trial IGT, but our conclusions generalize to any trial length. In a two-trial IGT, there are \( 4^2 \) possible choice sequences that are indexed by the parameter \( \theta \) (Table [6.1]).

The resulting payoffs are neglected because they do not influence the predictions of this model. Suppose that the observed choice sequence was (A, D). Consequently, the estimate for \( \theta \) is \( \hat{\theta} = 4 \). If we now apply the post hoc fit method using \( \hat{\theta} \), we obtain \( P[S_A(t = 1)] = .25 \) for the predicted choice probability on the first trial because all decks are equally likely on the first trial. For the predicted choice probability on the second trial, we obtain \( P[S_D(t = 2)] = 1 \). Note that, in the case of this deterministic model, this prediction is independent of the information on the choice.

108
Table 6.1: Illustration of the deterministic model proposed by Konstantinidis et al. All possible choice sequences are presented as a function of the model parameter $\theta$. Here we only consider a two-trial IGT, but our conclusion generalizes to any trial length.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>Choice sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A A</td>
</tr>
<tr>
<td>2</td>
<td>A B</td>
</tr>
<tr>
<td>3</td>
<td>A C</td>
</tr>
<tr>
<td>4</td>
<td>A D</td>
</tr>
<tr>
<td>5</td>
<td>B A</td>
</tr>
<tr>
<td>. .</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>D D</td>
</tr>
</tbody>
</table>

and payoff on the first trial. If we use a longer IGT, the predicted choice probabilities are always one for the chosen deck, and zero for the unchosen decks (except for the first trial; see Appendix D for a recipe of the post hoc fit method). Thus, the post hoc fit method would clearly favor this overly complex model.

KSSAS provided only one argument supporting the post hoc fit method, that is, that RL models should include choice inertia. However, choice inertia (also called choice perseveration) and choice mimicry refer to two different concepts. Choice inertia is a psychological process describing participants’ tendency to repeat previous choices (see for example [Worthy, Pang, & Byrne, 2013]). Choice mimicry, on the other hand, refers to a model’s ability to fine-tune its parameters to obtain an accurate fit to the exact sequences of observed payoffs and choices. This fine-tuning is facilitated by the fact that the post-hoc fit method uses the payoffs and choices twice; once to obtain the best-fitting parameters, and once to make one-step-ahead “predictions” given the payoffs and choices from all previous trials. Thus, it is possible to observe choice mimicry in the post hoc absolute fit performance of a model, even though the model does not incorporate choice inertia as one of the psychological processes underlying IGT performance.

In addition, one crucial drawback of model selection criteria, such as BIC and $G^2$ –criteria that are based on the post hoc fit method– is that these criteria do not fully account for model complexity. A more complete account for model complexity (i.e., number of parameters, functional form of the model, and extension of the parameter space; I. J. Myung & Pitt [1997]) is offered by the Bayes factor (Jeffreys [1961] see Steingroever, Wetzels, & Wagenmakers [2016] for a comparison of RL models using Bayes factors).

Finally, it is important to note that, in a strict sense, the post hoc fit method does not predict, but it post-dicts. This is because in a first step, we use the data (i.e., choices, wins, losses of a given participant) to estimate the model parameters. In a second step, we use the estimated parameters together with the choices, wins, losses observed up-to and including the current trial to “predict” the choice on the next trial. But this is a postdiction because we already used all of the data to estimate the model parameters, and thus use all information for the one-step-ahead “predictions”.

\[5\] See Wagenmakers, Grünwald, and Steyvers [2006] for a description of the accumulative prediction error—a method that, in contrast to the post hoc fit method, indeed assesses whether a model can predict the next choice given only the information from the previous trials.
6.3 Statistical Adequacy Versus Psychological Relevance

In cognitive modeling, it is important to differentiate between statistical aspects of model adequacy (e.g., good fit to the observed data and good predictions for new data), and psychological relevance of parameter estimates (e.g., good parameter recovery, and good performance on tests of selective influence). In other words, in a statistical sense a model can be adequate because it provides a good fit to the observed data or because it makes good predictions for the behavior in new environments. However, this statistical dimension of model adequacy is not necessarily related to the psychological dimension that considers the link between model parameters and psychological processes. For example, a model is psychologically inadequate if its parameters cannot be recovered, or if it performs poorly on a test of selective influence—a test that investigates whether experimental manipulations of certain psychological processes are reflected by the parameters that are sought to measure the corresponding psychological processes (e.g., Steingroever, Wetzels, & Wagenmakers, 2013b; Wetzels, Vandekerckhove, et al., 2010). It is therefore possible that a model has adequate statistical properties, but that the parameter recovery is relatively poor, compromising the model’s utility in applied settings (e.g., the Value-Plus-Perseveration model; Ahn et al., 2014; Worthy, Pang, & Byrne, 2013). On the other hand, a model with parameters that clearly link to distinct psychological processes does not necessarily need to be able to account for the finer details of performance in order to be practically useful. For example, the EZ diffusion model (Wagenmakers, van der Maas, & Grasman, 2007) cannot account for the fact that error responses are differently distributed than correct responses; nevertheless, the EZ diffusion model is useful for estimating model parameters, especially when sample size is low (van Ravenzwaaij & Oberauer, 2009).

6.4 Conclusion

To conclude, KSSAS provide a focused summary of the current state of model comparison efforts in the field of RL models for the IGT, and they make crucial suggestions for future efforts on finding a good model for the IGT. Even though we agree with KSSAS on these counts, the focus of Steingroever et al. (2014) was not on model comparison. But we focused on the assessment of absolute model performance. Our goal was to illustrate why applied researchers should carefully assess absolute model performance before they draw conclusions from the estimated parameters. In particular, premature conclusions can be avoided by carefully assessing both post hoc fit performance and simulation performance.
Bayes Factors for Reinforcement-Learning Models of the Iowa Gambling Task

This chapter has been published as:
Bayes factors for reinforcement-learning models of the Iowa gambling task.
*Decision, 3*, 115–131.

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

---

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); 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:

$$\text{BIC}_{M(.)} = -2 \log(L(.)_{(.)}) + k(.) \log(n),$$

where $L(.)_{(.)}$ is the maximum likelihood of model $M(.)_{(.)}$, $k(.)$ is the number of free parameters of model $M(.)_{(.)}$, 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 $-2 \log(L(.)_{(.)})$. 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 the context of RL models for the IGT, the post hoc fit criterion is also known as the one-step-ahead prediction method.

112
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. 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,

---

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, Bussemeyer & Stout, 2002, Yechiam & Bussemeyer, 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., 2009).
Table 7.1: Main characteristics of the payoff scheme of the traditional IGT as developed by Bechara et al. (1994).

<table>
<thead>
<tr>
<th></th>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bad deck</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with frequent losses</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Number of losses/10 cards</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Loss/10 cards</td>
<td>−1250</td>
<td>−1250</td>
<td>−250</td>
<td>−250</td>
</tr>
<tr>
<td>Net outcome/10 cards</td>
<td>−250</td>
<td>−250</td>
<td>250</td>
<td>250</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Model(s)</th>
<th>Model equation</th>
<th>Free parameters</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Utility</strong></td>
<td>EV</td>
<td>$u_k(t) = (1 - w) \cdot W(t) + w \cdot L(t)$</td>
<td>$w$: Attention weight</td>
<td>[0, 1]</td>
</tr>
<tr>
<td></td>
<td>PVL, PVL-Delta, &amp; VPP</td>
<td>$u_k(t) = \begin{cases} X(t)^A &amp; \text{if } X(t) \geq 0 \ -w \cdot</td>
<td>X(t)</td>
<td>^A &amp; \text{if } X(t) &lt; 0 \end{cases}$</td>
</tr>
<tr>
<td>Learning rule</td>
<td>EV, PVL-Delta, &amp; VPP</td>
<td>$Ev_k(t) = Ev_k(t - 1) + a \cdot (u_k(t) - Ev_k(t - 1))$</td>
<td>$a$: Updating</td>
<td>[0, 1]</td>
</tr>
<tr>
<td></td>
<td>PVL</td>
<td>$Ev_k(t) = a \cdot Ev_k(t - 1) + \delta_k(t) \cdot u_k(t)$</td>
<td>$a$: Recency</td>
<td>[0, 1]</td>
</tr>
<tr>
<td>Perseveration</td>
<td>VPP</td>
<td>$P_k(t) = \begin{cases} d \cdot P_k(t - 1) + \delta_k(t) \cdot \epsilon_{pos} &amp; \text{if } X(t) \geq 0 \ d \cdot P_k(t - 1) + \delta_k(t) \cdot \epsilon_{neg} &amp; \text{if } X(t) &lt; 0 \end{cases}$</td>
<td>$d$: Decay $\epsilon_{pos} \epsilon_{neg}$</td>
<td>[0, 1]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Ev_k(t) = w_{Ev} \cdot Ev_k(t) + (1 - w_{Ev}) \cdot P_k(t)$</td>
<td>$w_{Ev}$: Expectancy weight</td>
<td>[0, 1]</td>
</tr>
<tr>
<td>Choice rule</td>
<td>All</td>
<td>$Pr[S_k(t + 1)] = \frac{e^{\theta(t)Ev_k(t)}}{\sum_{j=1}^{4} e^{\theta(t)Ev_j(t)}}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>EV</td>
<td>$\theta(t) = (t/10)^c$</td>
<td>$c$: Consistency</td>
<td>[-2, 2]</td>
</tr>
<tr>
<td></td>
<td>PVL, PVL-Delta, &amp; VPP</td>
<td>$\theta(t) = 3^c - 1$</td>
<td>$c$: Consistency</td>
<td>[0, 5]</td>
</tr>
</tbody>
</table>

*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; Ahn et al., 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
7.2. The EV, PVL, PVL-Delta, and VPP Models

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, \( \delta_k \), 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.

The PVL model assumes a trial-independent sensitivity parameter \( \theta \), 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 \( \epsilon_{pos} \) or \( \epsilon_{neg} \) –depending on whether the net outcome on trial \( t \) is positive or negative— is added to the perseveration of deck \( k \). Positive values of \( \epsilon_{pos} \) or \( \epsilon_{neg} \) 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) $\epsilon_{\text{pos}}$ and (6) $\epsilon_{\text{neg}}$, which quantify the tendency to persevere given positive and negative net outcomes, respectively; (7) the weight of the expected utility $w_{\text{Ev}}$, 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, $M_1$ and $M_2$, 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 $p(M_1)/p(M_2)$ to posterior model odds $p(M_1 | y)/p(M_2 | y)$ brought about by the data $y$:

$$\frac{p(M_1 | y)}{p(M_2 | y)} = \frac{p(M_1)}{p(M_2)} \times \frac{m(y | M_1)}{m(y | M_2)} \tag{7.2}$$

Posterior model odds  Prior model odds  Bayes factor

Thus, the Bayes factor is the ratio of the marginal likelihoods of the two models: $BF_{12} = m(y | M_1)/m(y | M_2)$. The Bayes factor can range from zero to infinity. For ease of interpretation the Bayes factor is often transformed to the log scale, where $\log(BF_{12}) = 0$ indicates that observed data are equally likely to occur under both models, $\log(BF_{12}) = \log(10)$ indicates that the data are 10 times more likely to occur under model $M_1$ than model $M_2$, and $\log(BF_{12}) = \log(.10)$ indicates that the data are 10 times more likely to occur under model $M_2$ than model $M_1$. 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 $\log(BF_{12})$ between $\log(3)$ and $\log(10)$ is characterized as moderate evidence for $M_1$, whereas a $\log(BF_{12})$ between $\log(10)$ and $\log(30)$ is characterized as strong evidence for $M_1$.

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 $M_{(\cdot)}$ is obtained by integrating the likelihood over the prior:

$$m(y | M_{(\cdot)}) = \int p(y | \theta, M_{(\cdot)}) p(\theta | M_{(\cdot)}) \, d\theta, \tag{7.3}$$

with $\theta$ 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
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 \( m(y \mid M(\cdot)) \) 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 \( m(y \mid M(\cdot)) \):

\[
\frac{m(y \mid M(\cdot))}{\text{Marginal likelihood}} \approx \frac{1}{N} \sum_{i=1}^{N} p(y \mid \theta_i, M(\cdot)), \quad \theta_i \sim p(\theta \mid M(\cdot)) \quad .
\]  

(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 \( p(\theta \mid M(\cdot)) \), and then averaging the corresponding values for \( p(y \mid \theta_i, M(\cdot)) \). 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 \( m(y \mid M(\cdot)) \) 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 \( g(\theta \mid M(\cdot)) \) instead of the prior:

\[
\frac{m(y \mid M(\cdot))}{\text{Marginal likelihood}} = \int p(y \mid \theta, M(\cdot))p(\theta \mid M(\cdot))d\theta = \int p(y \mid \theta, M(\cdot))p(\theta \mid M(\cdot))g(\theta \mid M(\cdot)) \frac{g(\theta \mid M(\cdot))}{g(\theta \mid M(\cdot))}d\theta = \int \frac{p(y \mid \theta, M(\cdot))p(\theta \mid M(\cdot))}{g(\theta \mid M(\cdot))}g(\theta \mid M(\cdot))d\theta = \frac{1}{N} \sum_{i=1}^{N} \frac{p(y \mid \theta_i, M(\cdot))p(\theta_i \mid M(\cdot))}{g(\theta_i \mid M(\cdot))}, \quad \theta_i \sim g(\theta \mid M(\cdot)) .
\]  

(7.5)

This approximation is efficient if the importance density \( g(\theta \mid M(\cdot)) \) is similar to the posterior distribution but different otherwise.
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 \cite{Vandekerckhove2015}. 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 \cite{Vandekerckhove2015}. 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. 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. \cite{Steingroever2015}. 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.

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

\(^a\) 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., \(U(0,1)\)) 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](http://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.

<table>
<thead>
<tr>
<th>Model Comparison</th>
<th>Median SE [%]</th>
<th>25%, 75% Quantile [%]</th>
<th>2.5%, 97.5% Quantile [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE(BF_{EV,PVL})</td>
<td>1.89</td>
<td>[1.36, 2.77]</td>
<td>[0.93, 7.92]</td>
</tr>
<tr>
<td>SE(BF_{EV,PVL−D})</td>
<td>2.21</td>
<td>[1.54, 3.55]</td>
<td>[1.05, 12.32]</td>
</tr>
<tr>
<td>SE(BF_{EV,VPP})</td>
<td>5.69</td>
<td>[4.03, 8.06]</td>
<td>[2.14, 23.70]</td>
</tr>
<tr>
<td>SE(BF_{PVL,PVL−D})</td>
<td>2.13</td>
<td>[1.54, 3.26]</td>
<td>[1.10, 8.71]</td>
</tr>
<tr>
<td>SE(BF_{PVL,VPP})</td>
<td>5.47</td>
<td>[4.00, 7.97]</td>
<td>[2.09, 22.34]</td>
</tr>
<tr>
<td>SE(BF_{PVL−D,VPP})</td>
<td>5.74</td>
<td>[4.27, 8.27]</td>
<td>[2.25, 22.83]</td>
</tr>
</tbody>
</table>

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

$$P(\mathcal{M}(\cdot) \mid y) = \frac{BF_{(\cdot)B}}{\sum_{j=1}^{4} BF_{jB}},$$

(7.6)

where $B$ indicates the reference model which can be any of the four RL models \cite{Berger2005}. 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\footnote{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.} 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(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
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) indicates that the data are more likely to occur under the first model than under the second model, whereas a negative log(BF) 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.
Bayes Factors for Reinforcement-Learning Models of the Iowa Gambling Task

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.

<table>
<thead>
<tr>
<th>Model</th>
<th>MPMP</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>EV</td>
<td>.00</td>
<td>7</td>
</tr>
<tr>
<td>PVL</td>
<td>.04</td>
<td>25</td>
</tr>
<tr>
<td>PVL-Delta</td>
<td>.00</td>
<td>9</td>
</tr>
<tr>
<td>VPP</td>
<td>.64</td>
<td>59</td>
</tr>
</tbody>
</table>

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 $M_1$ than model $M_2$.

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(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
7.6 Discussion

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.

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).

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 $M_1$ over model $M_2$, this does not mean that model $M_1$ 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. 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.
Bayes Factors for Reinforcement-Learning Models of the Iowa Gambling Task

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.
Chapter 8

A Tutorial on Bridge Sampling

This chapter is currently in preparation as:
Quentin F. Gronau, Alexandra Sarafoglou, Dora Matzke, Alexander Ly, Udo Boehm, Maarten Marsman, David S. Leslie, Jonathan J. Forster, Eric-Jan Wagenmakers, Helen Steingroever
(in preparation).
A tutorial on bridge sampling.

Abstract

The marginal likelihood plays an important role in many areas of Bayesian statistics such as parameter estimation, model comparison, and model averaging. In most applications, however, the marginal likelihood is not analytically tractable and must be approximated using numerical sampling methods. Here we provide a tutorial on bridge sampling (Meng & Wong, 1996), a reliable and relatively straightforward sampling method that allows researchers to obtain the marginal likelihood for models of varying complexity. First, we introduce bridge sampling and three related sampling methods using the beta-binomial model as a running example. We then apply bridge sampling to estimate the marginal likelihood for the Expectancy Valence (EV) model—a popular model for reinforcement learning. Our results indicate that bridge sampling provides accurate estimates for both a single participant and a hierarchical version of the EV model. We conclude that bridge sampling is an attractive method for mathematical psychologists who typically aim to approximate the marginal likelihood for a limited set of possibly high-dimensional models.

Bayesian statistics has become increasingly popular in mathematical psychology (Andrews & Baguley, 2013; Bayarri, Benjamin, Berger, & Sellke, 2016; Poirier, 2006; Vanpaemel, 2016; Verhagen, Levy, Millsap, & Fox, 2015; Wetzels et al., 2016). The Bayesian approach is conceptually simple, theoretically coherent, and easily applied to relatively complex problems. These problems involve, for instance, hierarchical modeling (Matzke, Dolan, Batchelder, & Wagenmakers, 2015; Matzke & Wagenmakers, 2009; Rouder & Lu, 2005; Rouder et al., 2005, 2007) or the comparison of non-nested models (Lee, 2008; Lee & Wagenmakers, 2005; Pitt, Myung, & Zhang, 2002; Shiffrin et al., 2008). Three major applications of Bayesian statistics concern parameter estimation, model comparison, and Bayesian model averaging. In all three areas, the marginal likelihood—that is, the
probability of the observed data given the model of interest—plays a central role (see also Gelman & Meng, 1998).

First, in parameter estimation, we consider a single model and aim to quantify the uncertainty for a parameter of interest \( \theta \) after having observed the data \( y \). This is realized by means of a posterior distribution; here, the marginal likelihood of the data \( p(y) \) ensures that the posterior distribution is a proper probability density function (PDF) in the sense that it integrates to 1. This is evident from Bayes theorem:

\[
p(\theta | y) = \frac{p(y | \theta) p(\theta)}{\int p(y | \theta') p(\theta') d\theta'} = \frac{\text{likelihood}}{\int p(y | \theta') p(\theta') d\theta'} \cdot \frac{\text{prior}}{p(y)}.
\]

Equation 8.1 shows that the marginal likelihood is independent of the parameter \( \theta \), and is just a single number. This illustrates why in parameter estimation the marginal likelihood is referred to as a normalizing constant.

Second, in model comparison, we consider \( m \) (\( m \in \mathbb{N} \)) competing models, and are interested in the evidence that the data \( y \) provide for a particular model \( M_i \) (\( i \in \{1, 2, \ldots, m\} \)) given the prior probabilities of all models under consideration (see three special issues on this topic in the *Journal of Mathematical Psychology*: Mulder & Wagenmakers, 2016; J. J. Myung, Forster, & Browne, 2000; Wagenmakers & Waldorp, 2006). This evidence is quantified by the so-called posterior model probability \( p(M_i | y) \) of model \( M_i \) given the data \( y \) (Berger & Molina, 2005):

\[
p(M_i | y) = \frac{p(y | M_i) p(M_i)}{\sum_{j=1}^{m} p(y | M_j) p(M_j)},
\]

where the denominator is the sum of the marginal likelihood times the prior model probability of all \( m \) models. In model comparison, the marginal likelihood for a specific model is also referred to as evidence or likelihood of the model (Kass & Raftery, 1995; Ntzoufras, 2009).

If only two models \( M_1 \) and \( M_2 \) are considered, Equation 8.2 can be used to quantify the relative posterior model plausibility of model \( M_1 \) compared to model \( M_2 \). This relative plausibility is given by the ratio of the posterior probabilities of both models, and is referred to as the posterior model odds:

\[
\frac{p(M_1 | y)}{p(M_2 | y)} = \frac{p(M_1)}{p(M_2)} \cdot \frac{p(y | M_1)}{p(y | M_2)}.
\]

Equation 8.3 illustrates that the posterior model odds are the product of two factors: The first factor is the ratio of the prior probabilities of both models—the prior model odds. The second factor is the ratio of the marginal likelihoods of both models—the so-called Bayes factor (Etz & Wagenmakers, submitted; Jeffreys, 1961; Ly, Verhagen, & Wagenmakers, 2016). The Bayes factor plays an important role in model comparison and is referred to as the “standard Bayesian solution to the hypothesis testing and model selection problems” (Lewis & Raftery, 1997, p. 648) and “the primary tool used in Bayesian inference for hypothesis testing and model selection” (Berger, 2006, p. 378).

Third, the marginal likelihood also plays an important role in Bayesian model averaging (BMA; Hoeting, Madigan, Raftery, & Volinsky, 1999) where aspects of parameter estimation and model
comparison are combined. As in model comparison, BMA considers several models; however, it does not aim to identify a single best model. Instead, BMA aims to quantify the evidence for a specific component (e.g., a specific learning rule) that is assumed by several models under consideration. Specifically, BMA quantifies the model-averaged evidence in favor of that specific component by computing the so-called posterior inclusion probability. The posterior inclusion probability is given by the sum of the posterior model probabilities of all models that contain the specific component, and as such depends on the marginal likelihood of the models.

A problem that arises in all three areas –parameter estimation, model comparison and BMA– is that the marginal likelihood involves an integral (see Equation 8.1) that is tractable only in certain restricted examples. This is a pressing problem in mathematical psychology where models can be non-linear and equipped with a large number of parameters, especially when the models are implemented in a Bayesian hierarchical framework. Such a framework incorporates both commonalities and differences between participants of one group by assuming that the model parameters of each participant are drawn from a group-level distribution (for advantages of the Bayesian hierarchical framework see [Ahn et al., 2011], [Navarro et al., 2006], [Rouder & Liu, 2005], [Rouder et al., 2005, 2008], [Scheibehenne & Pachur, 2015], [Shiffrin et al., 2008], [Wetzels, Vandekerckhove, et al., 2010]). For instance, consider a four-parameter Bayesian hierarchical model with four group-level distributions each characterized by two parameters and a group size of 30 participants; this then results in $30 \times 4$ individual-level parameters and $2 \times 4$ group-level parameters for a total of 128 parameters. In sum, even simple models quickly become complex once hierarchical aspects are introduced and this frustrates the derivation of the marginal likelihood.

To overcome this problem, the marginal likelihood can be approximated using several Monte Carlo sampling methods. In this tutorial we focus on four such methods: the bridge sampling estimator and its three commonly used special cases—the naive Monte Carlo estimator, the importance sampling estimator, and the generalized harmonic mean estimator (for alternative methods see [Gamerman & Lopes, 2006] Chapter 7; and for alternative approximation methods relevant to model comparison and BMA see [Carlin & Chib, 1995], [Green, 1995]).

As we will illustrate throughout this tutorial, the bridge sampler is accurate, efficient, and relatively straightforward to implement (e.g., [DiCiccio, Kass, Raftery, & Wasserman, 1997], [Frühwirth-Schnatter, 2004], [Meng & Wong, 1996]).

The goal of this tutorial is to bring the bridge sampling estimator to the attention of mathematical psychologists. We aim to explain this estimator and facilitate its application by suggesting a step-by-step implementation scheme. To this end, we first show how bridge sampling and the three special cases can be used to approximate the marginal likelihood in a simple beta-binomial model. We begin with the naive Monte Carlo estimator and progressively work our way up—via the importance sampling estimator and the generalized harmonic mean estimator—to the most general case considered: the bridge sampling estimator. This order was chosen such that key concepts are introduced gradually and estimators are of increasing complexity and sophistication. The first three estimators are included in this tutorial with the sole purpose of facilitating the reader’s understanding of bridge sampling. In the second part of this tutorial, we outline how the bridge sampling estimator can be used to derive the marginal likelihood for the Expectancy Valence (EV; Busemeyer & Stout, 2002) model—a popular, yet relatively complex reinforcement-learning model for the Iowa gambling task (Bechara et al., 1994). We apply bridge sampling to both an individual-level and a hierarchical implementation of the EV model.

Throughout the article, we use the software package R to implement the bridge sampling

---

Footnote:

1 Appendix F gives a derivation showing that the first three estimators are indeed special cases of the bridge sampler.
estimator for the various models. The interested reader is invited to reproduce our results by
downloading the code and all relevant materials from our Open Science Framework folder at [osf.io/f9cq4].

8.1 Four Sampling Methods to Approximate the Marginal Likelihood

In this section we outline four standard methods to approximate the marginal likelihood. For more
detailed explanations and derivations, we recommend Ntzoufras (2009, Chapter 11) and Gamerman and Lopes (2006, Chapter 7); a comparative review of the different sampling methods is presented in DiCiccio et al. (1997). The marginal likelihood is the probability of the observed data \( y \) given a specific model of interest \( M \), and is defined as the integral of the likelihood over the prior:

\[
p(y | M) = \int p(y | \theta, M) p(\theta | M) \, d\theta,
\]

with \( \theta \) a vector containing the model parameters. Equation 8.4 illustrates that the marginal likelihood can be interpreted as a weighted average of the likelihood of the data given a specific value for \( \theta \) where the weight is the a priori plausibility of that specific value. Equation 8.4 can therefore be written as an expected value:

\[
p(y | M) = \mathbb{E}_{\text{prior}}(p(y | \theta, M)),
\]

where the expectation is taken with respect to the prior distribution. This idea is central to the four sampling methods that we discuss in this tutorial.

Introduction of the Running Example: The Beta-Binomial Model

Our running example focuses on estimating the marginal likelihood for a binomial model assuming a uniform prior on the rate parameter \( \theta \) (i.e., the beta-binomial model). Consider a single participant who answered \( k = 2 \) out of \( n = 10 \) true/false questions correctly. Assume that the number of correct answers follows a binomial distribution, that is, \( k \sim \text{Binomial}(n, \theta) \) with \( \theta \in (0, 1) \), where \( \theta \) represents the latent probability for answering any one question correctly. The probability mass function (PMF) of the binomial distribution is given by:

\[
\text{Binomial}(k | n, \theta) = \binom{n}{k} \theta^k (1 - \theta)^{n-k},
\]

where \( n \in \mathbb{Z}_{\geq 0} \), and \( k \leq n \). The PMF of the binomial distribution serves as the likelihood function in our running example.

In the Bayesian framework, we also have to specify the prior distribution of the model parameters; the prior distribution expresses our knowledge about the parameters before the data have been observed. In our running example, we assume that all values of \( \theta \) are equally likely a priori. This prior belief is captured by a uniform distribution across the range of \( \theta \), that is, \( \theta \sim \text{Uniform}(0, 1) \) which can equivalently be written in terms of a beta distribution \( \theta \sim \text{Beta}(1, 1) \). This prior distribution is represented by the dotted line in Figure 8.1. It is evident that the density of the prior distribution equals 1 for all values of \( \theta \). One advantage of expressing the
prior distribution by a beta distribution is that its two parameters (i.e., in its general form the shape parameters $\alpha$ and $\beta$) can be thought of as counts of “prior successes” and “prior failures”, respectively. In its general form, the PDF of a Beta($\alpha, \beta$) distribution ($\alpha, \beta > 0$) is given by:

$$
\text{Beta}(x | \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)},
$$

where $B(\alpha, \beta)$ is the beta function that is defined as: $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, and $\Gamma(n) = (n - 1)!$ for $n \in \mathbb{N}$.

Analytical derivation of the marginal likelihood

As we will see in this section, for the beta-binomial model the marginal likelihood is analytic. Assuming a general $k$ and $n$, we obtain the marginal likelihood as:

Figure 8.1: Prior and posterior distribution for the rate parameter $\theta$ from the beta-binomial model. The Beta(1, 1) prior on the rate parameter $\theta$ is represented by the dotted line; the Beta(3, 9) posterior distribution is represented by the solid line and was obtained after having observed 2 correct responses out of 10 trials.
\[ p(k \mid n) = \int_0^1 p(k \mid n, \theta) p(\theta) d\theta = \int_0^1 \binom{n}{k} \theta^k (1 - \theta)^{n-k} 1 d\theta = \binom{n}{k} B(k + 1, n - k + 1) = \frac{1}{n+1}, \]

where we suppress the “model” in the conditioning part of the probability statements because we focus on a single model in this running example. Using \( k = 2 \) and \( n = 10 \) of our example, we obtain:

\[ p(k = 2 \mid n = 10) = \frac{1}{11} \approx 0.0909. \]

The marginal likelihood can now be used to derive the posterior distribution for \( \theta \) after having observed the data. Using Bayes theorem, we obtain:

\[ p(\theta \mid k, n) = p(k \mid n, \theta) p(\theta) / p(k \mid n) = \frac{\binom{n}{k} \theta^k (1 - \theta)^{n-k} 1}{\binom{n}{k} B(k + 1, n - k + 1)} = \frac{\theta^k (1 - \theta)^{n-k}}{B(k + 1, n - k + 1)}, \]

which we recognize as the PDF of the Beta\((k + 1, n - k + 1)\) distribution. Thus, if we assume a uniform prior on \( \theta \) and observe \( k = 2 \) correct responses out of \( n = 10 \) trials, we obtain a Beta\((3, 9)\) distribution as posterior distribution. This distribution is represented by the solid line in Figure 8.1. In general, if \( k \mid n, \theta \sim \text{Binomial}(n, \theta) \) and \( \theta \sim \text{Beta}(1, 1) \), then \( \theta \mid n, k \sim \text{Beta}(k + 1, n - k + 1) \).

**Method 1: The Naive Monte Carlo Estimator of the Marginal Likelihood**

The simplest method to approximate the marginal likelihood is provided by the naive Monte Carlo estimator (Hammersley & Handscomb, 1964; Raftery & Banfield, 1991). This method uses the standard definition of the marginal likelihood (Equation 8.4), and relies on the central idea that the marginal likelihood can be written as an expected value with respect to the prior distribution:

\[ p(y) = \int p(y \mid \theta) p(\theta) d\theta = \mathbb{E}_{\text{prior}}(p(y \mid \theta)). \]

The expected value of the likelihood with respect to the prior can be approximated by evaluating the likelihood in \( N \) samples from the prior distribution for \( \theta \) and averaging the resulting values. This yields the naive Monte Carlo estimator \( \hat{p}_1(y) \):

\[ \hat{p}_1(y) = \frac{1}{N} \sum_{i=1}^{N} \frac{p(y \mid \theta_i)}{\mathbb{P} \sim p(\theta) \quad \text{average likelihood}} \]

where \( \mathbb{P} \) is a sample from the prior distribution.
8.1. Four Sampling Methods to Approximate the Marginal Likelihood

Running example

To obtain the naive Monte Carlo estimate of the marginal likelihood in our running example, we need \( N \) samples from the Beta(1, 1) prior distribution for \( \theta \). For illustrative purposes, we limit the number of samples to 12. We obtain the following samples:

\[
\{\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{12}\} = \{0.58, 0.76, 0.03, 0.93, 0.27, 0.97, 0.45, 0.46, 0.18, 0.64, 0.06, 0.15\},
\]

where we use the tilde symbol to emphasize that we refer to a sampled value. All sampled values are represented by the gray dots in Figure 8.2.

Figure 8.2: Illustration of the naive Monte Carlo estimator for the beta-binomial example. The dotted line represents the prior distribution and the solid line represents the posterior distribution that was obtained after having observed 2 correct responses out of 10 trials. The gray dots represent the 12 samples \( \{\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{12}\} \) randomly drawn from the Beta(1, 1) prior distribution.

Following Equation 8.6, the next step is to calculate the likelihood (Equation 8.5) for each \( \tilde{\theta}_i \), and then to average all obtained likelihood values. This yields the naive Monte Carlo estimate of the marginal likelihood:
\[ \hat{p}_1(y) = \frac{1}{12} \sum_{i=1}^{12} p(k = 2 \mid n = 10, \tilde{\theta}_i) = \frac{1}{12} \sum_{i=1}^{12} \binom{n}{k} (\tilde{\theta}_i)^k (1 - \tilde{\theta}_i)^{n-k} \]

\[ = \frac{1}{12} \left( \frac{10}{2} \right) (0.58^2(1 - 0.58)^8 + \ldots + 0.15^2(1 - 0.15)^8) \]

\[ = 0.0945, \]

where we use \( \hat{p}_1(y) \) to refer to \( \hat{p}_1(k = 2 \mid n = 10) \)—a notation that we adopt throughout this article.

**Method 2: The Importance Sampling Estimator of the Marginal Likelihood**

The naive Monte Carlo estimator introduced in the last section performs well if the prior and posterior distribution have a similar shape. However, the estimator is unstable if the posterior distribution is peaked relative to the prior (Gamerman & Lopes, 2006; Ntzoufras, 2009). In such a situation, most of the sampled values for \( \theta \) result in likelihood values close to zero and contribute only minimally to the estimate. This means that those few samples that result in high likelihood values dominate estimates of the marginal likelihood. Consequently, the variance of the estimator is increased (Newton & Raftery, 1994; Pajor, 2016).

The importance sampling estimator, on the other hand, overcomes this shortcoming by boosting sampled values in regions of the parameter space where the integrand of Equation 8.4 is large. This is realized by using samples from a so-called importance density \( g_{IS}(\theta) \) instead of the prior distribution. The advantage of sampling from an importance density is that values for \( \theta \) that result in high likelihood values are sampled most frequently, whereas values for \( \theta \) with low likelihood values are sampled only rarely.

To derive the importance sampling estimator, Equation 8.4 is used as starting point which is then extended by the importance density \( g_{IS}(\theta) \):

\[ p(y) = \int p(y \mid \theta) p(\theta) \, d\theta = \int p(y \mid \theta) p(\theta) \frac{g_{IS}(\theta)}{g_{IS}(\theta)} \, d\theta = \int \frac{p(y \mid \theta) p(\theta)}{g_{IS}(\theta)} g_{IS}(\theta) \, d\theta \]

This yields the importance sampling estimator \( \hat{p}_2(y) \):

\[ \hat{p}_2(y) = \frac{1}{N} \sum_{i=1}^{N} p(y \mid \tilde{\theta}_i) \frac{\hat{p}(\tilde{\theta}_i)}{g_{IS}(\tilde{\theta}_i)}, \quad \tilde{\theta}_i \sim g_{IS}(\theta). \tag{8.7} \]

A suitable importance density should (1) be easy to evaluate; (2) have the same domain as the posterior distribution; (3) closely resemble the posterior distribution, and (4) have fatter tails than the posterior distribution (Neal, 2001; Vandekerckhove et al., 2015). The latter criterion ensures that values in the tails of the distribution have little impact on the estimator (Neal, 2001).

\[^2\]The interested reader is referred to Pajor (2016) for a recent improvement on the calculation of the naive Monte Carlo estimator. The proposed improvement involves trimming the prior distribution in such a way that regions with low likelihood values are eliminated, thereby increasing the accuracy and efficiency of the estimator.
8.1. Four Sampling Methods to Approximate the Marginal Likelihood

Running example

To obtain the importance sampling estimate of the marginal likelihood in our running example, we first need to choose an importance density \( g_{IS}(\theta) \). An importance density that fulfills the four above mentioned desiderata is a mixture between a beta density that provides the best fit to the posterior distribution and a uniform density across the range of \( \theta \) (Vandekerckhove et al., 2015). The relative impact of the uniform density is quantified by a mixture weight \( \gamma \) that ranges between 0 and 1. The larger \( \gamma \), the higher the influence of the uniform density resulting in a less peaked distribution with thick tails. If \( \gamma = 1 \), the beta mixture density simplifies to the uniform distribution on \([0, 1]\) and if \( \gamma = 0 \), the beta mixture density simplifies to the beta density that provides the best fit to the posterior distribution.

In our specific example, we already know that the Beta(3, 9) density is the beta density that provides the best fit to the posterior distribution because this is the analytic expression of the posterior distribution. However, to demonstrate the general case, we show how we can find the beta distribution with the best fit to the posterior distribution using the method of moments. This method requires draws from the Beta(3, 9) posterior distribution. We obtain:

\[
\{\theta_1^*, \theta_2^*, \ldots, \theta_{12}^*\} = \{0.22, 0.16, 0.09, 0.35, 0.06, 0.27, 0.26, 0.41, 0.20, 0.43, 0.21, 0.12\},
\]

which yields a mean of \( \bar{\theta}^* = 0.232 \) and a variance of \( s_{\theta}^2 = 0.014 \). Note that here we use \( \theta_i^* \) to refer to the \( i^{th} \) sample from the posterior distribution to distinguish it from the previously used \( \theta_i \)—the \( i^{th} \) sample from a distribution other than the posterior distribution, such as a prior distribution or an importance density.

Knowing that, if \( X \sim \text{Beta}(\alpha, \beta) \), then \( E(X) = \alpha/(\alpha + \beta) \) and \( V(X) = \alpha\beta/[(\alpha + \beta)^2(\alpha + \beta + 1)] \), we obtain the following method of moment estimates for \( \alpha \) and \( \beta \):

\[
\hat{\alpha} = \bar{\theta}^* \left( \frac{\bar{\theta}^*(1 - \bar{\theta}^*)}{s_{\theta}^2} - 1 \right) = 0.232 \left( \frac{0.232(1 - 0.232)}{0.0142} - 1 \right) = 2.673,
\]

\[
\hat{\beta} = (1 - \bar{\theta}^*) \left( \frac{\bar{\theta}^*(1 - \bar{\theta}^*)}{s_{\theta}^2} - 1 \right) = (1 - 0.232) \left( \frac{0.232(1 - 0.232)}{0.0142} - 1 \right) = 8.865.
\]

Using a mixture weight on the uniform component of \( \gamma = 0.30 \)—a choice that was made to ensure that, visually, the tails of the importance density are clearly thicker than the tails of the posterior distribution—we obtain the following importance density:

\[
\gamma \times \text{Beta}(\theta \mid 1, 1) + (1 - \gamma) \times \text{Beta}(\theta \mid \hat{\alpha}, \hat{\beta}) = 0.3 + 0.7 \text{ Beta}(\theta \mid 2.673, 8.865).
\]

This importance density is represented by the dashed line in Figure 8.3. The figure also shows the posterior distribution (solid line). As is evident from the figure, the beta mixture importance density resembles the posterior distribution, but has fatter tails—a crucial criterion of the importance density in importance sampling.

In general, it is advised to choose the mixture weight on the uniform component \( \gamma \) small enough to make the estimator efficient, yet large enough to produce fat tails to stabilize the estimator. A suitable mixture weight can be realized by gradually minimizing the mixture weight and investigating whether stability is still guaranteed (i.e., robustness analysis).

Drawing \( N = 12 \) samples for \( \theta \) from our beta mixture importance density results in:

\[\text{In our running example, the importance sampling estimator then reduces to the naive Monte Carlo estimator.}\]
\{\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{12}\} = \{0.11, 0.07, 0.33, 0.25, 0.41, 0.39, 0.25, 0.13, 0.64, 0.26, 0.74, 0.92\}.

These samples are represented by the gray dots in Figure 8.3.

Figure 8.3: Illustration of the importance sampling estimator for the beta-binomial model. The dashed line represents our beta mixture importance density and the solid gray line represents the posterior distribution that was obtained after having observed 2 correct responses out of 10 trials. The gray dots represent the 12 samples \{\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{12}\} randomly drawn from our beta mixture importance density.

The final step is to compute the average adjusted likelihood for the 12 samples using Equation 8.7. This yields the importance sampling estimate of the marginal likelihood as:
8.1. Four Sampling Methods to Approximate the Marginal Likelihood

\[ \hat{p}_2(y) = \frac{1}{12} \sum_{i=1}^{12} \frac{p(k = 2 \mid n = 10, \hat{\theta}_i) \ p(\hat{\theta}_i)}{3 + .7 \ Beta(\hat{\theta}_i \mid 2.673, 8.865)} \]

\[ = \frac{1}{12} \left( \binom{10}{2} \frac{0.11^2 (1 - 0.11)^8 \times 1}{.3 + .7 \ Beta(0.11 \mid 2.673, 8.865)} + \ldots + \binom{10}{2} \frac{0.92^2 (1 - 0.92)^8 \times 1}{.3 + .7 \ Beta(0.92 \mid 2.673, 8.865)} \right) \]

\[ = \frac{1}{12} \binom{10}{2} (0.0021 + \ldots + 4.7 \times 10^{-9}) \]

\[ = 0.0829. \]

**Method 3: The Generalized Harmonic Mean Estimator of the Marginal Likelihood**

Just as the importance sampling estimator, the generalized harmonic mean estimator focuses on regions of the parameter space where the integrand of Equation [8.4] is large by using an importance density \( g_{IS}(\theta) \) (Gelfand & Dey, 1994). However, in contrast to the importance sampling estimator, the generalized harmonic mean estimator requires an importance density with thinner tails, and it only requires samples from one density (i.e., the posterior density).

To derive the generalized harmonic mean estimator, also known as reciprocal importance sampling estimator (Frühwirth-Schnatter, 2004), we use the following identity:

\[ \frac{1}{p(y)} = \int \frac{1}{p(y)} g_{IS}(\theta) \ d\theta = \int \frac{p(\theta \mid y)}{p(y \mid \theta) p(\theta)} g_{IS}(\theta) \ d\theta = \int \frac{g_{IS}(\theta)}{p(y \mid \theta) p(\theta)} \ p(\theta \mid y) \ d\theta \]

\[ = E_{post} \left( \frac{g_{IS}(\theta)}{p(y \mid \theta) p(\theta)} \right). \]

Rewriting results in:

\[ p(y) = \left( E_{post} \left( \frac{g_{IS}(\theta)}{p(y \mid \theta) p(\theta)} \right) \right)^{-1}, \]

which is used to define the generalized harmonic mean estimator \( \hat{p}_3(y) \) (Gelfand & Dey, 1994) as follows:

\[ \hat{p}_3(y) = \left( \frac{1}{N} \sum_{j=1}^{N} \frac{g_{IS}(\theta_j^y)}{p(y \mid \theta_j^y) p(\theta_j^y)} \right)^{-1}, \quad \theta_j^y \sim p(\theta \mid y). \quad \text{(8.8)} \]

Note that the generalized harmonic mean estimator is a more stable version of the harmonic mean estimator (Newton & Raftery, 1994). A problem of the harmonic mean estimator is that it is dominated by samples that have small likelihood values.
Note that the generalized harmonic mean estimator—in contrast to the importance sampling estimator—evaluates samples from the posterior distribution.

As for the importance sampling estimator, the importance density for the generalized harmonic mean estimator should (1) be easy to evaluate; (2) have the same domain as the posterior distribution; (3) closely resemble the posterior distribution; but (4) should have thinner tails than the posterior distribution—a crucial difference to the importance density of the importance sampling estimator [Newton & Raftery, 1994; DiCiccio et al., 1997].

Running example

To obtain the generalized harmonic mean estimate of the marginal likelihood in our running example, we need to choose a suitable importance density. In our running example, an importance density that fulfills the four above mentioned desiderata can be obtained by following four steps:

First, we draw \( N = 12 \) samples from the posterior distribution. Reusing the samples from the last section, we obtain:

\[
\{\theta^*_1, \theta^*_2, \ldots, \theta^*_n\} = \{0.22, 0.16, 0.09, 0.35, 0.06, 0.27, 0.26, 0.41, 0.20, 0.43, 0.21, 0.12\},
\]

Second, we probit-transform all posterior samples (i.e., \( \xi^*_j = \Phi^{-1}(\theta^*_j) \), \( j = \{1, 2, \ldots, 12\} \)). The result of this transformation is that the samples range across the entire real line instead of the \((0, 1)\) interval only. We obtain:

\[
\{\xi^*_1, \xi^*_2, \ldots, \xi^*_{12}\} = \{-0.77, -0.99, -1.34, -0.39, -1.55, -0.61, -0.64, -0.23, -0.84, -0.18, -0.81, -1.17\}.
\]

These probit-transformed samples are represented by the gray dots in Figure 8.4.

Third, we search for the normal distribution that provides the best fit to the probit-transformed posterior samples \( \xi^*_j \). Using the method of moments, we obtain as estimates \( \hat{\mu} = -0.793 \) and \( \hat{\sigma} = 0.423 \). Note that the choice of a normal importance density justifies step 2; the probit transformation was required to match the range of the posterior distribution to the one of the normal distribution.

Finally, as importance density we choose a normal distribution with mean \( \mu = -0.793 \) and standard deviation \( \sigma = 0.423/1.5 \). The choice of a smaller standard deviation ensures thinner tails of the importance density than of the probit-transformed posterior distribution (for a discussion of alternative importance densities see DiCiccio et al., 1997). We decided to divide \( \hat{\sigma} \) by 1.5 for illustrative purposes only. Our importance density is displayed in Figure 8.4 (dashed line) together with the probit-transformed posterior distribution (solid line).

The generalized harmonic mean estimate can now be obtained using either the original posterior samples \( \theta^*_j \) or the probit-transformed samples \( \xi^*_j \). Here we use the latter ones (see also Overstall & Forster, 2010). Incorporating our specific importance density and a correction for having used the probit-transformation, Equation 8.8 becomes:

\[ \text{[Details provided in Appendix F. Note that using the original posterior samples \( \theta^*_j \) would involve transforming the importance density that ranges across the real line to the (0, 1) interval.]} \]
8.1. Four Sampling Methods to Approximate the Marginal Likelihood

Figure 8.4: Illustration of the generalized harmonic mean estimator for the beta-binomial model. The solid line represents the probit-transformed Beta(3, 9) posterior distribution that was obtained after having observed 2 correct responses out of 10 trials, and the dashed line represents the importance density $N(\xi; \mu = -0.793, \sigma = 0.423/1.5)$. The gray dots represent the 12 probit-transformed samples $\{\xi_1^*, \xi_2^*, \ldots, \xi_{12}^*\}$ randomly drawn from the Beta(3, 9) posterior distribution.

$$
\hat{p}_3(y) = \left( \frac{1}{N} \sum_{j=1}^{N} \frac{1}{\sigma} \phi \left( \frac{\xi_j^* - \mu}{\sigma} \right) \frac{p(y | \Phi(\xi_j^*))}{p(\phi(\xi_j^*))} \right)^{-1}, \quad \xi_j^* = \Phi^{-1}(\theta_j^*) \text{ and } \theta_j^* \sim p(\theta | y). \quad (8.9)
$$

Equation 8.9 correctly suggests that the uniform prior on $\theta$ translates to a standard normal prior $\phi$ on $\xi$ because $p(\Phi(\xi)) = 1 \forall \xi$. For our beta-binomial model, we now obtain the generalized harmonic mean estimate of the marginal likelihood as:

141
Method 4: The Bridge Sampling Estimator of the Marginal Likelihood

Both the importance sampling estimator and the generalized harmonic mean estimator are sensitive to the tail behavior of the importance density relative to the posterior distribution. The bridge sampler, on the other hand, is robust to the tail behavior of the posterior and therefore produces more stable estimates (DiCiccio et al., 1997; Frühwirth-Schnatter, 2004; Gelman & Meng, 1998; Meng & Wong, 1996). Another distinctive feature of the bridge sampling estimator is that it relies on draws from two distributions to approximate the marginal likelihood: the posterior distribution \( p(\theta | y) \) and a proposal distribution \( g(\theta) \).

The proposal distribution \( g(\theta) \) is conceptually similar to an importance density and should resemble the posterior distribution. The distance between the proposal distribution and the posterior distribution is reduced by a so-called bridge function \( h(\theta) \). Due to this function, the bridge sampler is relatively robust to the choice of the proposal distribution.

Originally, bridge sampling was developed to directly estimate the Bayes factor, that is, the ratio of the marginal likelihoods of two models \( M_1 \) and \( M_2 \) (e.g., Jeffreys, 1961; Kass & Raftery, 1995):

\[
\frac{p(y | M_1)}{p(y | M_2)} = \frac{\int p(y | \theta, M_1) p(\theta | M_1) \, d\theta}{\int p(y | \theta, M_2) p(\theta | M_2) \, d\theta}.
\]

which is equivalent to:

\[
\frac{p(y | M_1)}{p(y | M_2)} = \frac{\int p(y | \theta, M_1) p(\theta | M_1) \, h(\theta) \, g(\theta) \, d\theta}{\int p(y | \theta, M_2) p(\theta | M_2) \, h(\theta) \, g(\theta) \, d\theta}.
\]  

In this tutorial, we use a version of bridge sampling that allows us to approximate the marginal likelihood of a single model (for an earlier application see for example Overstall & Forster, 2010). In the remainder, we therefore suppress the “model” in the conditioning part of the probability statements. Equation 8.10 then simplifies to:
In addition, we need to compute the Monte Carlo error is minimized if the bridge function is defined as:

$$ h(\theta) = \frac{1}{s_1 p(y | \theta)p(\theta) + s_2 p(y)g(\theta)} , $$

where $s_1 = \frac{N_1}{N_2+N_1}$, $s_2 = \frac{N_2}{N_2+N_1}$, and $C$ a constant. Equation 8.13 shows that the optimal choice of $h(\theta)$ depends on the marginal likelihood $p(y)$ which is the very entity we want to approximate. We can resolve this issue by applying an iterative scheme that updates an initial guess of the marginal likelihood until the estimate of the marginal likelihood has converged according to a predefined tolerance level. To do so, we insert the expression for the optimal bridge function (Equation 8.13) into Equation 8.11 with the marginal likelihood $p(y)$, and dividing the right side by its reciprocal results in:

$$ \frac{1}{N_2} \sum_{i=1}^{N_2} p(y | \tilde{\theta}_i) p(\tilde{\theta}_i) h(\tilde{\theta}_i) = \frac{1}{N_1} \sum_{j=1}^{N_1} h(\theta_j^*) g(\theta_j^*), \quad \tilde{\theta}_i \sim g(\theta), \quad \theta_j^* \sim p(\theta | y) . $$

Choosing the optimal bridge function

The purpose of the bridge function $h(\theta)$ is to reduce the distance between the proposal distribution $g(\theta)$ and the posterior distribution $p(\theta | y)$ by linking them together. Hence, the optimal bridge function should represent this link by supporting both distributions. Meng and Wong (1996, p. 837) showed that the Monte Carlo error is minimized if the bridge function is defined as:
in Equation 8.12 (Meng & Wong [1996]). The formula to approximate the marginal likelihood on iteration \( t + 1 \) is then specified as follows:

\[
\hat{p}(y)^{(t+1)} = \frac{1}{N_2} \sum_{i=1}^{N_2} \frac{p(y \mid \hat{\theta}_i)p(\hat{\theta}_i)}{s_1 p(y \mid \hat{\theta}_i)p(\hat{\theta}_i) + s_2 \hat{p}(y)^{(t)} g(\theta^*_j)} \cdot \frac{1}{N_1} \sum_{j=1}^{N_1} \frac{g(\theta^*_j)}{s_1 p(y \mid \theta^*_j)p(\theta^*_j) + s_2 \hat{p}(y)^{(t)} g(\theta^*_j)}
\]  

\[8.14\]

where \( \hat{p}(y)^{(t)} \) denotes the estimate of the marginal likelihood on iteration \( t \) of the iterative scheme. Extending the numerator of the right side of the Equation 8.14 with \( 1/g(\theta^*_j) \), and the denominator with \( 1/g(\theta^*_j) \), and subsequently defining \( l_{1,j} := \frac{p(y \mid \theta^*_j)p(\theta^*_j)}{g(\theta^*_j)} \) and \( l_{2,i} := \frac{p(y \mid \hat{\theta}_i)p(\hat{\theta}_i)}{g(\theta^*_j)} \), we obtain the formula for the iterative scheme of the bridge sampling estimator \( \hat{p}_4(y)^{(t+1)} \) at iteration \( t + 1 \) (Meng & Wong [1996] p. 837).

\[
\hat{p}_4(y)^{(t+1)} = \frac{1}{N_2} \sum_{i=1}^{N_2} \frac{p(y \mid \hat{\theta}_i)p(\hat{\theta}_i)}{s_1 l_{2,i} + s_2 \hat{p}_4(y)^{(t)}} \cdot \frac{1}{N_1} \sum_{j=1}^{N_1} \frac{g(\theta^*_j)}{s_1 l_{1,j} + s_2 \hat{p}_4(y)^{(t)}}
\]  

\[8.15\]

Equation 8.15 suggests that, in order to obtain the bridge sampling estimate of the marginal likelihood, a number of requirements need to be fulfilled. First, we need \( N_2 \) samples from the proposal distribution \( g(\theta) \) and \( N_1 \) samples from the posterior distribution \( p(\theta \mid y) \). Second, for all \( N_2 \) samples from the proposal distribution, we have to evaluate \( l_{2,i} \). This involves obtaining the value of the unnormalized posterior (i.e., the product of the likelihood times the prior) and of the proposal distribution for all samples. Third, we have to evaluate \( l_{1,j} \) for all \( N_1 \) samples from the posterior distribution. This is analogous to evaluating \( l_{2,i} \). Fourth, we have to determine the constants \( s_1 \) and \( s_2 \) that only depend on \( N_1 \) and \( N_2 \). Fifth, we need an initial guess of the marginal likelihood \( \hat{p}_4(y) \). Since some of these five requirements can be obtained easier than others, we will point to possible challenges.

A first challenge is that using a suitable proposal distribution may involve transforming the posterior samples. Consequently, as in the running example of the generalized harmonic mean estimator, we have to determine how the transformation affects the definition of the bridge sampling estimator for the marginal likelihood (Equation 8.15).
A second challenge is how to use the $N_1$ posterior samples for fitting the proposal distribution and for evaluating $l_{1,j}$ most efficiently. One option is to use all $N_1$ samples for both fitting the proposal distribution and for evaluating $l_{1,j}$. However, Overstall and Forster (2010) showed that it is more efficient to divide the posterior samples in two parts; the first part is used to obtain the best-fitting proposal distribution, and the second part is used for the function evaluations. Throughout this tutorial, we use two equally large parts. In the remainder we therefore state that we draw $2N_1$ samples from the posterior distribution. In addition, we obtain the first part by selecting only those posterior samples with even index numbers; posterior samples with odd index numbers constitute the second part (for an alternative way of splitting the posterior samples see Overstall & Forster [2010]).

Running example

To obtain the bridge sampling estimate of the marginal likelihood in the beta-binomial example, we follow the eight steps illustrated in Figure 8.5.

1. **We draw $2N_1 = 24$ samples from the Beta(3, 9) posterior distribution for $\theta$.**
   
   We obtain the following sample of 24 values:
   
   $$\{\theta_1^*, \theta_2^*, \ldots, \theta_{24}^*\} = \{0.22, 0.16, 0.09, 0.35, 0.06, 0.26, 0.27, 0.26, 0.41, 0.20, 0.43, 0.21, 0.12, 0.15, 0.21, 0.24, 0.18, 0.12, 0.22, 0.15, 0.22, 0.23, 0.26, 0.29, 0.28\}$$

   Note that the first 12 samples equal the ones used in the last section, whereas the last 12 samples were obtained from drawing again 12 values from the Beta(3, 9) posterior distribution for $\theta$. 

   ![Figure 8.5: Schematic illustration of the steps involved in obtaining the bridge sampling estimate of the marginal likelihood.](image-url)
2. We choose a proposal distribution.
   Here we opt for an approach that can be easily generalized to models with multiple parameters
   and select a normal distribution as the proposal distribution $g(\theta)$.

3. We transform the first batch of $N_1$ posterior samples.
   Since we use a normal proposal distribution, we have to transform the posterior samples
   from the rate scale to the real line so that the range of the posterior distribution matches
   the range of the proposal distribution. This can be achieved by probit-transforming the posterior
   samples, that is, $\xi_j^* = \Phi^{-1}(\theta_j^*)$ with $j \in \{2, 4, \ldots, 24\}$. We obtain:
   $$\{\xi_2^*, \xi_4^*, \ldots, \xi_{24}^*\} = \{-0.99, -0.39, -0.61, -0.23, -0.18, -1.17, -0.81, -0.92, -0.77, -0.77,$$
   $$-0.64, -0.58\}.$$  

4. We fit the proposal distribution to the first batch of $N_1$ probit-transformed posterior samples.
   We use the method of moment estimates $\hat{\mu} = -0.672$ and $\hat{\sigma} = 0.298$ from the first batch
   of $N_1$ probit-transformed posterior samples to obtain our proposal distribution $g(\xi \mid \mu = -0.672, \sigma = 0.298) = \frac{1}{0.298} \phi \left( \frac{\xi + 0.672}{0.298} \right)$.

5. We draw $N_2$ samples from the proposal distribution.
   We obtain:
   $$\{\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_{12}\} = \{-0.90, -0.55, -1.16, -0.53, -0.45, -0.60, -0.63, -0.48, -0.69,$$
   $$-1.20 - 0.65 - 0.79\}.$$  

6. We calculate $l_{2,i}$ for all $N_2$ samples from the proposal distribution.
   This step involves assessing the value of the unnormalized posterior and the proposal
   distribution for all $N_2$ samples from the proposal distribution. As in the running
   example for the generalized harmonic mean estimator, we obtain the unnormalized posterior
   as: $p(y \mid \Phi(\hat{\xi}_i)) p(\Phi(\hat{\xi}_i)) \phi(\hat{\xi}_i)$, where the extra term $\phi(\hat{\xi}_i)$ comes from using the
   change-of-variable method (see running example for the generalized harmonic mean estimator
   and Appendix F for details). Thus, as in the case of the generalized harmonic mean estimator,
   the uniform prior on $\theta$ translates to a standard normal prior on $\xi$. The values of the proposal
   distribution can easily be obtained (for example using the R software).

7. We transform the second batch of the $N_1$ posterior samples.
   As in step 2, we use the probit transformation and obtain:
   $$\{\xi_2^*, \xi_3^*, \ldots, \xi_{23}^*\} = \{-0.77, -1.34, -1.55, -0.64, -0.84, -0.81, -1.04, -0.71, -1.17, -1.04,$$
   $$-0.74, -0.55\}.$$  

8. We calculate $l_{1,j}$ for the second batch of $N_1$ probit-transformed samples from the posterior
   distribution.
   This is analogous to step 6.

---

6 There exist several candidates for the proposal distribution. Alternative proposal distributions are, for example,
the importance density that we used for the importance sampling estimator (i.e., a beta mixture importance density)
or for the generalized harmonic mean estimator (i.e., a normal distribution with best fit to the posterior distribution,
but with a smaller standard deviation), or the analytically derived Beta(3, 9) posterior distribution.
9. **We run the iterative scheme (Equation 8.15)** until our predefined tolerance criterion is reached.

As tolerance criterion we choose \( |\hat{p}_4(y)_{(t+1)} - \hat{p}_4(y)_{(t)}| / \hat{p}_4(y)_{(t+1)} \leq 10^{-10} \). This requires an initial guess for the marginal likelihood \( \hat{p}_4(y)_{(0)} \) which we set to 0.

The simplicity of the beta-binomial model allows us to calculate the bridge sampling estimate by hand. To determine \( \hat{p}_4(y)_{(t+1)} \) according to Equation 8.15, we need to calculate the constants \( s_1 \) and \( s_2 \). Since \( N_1 = N_2 \), we obtain: \( s_1 = s_2 = N_2/(N_2 + N_1) = 12/(12 + 12) = 0.5 \). In addition, we need to calculate \( l_{2,i} \) (\( i \in \{1, 2, \ldots , 12\} \)) for all samples from the proposal distribution, and \( l_{1,j} \) (\( j \in \{1, 3, \ldots , 23\} \)) for the second batch of the probit-transformed samples from the posterior distribution. Here we show how to calculate \( l_{2,1} \) and \( l_{1,1} \) for the first sample from the proposal distribution and the posterior distribution, respectively:

\[
\begin{align*}
\hat{p}_4(y)_{(t+1)} &= \frac{1}{12} \sum_{j=1}^{2N_1} \frac{l_{2,i}}{s_1 l_{2,i} + s_2 \hat{p}_4(y)_{(t)}} \\
&= \frac{1}{12} \left( \frac{0.080}{0.5 \cdot 0.080 + 0.5 \cdot \hat{p}_4(y)_{(t)}} + \ldots + \frac{0.071}{0.5 \cdot 0.085 + 0.5 \cdot \hat{p}_4(y)_{(t)}} \right).
\end{align*}
\]

Using \( t = 0 \), we obtain as updated estimate of the marginal likelihood \( \hat{p}_4(y)_{(1)} = 0.091 \). This iterative procedure has to be repeated until our predefined tolerance criterion is reached. For our running example, this criterion is reached after \( t = 6 \) iterations. We now obtain the bridge sampling estimate of the marginal likelihood as \( \hat{p}_4(y)_{(7)} = 0.0894 \).
Interim Summary

So far we used the beta-binomial model to illustrate the computation of four different estimators of the marginal likelihood. These four estimators were discussed in order of increasing sophistication, such that the first three estimators provided the proper context for understanding the fourth, most general estimator—the bridge sampler. This estimator is the focus in the remainder of this tutorial. The goal of the next sections is to demonstrate that bridge sampling is particularly suitable to estimate the marginal likelihood of popular models in mathematical psychology. Importantly, bridge sampling may be used to obtain accurate estimates of the marginal likelihood of hierarchical models (for a detailed comparison of bridge sampling versus its special cases see Frühwirth-Schnatter, 2004; Sinharay & Stern, 2005).

Assessing the Accuracy of the Bridge Sampling Estimate

In this section we show how to quantify the accuracy of the bridge sampling estimate. A straightforward approach would be to apply the bridge sampling procedure multiple times and investigate the variability of the marginal likelihood estimate. In practice, however, this solution is often impractical due to the substantial computational burden of obtaining the posterior samples and evaluating the relevant quantities in the bridge sampling procedure.

Frühwirth-Schnatter (2004) proposed an alternative approach that approximates the estimator’s expected relative mean-squared error:

\[ RE^2 = \frac{\mathbb{E} \left[ (\hat{p}_4(y) - p(y))^2 \right]}{p(y)^2}. \]  

(8.16)

The derivation of this approximate relative mean-squared error by Frühwirth-Schnatter takes into account that the samples from the proposal distribution \( g(\theta) \) are independent, whereas the MCMC samples from the posterior distribution \( p(\theta|y) \) may be autocorrelated. The approximate relative mean-squared error is given by:

\[ \overline{RE}^2 = \frac{1}{N_2} \frac{V_g(\theta)}{\mathbb{E}_g^2(\theta)} f_1(\theta) + \frac{\rho_{f_2}(0) V_{\text{post}}(f_2(\theta))}{N_1} \mathbb{E}_{\text{post}}^2(f_2(\theta)), \]  

(8.17)

where \( f_1(\theta) = \frac{p(\theta|y)}{\sqrt{\pi_1 p(\theta|y) + \pi_2 g(\theta)}}, \ f_2(\theta) = \frac{g(\theta)}{\sqrt{\pi_1 p(\theta|y) + \pi_2 g(\theta)}}, \ V_g(\theta) = \int (f_1(\theta) - \mathbb{E}[f_1(\theta)])^2 g(\theta) \, d\theta \) denotes the variance of \( f_1(\theta) \) with respect to the proposal distribution \( g(\theta) \) (the variance \( V_{\text{post}}(f_2(\theta)) \) is defined analogously), and \( \rho_{f_2}(0) \) corresponds to the normalized spectral density of the autocorrelated process \( f_2(\theta) \) at the frequency 0.

In practice, we approximate the unknown variances and expected values by the corresponding sample variances and means. Hence, for evaluating the variance and expected value with respect to \( g(\theta) \), we use the \( N_2 \) samples for \( \hat{\theta}_i \) from the proposal distribution. To evaluate the variance and expected value with respect to the posterior distribution, we use the second batch of \( N_1 \) samples \( \theta_j^* \) from the posterior distribution which we also use in the iterative scheme for computing the marginal likelihood. Because the posterior samples are obtained via an MCMC procedure and are hence autocorrelated, the second term in Equation (8.17) is adjusted by the normalized spectral density (for details see Frühwirth-Schnatter, 2004). To evaluate the normalized posterior density which appears in the numerator of \( f_1(\theta) \) and the denominator of both \( f_1(\theta) \) and \( f_2(\theta) \), we use the bridge sampling estimate as normalizing constant.

---

8We estimate the spectral density at frequency zero by fitting an autoregressive model using the spectrum0.ar() function as implemented in the coda R package (Plummer, Best, Cowles, & Vines, 2006).
Note that, under the assumption that the bridge sampling estimator \( \hat{p}_4(y) \) is an unbiased estimator of the marginal likelihood \( p(y) \), the square root of the expected relative mean-squared error (Equation 8.16) can be interpreted as the coefficient of variation (i.e., the ratio of the standard deviation and the mean; C. E. Brown, 1998). In the remainder of this article, we report the coefficient of variation to quantify the accuracy of the bridge sampling estimate.

8.2 Case Study: Bridge Sampling for Reinforcement Learning Models

In this section, we illustrate the computation of the marginal likelihood using bridge sampling in the context of a published data set (Busemeyer & Stout, 2002) featuring the Expectancy Valence (EV) model—a popular reinforcement learning (RL) model for the Iowa gambling task (IGT; Bechara et al., 1994). We first introduce the task and the model, and then use bridge sampling to estimate the marginal likelihood of the EV model implemented in both an individual-level and a hierarchical Bayesian framework. For the individual-level framework, we compare estimates obtained from bridge sampling to importance sampling estimates published in Steingroever et al. (2016). For the hierarchical framework, we compare our results to estimates from the Savage-Dickey density ratio test (Dickey, 1971; Dickey & Lientz, 1970; Wagenmakers et al., 2010; Wetzels, Grasman, & Wagenmakers, 2010).

The Iowa Gambling Task

In this section we describe the IGT (see also Steingroever et al., submitted; Steingroever, Wetzels, Horstmann, et al., 2013; Steingroever, Wetzels, & Wagenmakers, 2013a, 2013b; Steingroever et al., 2014, 2016). Originally, Bechara et al. (1994) developed the IGT to distinguish decision-making strategies of patients with lesions to the ventromedial prefrontal cortex from the ones of healthy controls (see also Bechara et al., 1998, 1999, 2000). During the last decades, the scope of application of the IGT has increased tremendously covering clinical populations with, for example, pathological gambling (Cavedini, Riboldi, Keller, et al., 2002), obsessive-compulsive disorder (Cavedini, Riboldi, D’Annucci, et al., 2002), psychopathic tendencies (Blair et al., 2001), and schizophrenia (Bark et al., 2005; Martino et al., 2007).

The IGT is a card game that requires participants to choose, over several rounds, cards from four different decks in order to maximize their long-term net outcome (Bechara et al., 1994, 1997). The four decks differ in their payoffs, and two of them result in negative long-term outcomes (i.e., the bad decks), whereas the remaining two decks result in positive long-term outcomes (i.e., the good decks). After each choice, participants receive feedback on the rewards and losses (if any) associated with that card, as well as their running tally of net outcomes over all trials so far. Unbeknownst to the participants, the task (typically) contains 100 trials.

A crucial aspect of the IGT is whether and to what extent participants eventually learn to prefer the good decks because only choosing from the good decks maximizes their long-term net outcome. The good decks are typically labeled as decks C and D, whereas the bad decks are labeled as decks A and B. Table 8.1 presents a summary of the traditional payoff scheme as developed by Bechara et al. (1994). This table illustrates that decks A and B yield high constant rewards, but even higher unpredictable losses: hence, the long-term net outcome is negative. Decks C and D, on the other hand, yield low constant rewards, but even lower unpredictable 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: decks A and C yield frequent losses, while decks B and D yield infrequent losses.

The Expectancy Valence Model

In this section, we describe the EV model (see also Steingroever, Wetzels, & Wagenmakers, 2013a; Steingroever et al., 2014, 2016, submitted). Originally proposed by Busemeyer and Stout (2002), the EV model is arguably the most popular model for the IGT (for references see Steingroever, Wetzels, & Wagenmakers, 2013a, and for alternative IGT models see Ahn et al., 2008; Dai, Kerestes, Upton, Busemeyer, & Stout, 2015; Steingroever et al., 2014; Worthy, Pang, & Byrne, 2013; Worthy & Maddox, 2014). The model formalizes participants’ performance on the IGT through the interaction of three model parameters that represent distinct psychological processes. The first model assumption is that after choosing a card from deck \( k \in \{1, 2, 3, 4\} \) on trial \( t \), participants compute a weighted mean of the experienced reward \( W(t) \) and loss \( L(t) \) to obtain the utility of deck \( k \) on trial \( t \), \( u_k(t) \):

\[
u_k(t) = (1 - w)W(t) + wL(t).
\]

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 rewarding 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 further 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):

\[
Ev_k(t) = Ev_k(t-1) + a(u_k(t) - Ev_k(t-1)).
\]

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

<table>
<thead>
<tr>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad deck with frequent losses</td>
<td>Bad deck with infrequent losses</td>
<td>Good deck with frequent losses</td>
<td>Good deck with infrequent losses</td>
</tr>
<tr>
<td>Reward/trial</td>
<td>100</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Number of losses/10 cards</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Loss/10 cards</td>
<td>−1250</td>
<td>−1250</td>
<td>−250</td>
</tr>
<tr>
<td>Net outcome/10 cards</td>
<td>−250</td>
<td>−250</td>
<td>250</td>
</tr>
</tbody>
</table>
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.

In the next step, the model assumes that the expected utilities of each deck guide participants’ choices on the next trial \( t + 1 \). This assumption is formalized by the softmax choice rule, also known as the ratio-of-strength choice rule (Luce, 1959):

\[
P[S_k(t + 1)] = \frac{e^{\theta(t) \cdot Ev_k(t)}}{\sum_{j=1}^{4} e^{\theta(t) \cdot Ev_j(t)}}.
\]

The EV model uses this rule to compute the probability of choosing each deck on each trial. This rule contains a sensitivity parameter \( \theta \) that indexes the extent to which trial-by-trial choices match the expected deck utilities. Values of \( \theta \) close to zero indicate random choice behavior (i.e., strong exploration), whereas large values of \( \theta \) indicate choice behavior that is strongly determined by the expected utilities (i.e., strong exploitation).

The EV model uses a trial-dependent sensitivity parameter \( \theta(t) \), which also depends on the final model parameter, response consistency \( c \in [-5, 5] \):

\[
\theta(t) = (t/10)^c.
\]

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 \( w \in [0, 1] \), which quantifies the weight of losses over rewards; (2) the updating parameter \( a \in [0, 1] \), which determines the memory for past expectancies; and (3) the response consistency parameter \( c \in [-2, 2] \), which determines the balance between exploitation and exploration.

Data

We applied bridge sampling to a data set published by Busemeyer and Stout (2002). The data set consists of 30 healthy participants each contributing \( T = 100 \) IGT card selections (see Busemeyer and Stout for more details on the data sets).\(^9\)

Application of Bridge Sampling to an Individual-Level Implementation of the EV Model

In this section we describe how we use bridge sampling to estimate the marginal likelihood of an individual-level implementation of the EV model. This implementation estimates model parameters for each participant separately. Accordingly, we also obtain a marginal likelihood of the EV model for every participant.

\(^9\)Note that we excluded three participants due to incomplete choice data.
Schematic execution of the bridge sampler

To obtain the bridge sampling estimate of the marginal likelihood for each participant, we follow the steps outlined in Figure 8.5.

For each participant \( s, s \in \{1, 2, \ldots, 30\} \), we proceed as follows:

1. **For each parameter, we draw \( 2N_1 \) samples from the posterior distribution.**

   Since Steingroever et al. (2016) already fit an individual-level implementation of the EV model separately to the data of each participant in Busemeyer and Stout (2002), we reuse their posterior samples (see Steingroever et al. 2016 for details on the implementation). Therefore, \( 2N_1 \) matches the number of samples obtained from Steingroever et al. (2016) which was at least 5,000; however, whenever this number of samples was insufficient to ensure convergence of the Hamiltonian Monte Carlo (HMC) chains, Steingroever et al. (2016) repeated the fitting routine with 5,000 additional samples. Steingroever et al. (2016) confirmed convergence of the HMC chains by reporting that all \( \hat{R} \) statistics were below 1.05.

2. **We choose a proposal distribution.**

   We use a multivariate normal distribution as a proposal distribution because it is easy to fit, easy to evaluate, and easy to sample from.

3. **We transform the first batch of \( N_1 \) posterior samples.**

   Since we use a multivariate normal distribution as a proposal distribution, we have to transform all posterior samples to the real line using the probit transformation, that is, \( \omega_{s,j}^* = \Phi^{-1}(w_{s,j}^*), \alpha_{s,j}^* = \Phi^{-1}(a_{s,j}^*), \gamma_{s,j}^* = \Phi^{-1}\left( (c_{s,j}^* + 2) / 4 \right) \), \( j = \{2, 4, \ldots, 2N_1\} \).

4. **We fit the proposal distribution to the first batch of \( N_1 \) probit-transformed posterior samples.**

   We use method of moment estimates for the mean vector and the covariance matrix obtained from the first batch of \( N_1 \) probit-transformed posterior samples to specify our multivariate normal proposal distribution.

5. **We draw \( N_2 \) samples from the proposal distribution.**

   We use the R software to randomly draw \( N_2 \) samples from the proposal distribution obtained in step 4. We obtain \((\tilde{\omega}_{s,i}, \tilde{\alpha}_{s,i}, \tilde{\gamma}_{s,i})\) with \( i \in \{1, 2, \ldots, N_2\} \).

6. **We calculate \( l_{2,i} \) for all \( N_2 \) samples from the proposal distribution.**

   This step involves assessing the value of the unnormalized posterior and the proposal distribution for all \( N_2 \) samples from the proposal distribution. Before we can assess the value of the unnormalized posterior (i.e., the product of the likelihood and the prior), we have to derive how our transformation in step 3 affects the unnormalized posterior.

   First, we derive how our transformation affects the likelihood. Following the change-of-variable method, we have to obtain the value of the likelihood in \((\Phi(\tilde{\omega}_{s,i}), \Phi(\tilde{\alpha}_{s,i}), \Phi(\tilde{\gamma}_{s,i}))\). Before formalizing the likelihood of the observed choices of participant \( s \), we define the following variables: We define \( Ch_s(t) \) as a vector containing the sequence of choices made by participant \( s \) up to and including trial \( t \), and \( X_s(t) \) as a vector containing the corresponding sequence of net outcomes. We now obtain the following expression for the likelihood of the observed choices of participant \( s \):
8.2. Case Study: Bridge Sampling for Reinforcement Learning Models

\[ p(Ch_s(T) \mid \Phi(\tilde{\omega}_{s,i}), \Phi(\tilde{\alpha}_{s,i}), \Phi(\tilde{\gamma}_{s,i}), X_s(T)) = \prod_{t=1}^{T} \prod_{k=1}^{4} Pr(S_k(t+1) \mid Ch_s(t), X_s(t)) \cdot \delta_k(t+1). \]  

(8.18)

Here \( T \) is the total number of trials, \( Pr(S_k(t+1) \mid Ch_s(t), X_s(t)) \) is the probability of choosing deck \( k \) on trial \( t + 1 \) given all previous choices and net outcomes, and \( \delta_k(t+1) \) is a dummy variable which is 1 if deck \( k \) is chosen on trial \( t + 1 \) and 0 otherwise.

Second, we have to derive how our transformation affects the priors on each EV model parameter to yield priors on the probit-transformed model parameters. Since Steingroever et al. (2016) used independent uniform priors on \([0, 1]\) – a prior setting that requires transforming the \( c \) parameter to the \([0, 1]\) interval prior to model fitting – we obtain standard normal priors on the probit-transformed model parameters (see beta-binomial example and Appendix F for an explanation).

Finally, we note that the value of the proposal distribution can be obtained using the R software.

7. We transform the second batch of \( N_1 \) posterior samples. This is analogous to step 2.

8. We calculate \( l_{1,j} \) for the second batch of \( N_1 \) probit-transformed samples from the posterior distribution. This is analogous to step 6.

9. We run the iterative scheme (Equation 8.15) until our predefined tolerance criterion is reached. We use a tolerance criterion and initialization analogous to the running example. Once convergence is reached, we receive an estimate of the marginal likelihood for each participant. We use these estimates to derive the coefficient of variation for each participant. The largest coefficient of variation is 1.94% suggesting that the bridge sampler has low variance.\footnote{Note that this measure relates to the marginal likelihoods, not to the log marginal likelihoods.}

Assessing the accuracy of our implementation

To assess the accuracy of our implementation, we compared the marginal likelihood estimates obtained with our bridge sampler to the estimates obtained with importance sampling (Steingroever et al., 2016). Figure 8.6 shows the log marginal likelihoods for the 30 participants of Busemeyer and Stout (2002) obtained with bridge sampling (x-axis) and importance sampling reported by Steingroever et al. (2016; y-axis). The two sets of estimates correspond almost perfectly. These results indicate a successful implementation of the bridge sampler. Thus, this section emphasizes that both the importance sampler and bridge sampler can be used to estimate the marginal likelihood for the data of individual participants. However, when we want to estimate the marginal likelihood of a Bayesian hierarchical model, it may be difficult to find a suitable importance density. The bridge sampler, on the other hand, can be applied more easily and more efficiently.
Application of Bridge Sampling to a Hierarchical Implementation of the EV Model

Using the Busemeyer and Stout (2002) data set in this section again, we illustrate how we used bridge sampling to estimate the marginal likelihood of a hierarchical EV model. In the hierarchical case of the EV model, we account for the hierarchical structure of the data and hence incorporate both the differences and the similarities between participants. Accordingly, we assume that the parameters $w$, $a$ and $c$ from each participant are drawn from a group-level distribution.

Schematic execution of the bridge sampler

To compute the marginal likelihood, we again follow the steps outlined in Figure 8.5 with a few minor modifications.

1. For each parameter, that is, all individual-level and group-level parameters, we draw $2N_1 = 60,000$ samples from the posterior distribution.

To obtain the posterior samples, we fit a hierarchical Bayesian implementation of the EV model to the Busemeyer and Stout (2002) data set using the software JAGS (Plummer, 2003). We assume that, for each participant $s$, $s \in \{1, 2, \ldots, 30\}$, each probit-transformed individual-level parameter (i.e., $\omega_s = \Phi^{-1}(w_s)$, $\alpha_s = \Phi^{-1}(a_s)$, $\gamma_s = \Phi^{-1}((c_s + 2)/4)$) is drawn from a group-level normal distribution characterized by a group-level mean and

---

11We used a model file that is an adapted version of the model file used by Ahn et al. (2011).
standard deviation parameter. For all group-level mean parameters \( \mu_\omega, \mu_\alpha, \mu_\gamma \) we assume a standard normal distribution, and for all group-level standard deviation parameters \( \sigma_\omega, \sigma_\alpha, \sigma_\gamma \) a uniform distribution ranging from 0 to 1.5. For a detailed explanation of the hierarchical implementation of the EV model, see [Wetzels, Vandekerckhove et al. (2010)].

To reach convergence and reduce autocorrelation, we collect two MCMC chains, each with 120,000 samples from the posterior distributions after having excluded the first 30,000 samples as burn-in. Out of these 120,000 samples per chain, we retained every 4th value yielding 30,000 samples per chain. This setting resulted in all \( \hat{R} \) statistics below 1.05 suggesting that all chains have successfully converged from their starting values to their stationary distributions.

2. **We choose a proposal distribution.**

   We use a multivariate normal distribution as a proposal distribution.

3. **We transform the first batch of \( N_1 \) posterior samples.**

   As before, we ensure that the range of the posterior distribution matches the range of the proposal distribution by using the probit transformation, that is, \( \omega_{s,j}^* = \Phi^{-1}(w_{s,j}) \), \( \alpha_{s,j}^* = \Phi^{-1}(a_{s,j}) \), \( \gamma_{s,j}^* = \Phi^{-1}(\gamma_{s,j}) \), \( \tau_{s,j}^* = \Phi^{-1}(\tau_{s,j}) / 1.5 \), and \( \tau_{\gamma,j}^* = \Phi^{-1}(\tau_{\gamma,j}) / 1.5 \), \( j = \{2, 4, \ldots, 2N_1\} \). The group-level mean parameters do not have to be transformed because they already range across the entire real line.

4. **We fit the proposal distribution to the first batch of the \( N_1 \) probit-transformed posterior samples.**

   We use method of moment estimates for the mean vector and the covariance matrix obtained from the first batch of \( N_1 \) probit-transformed posterior samples to specify our multivariate normal proposal distribution.

5. **We draw \( N_2 \) samples from the proposal distribution.**

   We use the R software to randomly draw \( N_2 \) samples from the proposal distribution obtained in step 4. We obtain \( (\tilde{\omega}_{s,i}, \tilde{\alpha}_{s,i}, \tilde{\gamma}_{s,i}) \) and \( (\tilde{\mu}_\omega, i, \tilde{\tau}_\omega, i, \tilde{\mu}_\alpha, i, \tilde{\tau}_\alpha, i, \tilde{\mu}_\gamma, i, \tilde{\tau}_\gamma, i) \) with \( i \in \{1, 2, \ldots, N_2\} \) and \( s \in \{1, 2, \ldots, 30\} \).

6. **We calculate \( l_{2,i} \) for all \( N_2 \) samples from the proposal distribution.**

   This step involves assessing the value of the unnormalized posterior and the proposal distribution for all \( N_2 \) samples from the proposal distribution. The unnormalized posterior is defined as:

   \[
   \left( \prod_{s=1}^{30} p(CH_s(T) \mid \Phi(\tilde{\kappa}_{s,j}), X_s(T)) p(\tilde{\kappa}_{s,i} \mid \tilde{\zeta}_i) \right) p(\tilde{\zeta}_i), \]

   where \( CH_s(T) \) refers to all choices of subject \( s \), \( X_s(T) \) to the corresponding net outcomes, \( \tilde{\kappa}_{s,i} = (\tilde{\omega}_{s,j}, \tilde{\alpha}_{s,j}, \tilde{\gamma}_{s,j}) \) to the \( i^{th} \) sample from the proposal distribution for the individual-level parameters of subject \( s \), and \( \tilde{\zeta}_i \) to the \( i^{th} \) sample from the proposal distribution for all group-level parameters (e.g., \( \tilde{\zeta}_i = (\tilde{\mu}_\omega, i, \tilde{\tau}_\omega, i, \tilde{\mu}_\alpha, i, \tilde{\tau}_\alpha, i, \tilde{\mu}_\gamma, i, \tilde{\tau}_\gamma, i) \)).

   The likelihood function for a given participant is the same as in the individual case. However, for each participant we now have to add besides the prior on the individual-level parameters also the prior on the group-level parameters. The product of the likelihood and the priors gives the unnormalized posterior density (see Appendix [I] for details).

   Finally, we note that the value of the proposal distribution can be obtained using the R software.

7. **We follow steps 7 – 9, as outlined for the bridge sampler of the individual-level implementation of the EV model.**
Table 8.2: Bayes Factors Comparing the Full EV Model to the Restricted EV Models, log Marginal Likelihoods, and Coefficient of Variation (With Respect to the Marginal Likelihood) Expressed as an Percentage

<table>
<thead>
<tr>
<th>Model</th>
<th>Bayes factor</th>
<th>log marginal likelihood</th>
<th>CV [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>full model</td>
<td>–</td>
<td>-3801.877</td>
<td>10.53</td>
</tr>
<tr>
<td>restricted at $\mu_\alpha = -0.334$</td>
<td>0.729</td>
<td>-3801.561</td>
<td>14.21</td>
</tr>
<tr>
<td>restricted at $\mu_\alpha = -0.604$</td>
<td>0.826</td>
<td>-3801.686</td>
<td>9.99</td>
</tr>
<tr>
<td>restricted at $\mu_\gamma = 0.92$</td>
<td>0.710</td>
<td>-3801.535</td>
<td>13.15</td>
</tr>
</tbody>
</table>

Assessing the accuracy of our implementation

To investigate the accuracy of our implementation, we compare Bayes factors obtained with bridge sampling to Bayes factors obtained from the Savage-Dickey density ratio test (Dickey & Lientz, 1970; Dickey, 1971; for a tutorial, see Wagenmakers et al., 2010). The Savage-Dickey density ratio is a simple method for computing Bayes factors for nested models. We artificially create three nested models by taking the full EV model $M_f$ in which all parameters are free to vary, and then restrict one of the three group-level mean parameters, that is, $\mu_\omega$, $\mu_\alpha$, or $\mu_\gamma$, to a predefined value. For these values we choose the intersection point of the prior and posterior distribution of each group-level mean parameter. To obtain these intersection points, we fit the full EV model and then use a nonparametric logspline density estimator (Stone, Hansen, Kooperberg, Truong, et al., 1997). The obtained values are presented in Table 8.2. Since we compare the full model to each restricted model, we obtain three Bayes factors.

According to the Savage-Dickey density ratio test, the Bayes factor for the full model versus a specific restricted model $M_r$ can be obtained by dividing the height of the prior density at the predefined parameter value $\theta_0$ by the height of the posterior at the same location:

$$BF_{M_f, M_r} = \frac{p(y \mid M_f)}{p(y \mid M_r)} = \frac{p(\theta = \theta_0 \mid M_f)}{p(\theta = \theta_0 \mid y, M_f)}.$$  \(8.19\)

Since we choose $\theta_0$ to be the intersection point of the prior and posterior distribution, $BF_{M_f, M_r}$ equals 1. This Savage-Dickey Bayes factor of 1 indicates that the marginal likelihood under the full model equals the marginal likelihood under the restricted model. Figure 8.7 illustrates the Savage-Dickey Bayes factor comparing the full model to the model assuming $\mu_\alpha$ fixed to $-0.604$.

The computation of the three bridge sampling Bayes factors, on the other hand, works as follows: First, we follow the steps outlined above to obtain the bridge sampling estimate of the full EV model. Second, we obtain the bridge sampling estimate of the marginal likelihood for the three restricted models. This requires adapting the steps outlined above to each of the three restricted models. Lastly, we use Equation 8.19 to obtain the three Bayes factors.

The Bayes factors derived from bridge sampling are reported in Table 8.2. It is evident that Bayes factors derived from bridge sampling closely approximate the Savage-Dickey Bayes factors of 1. These results suggest a successful implementation of the bridge sampler. This is also reflected by the close match between the log marginal likelihoods of the four models presented in the third column of Table 8.2.
Figure 8.7: Prior and posterior distributions of the group-level mean $\mu_\alpha$ in the Busemeyer and Stout (2002) data set. The figure shows the posterior distribution (solid line) and the prior distribution (dotted line). The gray dot indicates the intersection of the prior and the posterior distributions, for which the Savage-Dickey Bayes factor equals 1.

Finally, we confirm that the bridge sampler has low variance; the coefficient of variation with respect to the marginal likelihood of the full model and the three restricted models ranges between 9.99 and 14.21%.

8.3 Discussion

In this tutorial, we explained how bridge sampling can be used to estimate the marginal likelihood of popular models in mathematical psychology. As a running example, we used the beta-binomial model to illustrate step-by-step the bridge sampling estimator. To facilitate the understanding of the bridge sampler, we first discussed three of its special cases—the naive Monte Carlo estimator, the importance sampling estimator, and the generalized harmonic mean estimator. Consequently, we introduced key concepts that became gradually more complicated and sophisticated. In the second part of this tutorial, we showed how bridge sampling can be used to estimate the marginal likelihood of both an individual-level and a hierarchical implementation of the Expectancy Valence (EV; Busemeyer & Stout, 2002) model—a popular reinforcement-learning model for the Iowa gambling task (IGT; Bechara et al., 1994). The running example and the application of bridge sampling to the EV model demonstrated the positive aspects of the bridge sampling estimator, that is, its accuracy, reliability, practicality, and ease-of-implementation (DiCiccio et al., 1997; Frühwirth-Schnatter, 2004; Meng & Wong, 1996).

The bridge sampling estimator is superior to the naive Monte Carlo estimator, the importance
sampling estimator, and the generalized harmonic mean estimator for several reasons. First, Meng and Wong (1996) showed that, among the four estimators discussed in this article, the bridge sampler minimizes the Monte Carlo errors because it uses the optimal bridge function. Second, in bridge sampling, choosing a suitable proposal distribution is much easier than choosing a suitable importance density for the importance sampling estimator or the generalized harmonic mean estimator. In particular, the specific shape of the proposal distribution in bridge sampling is not as relevant because the bridge function corrects for the deviances between the posterior and proposal distribution. Third, due to its robustness to the tail behavior of the proposal distribution relative to the posterior distribution, the bridge sampler can be applied to higher dimensional and complex models. This characteristic of the bridge sampler combined with the popularity of higher dimensional and complex models in mathematical psychology suggests that bridge sampling can advance model comparison exercises in many areas of mathematical psychology (e.g., reinforcement-learning models, response time models, multinomial processing tree models, etc.). Fourth, bridge sampling is relatively straightforward to implement. In particular, our step-by-step procedure can be easily applied to other models with only minor changes of the code (i.e., the unnormalized posterior and the proposal function have to be adapted).

Despite the numerous advantages of the bridge sampler, the take-home message of this tutorial is not that the bridge sampler should be used blindly. There exist a large variety of methods to approximate the marginal likelihood that differ in their efficiency. The most appropriate method optimizes the trade-off between accuracy and implementation effort. This trade-off depends on a number of aspects such as the complexity of the model, the number of models under consideration, the statistical experience of the researcher, and the time available. This suggests that the choice of the method has to be reconsidered each time a marginal likelihood needs to be obtained. Obviously, when the marginal likelihood can be determined analytically, bridge sampling is not needed at all. If the goal is to compare (at least) two nested models, the Savage-Dickey density ratio test (Dickey & Lientz 1970; Dickey 1971) might be a better alternative. If only an individual-level implementation of a model is used, importance sampling may be easier to implement and may require less computational effort. If the goal is to obtain the marginal likelihood of a large number of relatively simple models, the product space or reversible jump method might be more appropriate (Carlin & Chib 1995; Green 1995). If a researcher with a limited programming background and/or little time resources wants to conduct a model comparison exercise, rough approximations of the Bayes factor, such as the Bayesian information criterion, might be more suitable (Schwarz 1978). On the other hand, a researcher with an extensive background in programming and mathematical statistics might consider using path sampling (Gelman & Meng 1998). This method generalizes bridge sampling by using an infinite number of bridges.

To conclude, in this tutorial we showed that bridge sampling offers a reliable and easy-to-implement approach to estimate a model’s marginal likelihood. Bridge sampling can be profitably applied to a wide range of problems in mathematical psychology involving parameter estimation, model comparison, and Bayesian model averaging.

**Acknowledgements**

We thank Busemeyer and Stout (2002) for providing the data used in this article. This research was supported by a Netherlands Organisation for Scientific Research (NWO) grant to UB (406-12-125) and to HS (404-10-086), a European Research Council (ERC) grant to EJW (283876), and a Veni grant (451-15-010) from the NWO to DM.
Chapter 9

Bayesian Techniques for Analyzing Group Differences in the Iowa Gambling Task: A Case Study of Intuitive and Deliberate Decision Makers

This chapter has been submitted for publication as:
Helen Steingroever, Thorsten Pachur, Martin Šmíra, and Michael D. Lee (2016). Bayesian techniques for analyzing group differences in the Iowa gambling task: A case study of intuitive and deliberate decision makers.

Abstract

The Iowa gambling task (IGT) is one of the most popular experimental paradigms for assessing real-world decision making. In order to understand the psychological processes that underlie IGT performance, several cognitive models have been developed. For comparing how groups of individuals make decisions on the IGT, the standard approach is to estimate the maximum likelihood model parameters for each person, and then conduct frequentist tests to compare the parameters across groups. Here, we present a Bayesian alternative. In addition to a Bayesian repeated measures ANOVA for comparing behavioral performance, we propose a suite of three complementary model-based methods for assessing the cognitive variables and processes underlying IGT performance: (1) Bayesian hierarchical parameter estimation, (2) Bayes factor model comparison, and (3) Bayesian latent-mixture modeling. To illustrate these Bayesian analysis techniques, we test the extent to which differences in decision style (i.e., intuitive, affective vs. deliberate, planned) explain differences in IGT performance. Our results suggest that, on a behavioral level, intuitive and deliberate decision makers behave similarly on the IGT, and the modeling analyses consistently show that these behavioral preferences are driven by similar cognitive processes. Thus, our results challenge the notion that individual differences in intuitive and deliberative decision styles have a very broad impact on decision making, and that intuitive processes in healthy adults play a central role for IGT performance. The ability to draw such a conclusion demonstrates a major advantage of the Bayesian approach, that is, to be able to quantify evidence in favor of the null hypothesis.
The Iowa gambling task (IGT; Bechara et al., 1994)—a card game that consists of two bad decks (i.e., negative long-term outcomes) and two good decks (i.e., positive long-term outcomes)—is arguably the most popular neuropsychological paradigm for assessing decision-making deficits in clinical populations (Toplak et al., 2010). There is considerable evidence that the IGT performance of healthy decision makers (i.e., participants that do not have any neurological impairments) differs from that of clinical populations such as patients with lesions to the ventromedial prefrontal cortex (Bechara et al., 1998, 1999, 2000), pathological gambling (Cavedini, Riboldi, Keller, et al., 2002), obsessive-compulsive disorder (Cavedini, Riboldi, D’Amnucci, et al., 2002), psychopathic tendencies (Blair et al., 2001), or schizophrenia (Bark et al., 2005, Martino et al., 2007).

These studies have mainly relied on an analysis of the proportion of choices from the good decks as compared to the bad decks, with subsequent conclusions about group differences based on frequentist analysis techniques, such as t-tests and analyses of variance (ANOVAs). In addition, to investigate whether two groups differ in the psychological processes that underlie their performance, several reinforcement-learning (RL) models have been proposed. These models assume that card selection on the IGT results from an interaction of distinct psychological processes such as motivation, memory, and response consistency (Busemeyer et al., 2003). Using these models, it has been possible to infer group differences in cognitive processes even when the behavior of the groups at the aggregate level does not differ (e.g., Yechiam, Kanz, et al., 2008). Popular RL models for IGT data are the Expectancy Valence model (EV; Busemeyer & Stout, 2002; Yechiam, Kanz, et al., 2008) and the Prospect Valence Learning model (PVL; Ahn et al., 2008, 2011; see Steingroever, Wetzels, & Wagenmakers, 2013a, for additional references and a detailed description of the EV and PVL models). More recently, it has been shown that the hybrid PVL-Delta model outperforms the EV and PVL model in many model comparison analyses (Ahn et al., 2008; Fridberg et al., 2010; Steingroever, Wetzels, & Wagenmakers, 2013b; Steingroever et al., 2014, but see also Worthy, Pang, & Byrne, 2013 for the Value-Plus-Perseveration model).

The current standard approach for comparing model parameters between groups is to estimate the parameters for each participant separately using maximum likelihood, and then to use frequentist statistical tests, such as independent-samples t-tests, Friedman tests, or Mann-Whitney U tests, to compare the estimates across groups (e.g., Cella et al., 2012; Escartin et al., 2012; Yechiam, Kanz, et al., 2008; Yechiam, Hayden, et al., 2008). However, individual-level maximum likelihood results in inferior parameter inferences compared to Bayesian hierarchical parameter estimation (Ahn et al., 2011; Scheibehenne & Pachur, 2015; Shiffrin et al., 2008; Wetzels, Vandekerckhove, et al., 2010). In addition, there are several well-known problems inherent with frequentist tests, such as p-values overstating the evidence against the null hypothesis (Berger & Delampady, 1987; Edwards, Lindman, & Savage, 1963; V. E. Johnson, 2013; Sellke, Bayarri, & Berger, 2001), and that classical hypothesis testing cannot be used to quantify evidence in favor of the null hypothesis. The latter is a crucial disadvantage, given that many theories predict the absence of an effect (e.g., Gallistel, 2009; Rouder, Speckman, Sun, Morey, & Iverson, 2009). Finally, in contrast to Bayesian sequential testing, frequentist sequential testing is much less flexible, since it requires researchers to specify in advance the total duration of the data collection period (e.g., Reboissin, DeMets, Kim, & Lan, 2000), and the number of interim analyses (e.g., Pocock, 1977).

Here, we present a Bayesian approach to examine whether two groups differ in their IGT performance, encompassing both behavioral and model-based analyses. We illustrate the Bayesian approach by comparing IGT performance of decision makers with an intuitive (affective) decision style to those with a deliberate (planned) decision style, distinguished via established self-report instruments to measure decision style. This comparison is of theoretical interest because the prominent somatic marker hypothesis (Bechara et al., 1997; A. R. Damasio, Tranel, & Damasio, 1991; A. Damasio, 1994) suggests that intuitive processes are of particular importance for successful
9.1. The IGT and PVL-Delta Model

The IGT

In this section we describe the IGT (see also Steingroever, Wetzels, Horstmann, et al., 2013; Steingroever, Wetzels, & Wagenmakers, 2013a, 2013b; Steingroever et al., 2014, 2016). In the traditional version of the task, participants are initially given $2000 (hypothetically) and are presented with four decks of cards with different payoffs. Participants are instructed to choose, over several rounds, 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 losses (if any) associated with that card, as well as their running tally of rewards and losses over all trials so far.

Table 9.1: Summary of the payoff scheme of the traditional IGT as developed by Bechara et al. (1994).

<table>
<thead>
<tr>
<th>Deck A</th>
<th>Deck B</th>
<th>Deck C</th>
<th>Deck D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad deck with frequent losses</td>
<td>Bad deck with infrequent losses</td>
<td>Good deck with frequent losses</td>
<td>Good deck with infrequent losses</td>
</tr>
<tr>
<td>Reward/trial</td>
<td>100</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>Number of losses/10 cards</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Loss/10 cards</td>
<td>−1250</td>
<td>−1250</td>
<td>−250</td>
</tr>
<tr>
<td>Net outcome/10 cards</td>
<td>−250</td>
<td>−250</td>
<td>250</td>
</tr>
</tbody>
</table>

A crucial aspect of the IGT is whether and to what extent participants eventually learn to prefer the good decks because only choosing from the good decks maximizes their long-term net outcome. The good decks are typically labeled as decks C and D, whereas the bad decks are labeled as decks A and B. Table 9.1 presents a summary of the traditional payoff scheme as developed by Bechara et al. (1994). This table illustrates that decks A and B yield high constant rewards, but even higher unpredictable losses: hence, the long-term net outcome is negative. Decks C and D, on the other hand, yield low constant rewards, but even lower unpredictable 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: decks A and C yield frequent losses, while decks C and D yield infrequent losses.

The PVL-Delta Model

In this section, we describe the PVL-Delta model (see also Steingroever, Wetzels, & Wagenmakers, 2013b; Steingroever et al., 2016). The model formalizes participants' performance on the IGT through the interaction of four parameters that have natural psychological interpretations as controlling different psychological processes (Ahn et al., 2008; Fridberg et al., 2010; Steingroever et al., 2014).

The first assumption of the PVL-Delta model is that, after choosing a card from deck \( k \in \{1, 2, 3, 4\} \) on trial \( t \), participants evaluate the net outcome associated with the card using prospect theory (Tversky & Kahneman, 1992). Formally, the utility is given by

\[
 u_k(t) = \begin{cases} 
 X(t)^A & \text{if } X(t) \geq 0 \\
 -w \cdot |X(t)|^A & \text{if } X(t) < 0 
\end{cases}
\]

(9.1)

In this equation, \( X(t) \) represents the net outcome on trial \( t \), which is the sum of the experienced reward and loss (i.e., \( X(t) = W(t) - |L(t)| \)). The prospect utility function contains the first two model parameters. These are the loss aversion parameter \( w \in [0, 5] \), and the outcome sensitivity parameter \( A \in [0, 1] \).

The loss aversion parameter \( w \) quantifies the relative weight of net losses relative to net gains in participants' evaluation of the net outcome of a given card. A value of \( w \) greater than one indicates a larger impact of negative than of positive net outcomes, whereas a value of \( w \) approaching one indicates a similar impact of negative and positive outcomes. As \( w \) approaches zero, the model predicts that negative net outcomes will be neglected.

The outcome sensitivity parameter \( A \) quantifies the extent to which the subjective utility corresponds to the actual value, \( X(t) \). As \( A \) approaches one, the subjective utility \( u_k(t) \) increases in proportion to the actual net outcome. For values of \( A \) smaller than one, there is less differentiation between the positive and negative net outcomes. As \( A \) approaches zero, the sensitivity to differences in the positive and negative net outcomes continues to decrease towards the limit in which there is no sensitivity.

The PVL-Delta model also assumes that, having formed the utility of the card through Equation 9.1, people update their expected utility of the just-chosen deck, but keep the expected utilities of the remaining decks unchanged. This updating process is described by the delta learning rule:

\[
 Ev_k(t) = Ev_k(t - 1) + a \cdot (u_k(t) - Ev_k(t - 1)).
\]

(9.2)

The delta learning rule states that the expected utility of the chosen deck \( k \) is adjusted upward if the experienced utility \( u_k(t) \) is higher than expected. 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 an updating parameter \( a \in [0, 1] \). This parameter expresses the memory for past expectancies. 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.

In the next step, the PVL-Delta model assumes that the expected utilities of each deck guide participants’ choices on the next trial. This assumption is formalized by the softmax choice rule,\(^1\)

\[ Ev_k(0) = 0 \forall k. \]
also known as the ratio-of-strength choice rule (Luce, 1959). The PVL-Delta model uses this rule to compute the probability of choosing each deck on each trial (Equation 9.3). The softmax choice rule includes a sensitivity parameter \( \theta \) that controls the extent to which trial-by-trial choices match the expected deck utilities. Values of \( \theta \) close to zero indicate random choice behavior (i.e., strong exploration), whereas large values of \( \theta \) indicate choice behavior that is strongly determined by the expected utilities (i.e., choices strictly follow the expectancies of the decks).

\[
P[S_k(t + 1)] = \frac{e^{\theta \cdot E_{v_k}(t)}}{\sum_{j=1}^{4} e^{\theta \cdot E_{v_j}(t)}} \tag{9.3}
\]

The PVL-Delta model assumes a trial-independent sensitivity parameter \( \theta \), which depends on the final model parameter: the response consistency \( c \in [0, 5] \) (Equation 9.4). Small values of \( c \) lead to a small values of sensitivity \( \theta \) and thus to more random choice patterns, whereas large values of \( c \) lead to larger values of \( \theta \), and thus to more deterministic choice patterns.

\[
\theta = 3^c - 1 \tag{9.4}
\]

In sum, the PVL-Delta model has four parameters: (1) an outcome sensitivity parameter \( A \), which determines the shape of the utility function, (2) a loss aversion parameter \( w \), which quantifies the weight of net losses over net rewards, (3) an updating parameter \( a \), which determines the memory for past expectancies, and (4) a response consistency parameter \( c \), which determines the balance between exploration and exploitation in the deck choices.

**Bayesian Hierarchical Implementation of the PVL-Delta Model**

We used a Bayesian hierarchical implementation of the PVL-Delta model for the three cognitive modeling analyses (see Steingroever, Wetzels, & Wagenmakers, 2013b, for more details on the implementation, and Wetzels, Vandekerckhove, et al., 2010, for the same model specification in the case of the EV model). In a Bayesian hierarchical framework, the parameters of individual subjects of a specific group are assumed to stem from a group-level distribution. The Bayesian hierarchical framework thus naturally incorporates both the differences and commonalities between and within the participants of one group, and produces both inferences about individual-level and group-level parameter (Horn et al., 2015; Lejarraga et al., 2016; Navarro et al., 2006; Rouder & Lu, 2005; Rouder et al., 2005, 2008). To confirm that we correctly implemented the PVL-Delta model, we ran several parameter-recovery studies. The results of two such studies are presented in Appendix G.

In Bayesian parameter estimation, inferences about a parameter are based on the posterior distribution of the parameter values given the observed data. A posterior distribution expresses the uncertainty about the value of a parameter based on the modeling assumptions and the observed data.

In the Bayesian framework, Bayes factors are used to choose between models and to test hypotheses (Berger & Mortera, 1999; Edwards et al., 1963; Jeffreys, 1961; Kass & Raftery, 1995; Rouder, Morey, Speckman, & Province, 2012; Rouder et al., 2009; Wagenmakers, 2007; Wagenmakers et al., 2010; Wetzels, Raaijmakers, Jakab, & Wagenmakers, 2009). The Bayes factor quantifies the relative probability of the data under two competing models or hypotheses. In particular, \( BF_{01} \) quantifies the probability of the data under the null hypothesis (\( H_0 \)) relative to the probability of the data under the alternative hypothesis (\( H_1 \)). A Bayes factor can, for example, be used to quantify the evidence that the data provide for a model that assumes differences in the loss aversion parameter across two groups of decision makers (\( M_1 \)), compared to a model that
assumes no differences ($M_0$). If, for example, it was found that BF$_{01} = 10$, this would indicate that the data were 10 times more likely under $M_0$ than under $M_1$. Alternatively, if it were found that BF$_{01} = 1/10$, or equivalently, that BF$_{10} = 10$, this would indicate that the data were 10 times more likely under $M_1$ than under $M_0$. As these possibilities make clear, Bayes factors, in contrast to frequentist methods, allow the evidence for the null hypothesis or null model to be quantified (e.g., Rouder et al., 2009).

9.2 Proposed Methodology for Comparing Groups on the IGT

The IGT has often been used to investigate group differences in decision making. Group differences are interesting because the IGT is assumed to tap into a broad spectrum of distinct psychological processes; by comparing group differences in performance—and in particular by decomposing the behavior using cognitive modeling—there is the potential to identify which processes are different and which are the same. For example, Yechiam and colleagues found that drug and sex offenders over-weighted potential gains as compared with losses, whereas assault criminals tended to make less consistent choices and to focus on immediate outcomes (Yechiam, Kanz, et al., 2008). These findings required the use of a cognitive model because basic data analyses of the card selection behavior did not show substantial differences.

Accordingly, in this section, we present a set of Bayesian statistical analyses that can be applied to compare the performance of groups of people on the IGT. We start with a standard method for behavioral data analysis, before proposing a more novel set of complementary approaches for applying cognitive models.

Behavioral Data Analyses

Basic behavioral data analyses are usually based on general linear models. A standard IGT experiment involves repeated measures for a number of participants in two or more groups over two or more blocks of trials. Accordingly, a Bayesian block x group repeated measures ANOVA on the choices from the good decks (i.e., decks C and D) is appropriate. These sorts of ANOVA analyses can be conveniently performed in JASP (JASP Team, 2015; Rouder et al., 2012), which is user-friendly free software with a graphical user interface for conducting Bayesian data analysis.

Cognitive Modeling Analyses

We implemented all of our proposed model-based analyses using Stan (Stan Development Team, 2014b, 2014a; Hoffman & Gelman, 2014; see Chapter 9 of Stan Development Team, 2014c for a description on how to implement mixture models in Stan).

Bayesian hierarchical parameter estimation

The first model-based analysis involves inferring the posterior distributions of the group-level mean parameters across the two groups. These inferences can be made using the Bayesian hierarchical implementation of the PVL-Delta model described earlier, which assumes that the model parameters of each participant are drawn from a group-level distribution (Steingroever, Wetzels, & Wagenmakers, 2013b). To assess the account of the PVL-Delta model to the data we used the post hoc fit method as described in Steingroever et al. (2014).
Bayes factor model comparison

The second model-based analysis involves comparing the group-level mean parameters. This can be achieved by comparing models that assume differences in the group-level mean parameters across the two groups to a model that assumes no differences in these parameters (i.e., a null model). For models like the PVL-Delta that have more than one parameter of interest, multiple comparisons of this type are needed.

When we refer to a model that assumes differences in at least one group-level mean parameter, we index $M$ by the corresponding group-level mean parameter. $M_{\mu_w \mu_c}$, for example, refers to the model that assumes differences in the group-level mean parameter of the loss aversion parameter $w$ and of the consistency parameter $c$ (i.e., $\mu_{w,1} \neq \mu_{w,2}$ and $\mu_{c,1} \neq \mu_{c,2}$, where the second index refers to the group), but no differences in group-level mean parameter of the outcome sensitivity parameter $A$ and of the updating parameter $a$ (i.e., $\mu_{A,1} = \mu_{A,2}$ and $\mu_{a,1} = \mu_{a,2}$).

Since the PVL-Delta model has four parameters, we compared a total of $2^4 = 16$ models. These 16 models represent all possible combinations of the four group-level mean parameters of the PVL-Delta model that can either be the same or differ across the two groups. Note that, for all of the model comparisons, we assumed that the group-level standard deviations are the same across the two groups (i.e., $\sigma_{A,1} = \sigma_{A,2}$, $\sigma_{w,1} = \sigma_{w,2}$, $\sigma_{a,1} = \sigma_{a,2}$, and $\sigma_{c,1} = \sigma_{c,2}$). To quantify the evidence that the data provide for each of the 16 models, we used Bayes factors under the assumption of equal prior model probabilities of all models. Due to this assumption, the Bayes factor $BF_{01}$ simplifies to the posterior model odds, that is, the ratio of the posterior probability of model $M_0$ relative to the posterior probability of model $M_1$. The posterior probability of a specific model $M$ was estimated by means of the product space method (Carlin & Chib, 1995; Lodewyckx et al., 2011; see Appendix G for more details on the product space method, and see Steingroever et al., 2016, for more details on Bayes factors).

To confirm the stability of the Bayes factor estimates, we undertook several tests. First, we confirmed good sampling behavior of the model indicator variable $z$ (i.e., good mixing and low autocorrelations, that is, frequent model switches; Lodewyckx et al., 2011). Secondly, we repeated the product space method with fewer iterations (i.e., 5,000 samples instead of 7,000 of each chain after having discarded the first 1,000 samples of each chain as burn-in). The stability of the Bayes factor estimates was confirmed because the difference in corresponding estimated posterior model probabilities was smaller than 0.01. Thirdly, the correctness of our Stan model file was discussed on the Stan users mailing list.

Latent-mixture modeling

The first two model-based analyses focus on parameter estimation and model selection, respectively. Though relatively standard approaches in the general Bayesian statistics literature, they are the exception in the context of the IGT and associated cognitive modeling. The third model-based analysis, which combines elements of parameter estimation and model selection in a complementary way, is novel both in the context of the IGT and in Bayesian applications more generally. This analysis involves a two-group latent hierarchical mixture model (Lee et al., 2015; Chapter 6 in Lee & Wagenmakers, 2013).

The goal of the latent-mixture analysis is to infer the group membership of each participant, based on the cognitive model and the data, but without knowledge of the group membership of each participant. Instead, the group membership of each participant is represented by a latent

---

2The discussion can be found here [https://groups.google.com/forum/?hl=cs#!searchin/stan-users/reinforcement/stan-users/TjY5wQqUS2g/cff21Wo8Ur0J](https://groups.google.com/forum/?hl=cs#!searchin/stan-users/reinforcement/stan-users/TjY5wQqUS2g/cff21Wo8Ur0J)
indicator variable, and the analysis infers each participant’s probability of belonging to either of the two groups.

Formally, in the two-group case, group membership is indexed by a binary indicator variable $z_i$ (i.e., $z_i = 0$ and $z_i = 1$ indicate that the $i$-th participant belongs to the first and second group, respectively). The prior for these indicator parameters is $z_i \sim \text{Bernoulli}(\psi)$, so that $\psi \sim \text{Uniform}(0, 1)$ corresponds to the base-rate of membership to the second group. This choice of priors means that each participant is a priori equally likely to be assigned to either group. The latent-mixture model analysis provides a posterior distribution of the base-rate, and, for all participants separately, the probability with which they are assigned to each of the groups.

One way to apply this latent-mixture analysis is to use the same priors for model parameters as used in the first analysis to make group-level inferences. In this case, the inferences made by the latent-mixture analysis about the group membership of each participant reflect how people would be classified without knowing the true memberships. If these inferred group memberships agree with the actual one, the analysis provides strong evidence that the behavioral data and model separate people into the proposed groups.

A second way to apply the latent-mixture model approach is to use highly informative priors from the first analysis, so that each group is defined in terms of group-level parameter inferences based on the true memberships. That is, we assumed that probit transformed individual-level parameters were drawn from a group-level normal distribution: $z'_i \sim \text{N}(\mu_{z'}, \sigma_{z'})$. Note that we use $z_i$ to refer to a specific individual-level parameter of the PVL-Delta model (i.e., $z_i \in \{A_i, w_i, a_i, c_i\}$), and $z'_i$ to refer to its probit transformed version (i.e., $z'_i = \Phi^{-1}(z_i)$ with $\Phi^{-1}$ being the inverse of the cumulative standard normal distribution function). We assigned a normal prior to the group-level means $\mu_{z'}$, and a truncated normal prior (allowing for only positive values) to the group-level standard deviations, $\sigma_{z'}$. These prior distributions were characterized by means and standard deviations obtained from the first analysis. That is, we used the mean and the standard deviation of the posterior distribution of $\mu_{z'}$ obtained from the first analysis to specify the prior distribution on $\mu_{z'}$ in the informed latent-mixture model approach, and analogous for the prior distribution on $\sigma_{z'}$. This analysis obviously re-uses the behavioral data, and so cannot be used to make inferences about model parameters. It does, however, potentially provide a strong test of patterns of group membership. In particular, if the true group memberships of participants cannot be inferred under these ideal conditions, there is strong evidence that the model and data do not distinguish the participants into the proposed groups.

### 9.3 Case Study: Intuitive versus Deliberate Decision Making

Whereas many early applications of the IGT focused on comparing clinical to control groups, the task has increasingly also been used to study how individual differences in cognitive abilities (i.e., executive functions, intelligence), state mood, or personality characteristics among healthy participants can explain differences in decision making (Buelow & Suhr, 2009; Suhr & Tsanadis, 2007; Toplak et al., 2010). One interesting variable whose impact on decision behavior has recently received much attention is decision style (e.g., Phillips, Fletcher, Marks, & Hine, 2016), which measures whether people prefer making decisions using an intuitive or a deliberate decision mode (Betsch, 2004).

Thus, we considered possible IGT differences between intuitive and deliberative decision makers as a case study for our proposed methodology. The IGT seems a particularly promising context for this purpose because Bechara et al. (1997) proposed that intuitive, affective processes are important for good performance on this task. Even though Betsch (2004) showed that intuitive
9.3. Case Study: Intuitive versus Deliberate Decision Making

and deliberative decision styles are linked to behavioral differences in many decision tasks, such as the valuation of consumer items and monetary lotteries \citep{SchunkBetsch2006, BetschKunz2008}, we believe ours is the first study that investigates whether decision style also impacts IGT performance. Different predictions about the nature of an effect are conceivable. Based on the somatic maker hypothesis it is reasonable to expect that affective, intuitive processes lead to successful IGT performance and that intuitive and deliberate decision makers might thus also differ with regard to the processes underlying their card selection \citep{Becharaetal1997, ADamasioetal1991, ADamasio1994}. The results presented by \cite{MaiaMcClean2004}, however, suggest that consciously accessible, deliberate processes are crucial for good IGT performance.

Data

Seventy students from the University of Basel (average age 24.9, $SD = 5.8$, range = 19 – 51 years, 49 female) participated in the study. To measure individual participants’ decision style, we used a self-report inventory complied by \cite{BetschIannello2010} consisting of 12 subscales. Based on the mean score for each participant on each subscale, we conducted a principal component analysis with a rotation based on the varimax method. The Kaiser criterion suggested a three-factor solution (i.e., a deliberation factor, an intuition factor, and a spontaneity factor). Following previous research \citep{BetschKunz2008}, we classified participants as intuitive if they had both a factor score above the median of the intuition factor and a factor score below the median of the deliberation factor. Participants with the opposite pattern were classified as deliberate. This classification scheme yielded 19 participants in the intuitive group and 19 participants in the deliberate group. Thirty two participants thus remained unclassified. Using an alternative classification, which included all participants, and simply distinguished between intuitive and deliberate decision makers based on whether or not their score on the intuition factor was higher than the median, resulted in qualitatively identical conclusions. Appendix G provides more details.

Behavioral Data Analyses

In order to obtain a visual impression of the group-level deck preferences across trials, Figure 9.1 shows the proportion of choices from each deck as a function of 10 blocks, and the proportion of choices from the good and bad decks, separately for intuitive and deliberate decision makers. The figure suggests similar deck preferences for both groups. Specifically, although both groups failed to develop a clear avoidance of bad deck B, overall they learned to make more choices from the good decks than from the bad decks. There appears to be a slight trend for stronger learning in the group of intuitive decision makers.

We applied our proposed Bayesian data analysis, in the form of a 10 (block) x 2 (decision style) repeated measures ANOVA. The results of this analysis showed that the data are 3.64 times more likely under the null model that assumes no group differences in the number of choices from the good decks than under the alternate model that does assume group differences (i.e., the Bayes factor is 3.64 in favor of the model that includes no main effect of group). According to the classification scheme of \cite{Jeffreys1961}, this can be considered as moderate evidence for the null model. In addition, the data are about five times more likely under the model that assumes that there is no interaction between block and decision style than under the model that assumes that there is such an interaction effect (i.e., the Bayes factor is 5.38 in favor of the model that includes no interaction effect between block and decision style). This can also be classified as moderate

\footnote{The frequentist repeated measures ANOVA revealed that neither the main effect of decision style ($F(1,36) = .404, p = .529$) nor the interaction between block and decision style ($F(9,324) = 1.466, p = .159$) was significant.}
9. Bayesian Techniques for Analyzing Group Differences in the Iowa Gambling Task: A Case Study of Intuitive and Deliberate Decision Makers

Figure 9.1: Mean proportion of choices from each deck within 10 blocks of both groups of decision makers (first column). Each block contains 10 trials. The second column shows the mean proportion of choices from the good and bad decks within 10 blocks of both groups of decision makers. The third column shows the predictions of the PVL-Delta model for both groups of decision makers. The predictions were obtained by computing the mean probabilities of choosing the good decks or the bad decks on each trial according the post hoc absolute fit method (see Steingroever et al., 2014).
evidence for the null model [Jeffreys, 1961]. These results suggest that deliberate and intuitive decision makers show similar learning curves on the IGT.

### Cognitive Modeling Analyses

Even though the behavioral data analysis suggests that intuitive and deliberate decision makers show similar deck preferences on the IGT, it might still be the case that there are differences in the underlying cognitive processes driving the decisions of the two groups. To investigate this possibility, we next decompose the IGT performance of the two groups using three different cognitive modeling analyses.

In each of the three cognitive modeling analyses, we used random starting values, and run at least three HMC chains. We collected 4,000, 7,000 and 9,000 samples of each chain after having discarded the first 2,000, 1,000 and 1,000 samples of each chain as burn-in in the case of first, second, and third analysis, respectively. Visual inspection of the chains and confirmation that all parameters had \( R \) values below 1.05 suggested that the collected samples provided a valid approximation to the joint posterior parameter distribution.

**Bayesian hierarchical parameter estimation**

Before interpreting the model parameters, we first assessed whether the PVL-Delta model sufficiently accounts for the data of both groups using the post hoc absolute fit method (see Steingroever et al., 2014). The post hoc fit performance of the PVL-Delta model is presented in the third column of Figure 9.1. Comparing the second and third columns of Figure 9.1, it is apparent that the PVL-Delta model nicely captures the qualitative choice pattern in both groups. In particular, as the task proceeds, the model predicts that both groups learn to make more choices from the good decks, and that intuitive decision makers make slightly more choices from the good decks. The PVL-Delta model thus captures key trends in the data, and therefore provides a sufficient account to the data of both groups allowing for meaningful conclusions from the model parameters.

Figure 9.2 shows the posterior distributions of the group-level mean parameters of the PVL-Delta model, separately for the intuitive and the deliberate decisions makers. The posterior distributions show that deliberate decision makers tend to have a higher outcome sensitivity parameter \( \mu_A \) (i.e., a better correspondence between the objective and the subjective utilities of the decks), but a lower updating parameter \( \mu_a \) (i.e., less forgetting and weaker recency effects) than intuitive decision makers. In addition, the posterior distributions suggest that the groups differ neither on the loss aversion parameter \( \mu_w \) nor on the choice consistency parameter \( \mu_c \). Note that these conclusions are based only on a visual comparison of the posterior distributions.

**Bayes factor model comparison**

In this section, we discuss the results of the Bayes factor model comparison. We start by discussing the posterior model probabilities, and then derive Bayes factors using this formula: 

\[
BF_{\Omega,abcd} = \frac{\hat{p}(M_{\Omega}|D)}{\hat{p}(M_{abcd}|D)},
\]

that is, the ratio of the posterior model probability of model \( \Omega \) and \( M_{abcd} \). Tables 9.2 and 9.3 show the posterior model probabilities for eight of the models under the assumption of equal prior model probabilities of all models. The posterior model probabilities of the remaining models are below 0.05 and are not shown. The posterior probability of a specific

---

4The Bayes factors discussed in this article are based on unrounded posterior model probabilities and may therefore slightly differ from Bayes factors calculated using the posterior model probabilities presented in Table 9.2.
Figure 9.2: Posterior distributions of the group-level parameters of both groups obtained from fitting the PVL-Delta model to the data of each group separately.

The model quantifies the evidence that the data provide for that model relative to all other models under consideration (i.e., 15 alternative models). From the tables it is evident that the data provide most evidence for the null model $M_{\Omega}$, which assumes no differences between intuitive and deliberate decision makers in the group-level mean parameters. The evidence for the null model is weakest when it is compared to the model that assumes differences between intuitive and deliberate participants in outcome sensitivity parameter (i.e., model $M_{\mu_A}$; $BF_{\Omega,\mu_A} = 1.36$) and the model assuming differences in the updating parameter (i.e., model $M_{\mu_a}$; $BF_{\Omega,\mu_a} = 1.23$). According to Jeffreys (1961), the evidence for the null model compared to these two models can be characterized as anecdotal. When compared to model $M_{\mu_w}$ (i.e., the model that assumes differences in the loss aversion parameter), the Bayes factor analysis suggests that there is about three times as much evidence for the null model ($BF_{\Omega,\mu_w} = 2.84$); according to Jeffreys (1961), this level of evidence is also anecdotal. In addition, the data provide moderate evidence for the null model compared to model $M_{\mu_c}$ (i.e., the model that assumes differences in the consistency parameter; $BF_{\Omega,\mu_c} = 6.31$). These findings are consistent with Figure 9.2, where the largest differences in the posterior distributions were on the group-level mean of the outcome sensitivity parameter and the updating parameter; the group-level means for the loss aversion parameter and the consistency parameter had posterior distributions that were highly overlapping.
Table 9.2: Posterior model probabilities of the null model and models that assume differences in only one group-level mean parameter under the assumption of equal prior model probabilities. The posterior model probabilities of models that are neither shown in this table nor in Table 9.3 are less than .05.

| \( \hat{p}(M_\Omega | D) \) | \( \hat{p}(M_{\mu_A | D}) \) | \( \hat{p}(M_{\mu_w | D}) \) | \( \hat{p}(M_{\mu_a | D}) \) | \( \hat{p}(M_{\mu_c | D}) \) |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0.20            | 0.15            | 0.07            | 0.16            | 0.03            |

Table 9.3: Posterior model probabilities of models that assume differences in two group-level mean parameters under the assumption of equal prior model probabilities. The posterior model probabilities of models that are neither shown in this table nor in Table 9.3 are less than .05.

| \( \hat{p}(M_{\mu_A \mu_w | D}) \) | \( \hat{p}(M_{\mu_A \mu_a | D}) \) | \( \hat{p}(M_{\mu_w \mu_a | D}) \) |
|-----------------|-----------------|-----------------|
| 0.05            | 0.12            | 0.06            |

When comparing the null model to models that assume differences in two parameters as in Table 9.3, the null model is generally more strongly supported by the data than when compared to models that assume differences in only one parameter as in Table 9.2. In particular, the data provide anecdotal evidence for the null model compared to the model that assumes differences in both the outcome sensitivity and the updating parameter (i.e., model \( M_{\mu_A \mu_w} \)), and moderate evidence for the null model compared to models \( M_{\mu_A \mu_w}, M_{\mu_w \mu_a}, M_{\mu_A \mu_c}, M_{\mu_w \mu_c}, \) and \( M_{\mu_A \mu_w \mu_a} \), respectively. For all of the other model comparisons the Bayes factors are greater than 11, suggesting strong evidence for the null model. Thus, our model selection analyses of the data suggest that it is very unlikely that the two groups differ in three or more parameters.

In sum, out of all of the models considered, the null model—that is, the model that assumes no differences in the group-level mean parameters of the intuitive and deliberate decision makers—received most support. Out of all of the models that do assume differences in the group-level mean parameters between the two groups, the model that assumes differences in the outcome sensitivity and the update parameter, respectively, received the most evidence.

**Latent-mixture modeling**

Figure 9.3 shows the posterior means of the \( z_i \) variables for each participant. Since these expectations are naturally interpreted as group membership probabilities, a low posterior mean of \( z_i \) suggests that participant \( i \) is very likely to belong to the group of deliberate decision makers, whereas a large value suggests that that participant is very likely to belong to the group of intuitive decision makers. According to the group membership established with the decision-style inventory, participants 1–19 were classified as deliberate decision makers (i.e., unfilled bars), whereas participants 20–38 were classified as intuitive decision makers (i.e., grey bars). The horizontal line represents a posterior classification probability of 0.5.

If deliberate versus intuitive decision style has a crucial impact on IGT performance, the latent-mixture model should make inferences consistent with the group membership according to the decision-style inventory. Specifically, for participants 1–19 the posterior mean of the \( z_i \) variable should be below the horizontal line, whereas it should be above this line for participants 20–38. However, it is evident in Figure 9.3 that the group membership inferred from the latent-mixture modeling analysis does not coincide with the ground truth distinction between intuitive and deliberate decision makers.
9. Bayesian Techniques for Analyzing Group Differences in the Iowa Gambling Task: A Case Study of Intuitive and Deliberate Decision Makers

Figure 9.3: Posterior classification as belonging to the group of intuitive decision makers. According to the inventories, participants 1-19 were classified as deliberate decision makers (i.e., white bars), whereas participants 20-38 were classified as intuitive decision makers (i.e., grey bars). The horizontal line represents a posterior classification of .5.

9.4 Discussion

We presented a Bayesian approach for analyzing whether two groups differ in their behavior on the IGT, and for using cognitive models to understand whether their behavior is driven by different psychological processes. For the latter goal, we used three complementary analyses to “triangulate” the research question: Bayesian hierarchical parameter estimation, Bayes factor model comparison, and latent-mixture modeling (see also Lee et al., 2015).

We illustrated this Bayesian approach with a comparison of the choice behavior of intuitive and deliberate decision makers on the IGT. This comparison is interesting because Bechara et al. (1997) proposed that intuitive, affective processes are important for good performance on this task. In addition, intuitive versus deliberate decision style has been found to be linked to behavioral differences in several decision tasks, such as valuation of consumer items and monetary lotteries (Schunk & Betsch, 2006; Betsch & Kunz, 2008). To our knowledge, however, this is the first study that investigates whether decision style also impacts IGT performance.

The application of our Bayesian analysis approach to data from intuitive and deliberate decision makers revealed that, on a behavioral level, intuitive and deliberate decision makers show similar deck preferences on the IGT. Our Bayesian modeling techniques consistently revealed that similar cognitive processes drive performance of intuitive and deliberate decision makers on the IGT. The three different ways of formalizing the basic research question resulted in consistent findings, and, in our view, permit stronger conclusions than could be made based on any one approach individually.

Methodological Contribution

Even though the Bayes factor is “the standard Bayesian solution to the hypothesis testing and model selection problems” (Lewis & Raftery, 1997, p. 648), to our knowledge this is the first time that using Bayes factors is proposed to compare IGT performance at both the behavioral level
and using cognitive models. The current standard approach to analyze IGT data is to rely on frequentist methods. In the extensive reviews by Sevy et al. (2007) and Toplak et al. (2010) many non-significant results are reported, not allowing any insightful conclusions, since it can only be concluded that the null hypothesis cannot be rejected. For example, given an alternative hypothesis that states that two groups differ in the number of choices from the good decks as opposed to a null hypothesis that states that there is not such a difference, and a $p$-value larger than 0.05, one cannot conclude that both groups choose equally often from the good decks. This disadvantage does not hold for the Bayesian approach. The Bayes factor allows an inference about whether the data are informative enough to draw strong conclusions, and, in the case of informative data, further provides an inference about the relative probability of the data under the two competing hypotheses (for more advantages on the Bayesian approach see for example Rouder et al. 2009; Wagenmakers 2007; Wagenmakers, Lee, Lodewyckx, & Iverson 2008).

Being able to use Bayes factors to compare not only the behavioral performance of two groups (i.e., by means of repeated measures ANOVA), but also to investigate whether two groups differ in PVL-Delta model parameters (i.e., by means of the product space method and latent-mixture modeling) offers a crucial contribution. The current standard in the field of reinforcement-learning models for the IGT is to estimate parameters for each subject separately using maximum likelihood methods, and then use frequentist tests (e.g., independent-samples $t$-tests, Friedman tests, or Mann-Whitney $U$ tests) to compare the parameter estimates across groups (e.g., Cella et al. 2012; Escartin et al. 2012; Yechiam, Kanz, et al. 2008; Yechiam, Hayden, et al. 2008). However, many studies have shown that maximum likelihood procedures can result in inferior parameter estimates compared to Bayesian hierarchical parameter estimation (Ahn et al. 2011; Scheibehenne & Pachur 2015; Shiffrin et al. 2008; Wetzels, Vandekerckhove, et al. 2010). Thus, the possibility to derive Bayes factors to compare model parameters across groups addresses two major shortcomings of the current standard approach. It allows for a better estimation of the model parameters by avoiding maximum likelihood estimates, and it circumvents shortcomings inherent with classical hypothesis testing.

**Theoretical Contribution**

Our results offer important insights relating to the discussion of how intuitive and deliberate processes impact decision making. It has been argued that there are robust differences between decision makers in their tendency to rely on the intuitive and the deliberate system (Betsch 2004), and that these decision styles are linked to behavioral differences in several decision tasks, such as valuation of consumer items and monetary lotteries (Schunk & Betsch 2006; Betsch & Kunz 2008). Our results, however, suggest that a person’s decision style has no substantial bearing on IGT performance and is thus not a key driver behind the substantial individual differences typically observed in IGT performance.

These results are relevant to previous research in several ways. First, they seem to be inconsistent with the somatic maker hypothesis, according to which a stronger reliance on an intuitive decision mode results in better IGT performance because of the crucial role of the emotional, intuitive system for learning to make good decisions on the IGT (A. Damasio 1994). Secondly, they also contradict the hypothesis that deliberate decision makers perform better on the IGT because of a strong association between conscious awareness and choosing from the good decks (Maia & McClelland 2004). Instead, both systems—to the degree that they can be dissociated—seem to equally contribute to performance on the IGT. Finally, our data do not allow one to conclusively answer the question of whether or not deliberate decision makers have a less curved utility function than intuitive decision makers (i.e., a higher outcome sensitivity
parameter; Schunk & Betsch (2006). On the other hand, note that the weak association between decision style and IGT performance might also be due to the way decision styles are typically assessed. While standard decision-style inventories tap into decision making in the rather abstract and domain-general fashion, there is some indication for considerable domain-specificity of decision style (Pachur & Spaar, 2015). As a consequence, domain-general decision style might only weakly predict decision style in a financial risk task, such as the IGT.

If decision style is only little (if at all) associated with performance on the IGT, what other factors might account for individual variability commonly observed? One possibility is that capacities such as working memory, intelligence, and inhibition play a crucial role. On the other hand, although some studies have indeed found IGT performance to be linked to variables such as working memory, inhibition, intelligence, and personality (e.g., Crone, Vendel, & van der Molen, 2003; Demaree, Burns, & DeDonno, 2010; Franken & Muris, 2005; Suhr & Tsanadis, 2007), such links seem to emerge inconsistently and are, overall, rather weak (e.g., Dunn et al., 2006; Toplak et al., 2010). In light of these results, one cannot rule out that the difficulty in explaining individual differences in IGT performance is also due to characteristics of the task itself. Specifically, Schonberg, Fox, and Poldrack (2011) pointed out that due to its complex nature, tapping into multiple psychological processes (learning, evaluation, and search), the IGT mimics real-world decision making closely. However, this complexity could also have the effect that individual differences unfold in very complex ways on the IGT, making it difficult to identify clear associations between properties of patterns in people’s behavior and individual difference measures.

Conclusion

We proposed a set of Bayesian analyses for comparing IGT performance between groups. The application of these techniques to compare decision makers with a deliberate or an intuitive decision style showed not only that both groups of decision makers perform similarly on the IGT, but that their performance is also driven by similar cognitive processes. Our refined analysis approach could easily be adapted to other decision making tasks, and cognitive models of behavior on those tasks. All of the relevant code is available online, and all of the required programs are free to download. Due to the advantages of Bayesian analyses, we encourage using our proposed methodology to investigate group differences in IGT data, or in similar decision-making tasks.

Author Note

The code for all cognitive modeling analyses and the data are available on www.helensteingroever.com. The data are also published in Steingroever, Davis, et al. (2015). Additional results, such as tests that confirm the stability of the Bayes factor estimates, and assessment of absolute model account based on the post hoc absolute fit and simulation method for each deck separately, can be requested from the first author. We also repeated the analyses using an alternative classification, which included all participants (distinguishing between intuitive and deliberate decision makers simply based on whether their score on the intuition factor was higher or lower than/equal to the median); results from this analysis can also be requested from the first author.

Acknowledgements

We thank Noemie J. Eichorn for collecting the data. This research was supported by a Netherlands Organisation for Scientific Research (NWO) grant to HS (404-10-086).
Chapter 10

Using Bayesian Regression to Incorporate Covariates into Hierarchical Cognitive Models

This chapter has been submitted for publication as:
Using Bayesian regression to incorporate covariates into hierarchical cognitive models.

Abstract

An important tool in the advancement of cognitive science are quantitative models that represent different cognitive processes in terms of model parameters. To evaluate such models, their parameters are typically tested for relationships with behavioral and physiological variables that are thought to reflect specific cognitive processes. However, many models do not come equipped with the statistical framework needed to relate model parameters to covariates. Instead, researchers often revert to classifying participants into groups depending on their values on the covariates, and subsequently comparing the estimated model parameters between these groups. Here we develop a comprehensive solution to the covariate problem in the form of a Bayesian regression framework. Our framework can be easily added to existing cognitive models and allows researchers to quantify the evidential support for relationships between covariates and model parameters using Bayes factors. Moreover, we present a simulation study that demonstrates the superiority of the Bayesian regression framework to the conventional classification-based approach.

One of the major goals of cognitive science is to describe how cognition shapes human behavior. Cognitive models are an important tool in this endeavor as they offer a formal account of the relationship between cognitive processes, physiology, and behavior (Turner, Forstmann, Love, Palmeri, Thomas, & van Maanen, in press). More specifically, model parameters are often used to describe latent cognitive processes that are hypothesized to find their overt expression in physiological variables such as heart rate or EEG activity, and behavioral manifestations such as reaction times or confidence ratings (Forstmann, Wagenmakers, Eichele, Brown, & Serences, 2011).
However, cognitive models usually do not come equipped with a principled way of evaluating these hypothesized relationships. Instead of directly modeling and assessing the relationship between latent cognitive processes and measurable covariates at the population level, researchers often resort to a multi-step procedure; they first fit a cognitive model to participants’ behavioral data, then group participants according to the values on a set of covariates, and subsequently test the groups of participants for differences in their estimated model parameters. For instance, Cooper, Worthy, and Maddox (2015) asked participants to fill out the Regulatory Focus Questionnaire (Higgins et al., 2001) which consists of two scales that measure participants’ tendency to either avoid new tasks for fear of failure (prevention focus) or approach new tasks with an anticipation of success (promotion focus). Cooper et al. categorized participants into two groups based on whether they scored higher on the prevention focus scale or on the promotion focus scale. Subsequently, participants performed 250 trials of the Mars Farming task in which they have to choose between two options with choice-dependent rewards. Cooper et al. analyzed their data by first fitting a reinforcement-learning model to participants’ choice data and in a second step used an ANOVA to compare the estimated model parameters between the two groups of participants. Although this multi-step procedure might seem a reasonable way of testing which covariates are related to which model parameters, it is cumbersome and potentially misleading.

There are two major statistical pitfalls associated with a multi-step analysis procedure. First, when researchers fit hierarchical cognitive models to data and subsequently extract parameter estimates for individual participants, they risk overstating the evidence for differences between groups of participants. Such practices have unfortunately become more frequent with the advent of hierarchical Bayesian models in recent years (e.g., Ahn et al., 2014; Chan et al., 2013; Chevalier, Chatham, & Munakata, 2014; Vassileva et al., 2013). We will give a brief intuition of the mechanism underlying this pitfall here (for an in-depth discussion of the problem see Boehm, Marsman, Matzke, & Wagenmakers, submitted). Hierarchical Bayesian methods use information about the group of participants to inform parameter estimates for individuals. Individual estimates that are less reliable are more strongly pulled towards the group mean, a phenomenon known as shrinkage (Efron & Morris, 1977; Gelman et al., 2013; Rouder & Lu, 2005). One side effect of this correction of individual estimates is that their variance is reduced because all estimates are pulled towards a common value. As a consequence, effect sizes and test statistics, which are usually computed as the ratio of differences between group means to the variance within groups, are inflated and consequently overstate the evidence against the null hypothesis. This problem can be easily avoided if differences between groups of participants are incorporated into the hierarchical structure of the cognitive model and statistical tests are carried out at the group-level, instead of applying the model separately to each group of participants and carrying out tests on the participant-level.

The second pitfall associated with the multi-step procedure is the dichotomization of continuous covariates which can lead to biased statistical tests. This problem has been discussed repeatedly in the context of frequentist statistics (Altman & Royston, 2006; Austin & Brunner, 2004; Cohen, 1983; MacCallum, Zhang, Preacher, & Rucker, 2002; Maxwell & Delaney, 1993; Royston, Altman, & Sauerbrei, 2006). Despite these repeated warnings, several authors have recently applied dichotomization of continuous covariates to test for relationships with model parameters (e.g., Cooper et al., 2015; Kwak, Pearson, & Huettel, 2014; Steingroever et al., submitted). The type of bias introduced by such dichotomization-based tests depends on the correlation between covariates; uncorrelated covariates lead to reduced power (i.e., tests missing true relationships between covariates and model parameters) whereas correlated covariates lead to an inflation of the Type I error rate (i.e., tests detecting spurious relationships between covariates and model parameters). Maxwell and Delaney (1993) provide an accessible explanation of the mechanisms underlying these biases which we briefly summarize here. In our explanation of the mechanisms
we will use linear regression as comparison standard which is the correct analysis approach for the scenario with continuous, normally distributed model parameters and covariates considered here.

There are two possible scenarios for the correlation structure among covariates and model parameters that give rise to two different types of bias. In the first scenario, assume a researcher measures two uncorrelated continuous covariates, one of which is correlated with a model parameter whilst the other is not. For example, the researcher might administer a questionnaire with two uncorrelated subscales that measure participants’ preference for deliberate and intuitive decision making, respectively, and ask participants to complete 100 trials of a risky decision-making task. The researcher then fits a reinforcement-learning model with a loss aversion parameter to participants’ choice data, where, in fact, the loss aversion parameter is related to participants’ preference for intuitive decision making, but not to their preference for deliberate decision making. To test for relationships between the model parameter and the covariates, the researcher splits participants’ scores on each subscale into two halves based on, say, the median score of each subscale, and, for each subscale, uses a t-test to compare the parameter values of participants scoring above-median to the values of participants scoring below-median. Panel A of Figure 10.1 illustrates this scenario for the intuition scale which is positively correlated with the loss aversion parameter. The two horizontal lines show the mean parameter values of each group, the black diagonal line is the result of the correct regression analysis. As can be seen, within each group the deviation of most individual data points from the regression line, that is, the error variance, is much smaller than the deviation from the corresponding group mean. Consequently, a t-test for a difference in group means, which is just the ratio of the mean differences to the error variance, will be biased towards the null hypothesis. A t-test of the regression slope, on the other hand, uses the correct estimate for the error variance and will therefore not show such a bias.

For the second scenario, assume a researcher measures two correlated continuous covariates, one of which is correlated with a model parameter whilst the other is not. In our previous example, the deliberate decision-making subscale and the intuitive decision-making subscale might be correlated with each other, and the loss aversion model parameter might be correlated with the intuitive decision-making subscale, but not with the deliberate decision-making subscale. To test for relationships between the model parameter and the covariates, the researcher again splits each subscale into two halves and, for each subscale, uses a t-test to compare the parameter values of participants scoring above-median to the values of participants scoring below-median. In this case the covariate of interest is the deliberation subscale which is not correlated with the loss aversion parameter. Panel B of Figure 10.1 shows a scatterplot of participants’ scores on the two subscales with the intuition scale on the x-axis and the deliberation scale on the y-axis; the dark grey squares indicate the means of both subscales for each group created by splitting the deliberation scale into two halves. As can be seen, the mean value on the intuition scale is higher for one group than for the other. However, because the two subscales are correlated, the two groups also differ in their mean on the deliberation scale, which is correlated with the loss aversion parameter. Therefore, a t-test for a mean difference in the model parameter between the two groups might suggest a relationship between the deliberation scale and the model parameter due to the difference in means on the intuition scale. A regression analysis, on the other hand, avoids this problem because it partials out the correlation between the two covariates before relating the deliberation scale to the model parameter.

It should be clear from the above examples that dichotomization of continuous covariates is a problematic practice and the associated biases can be easily avoided by using an appropriate regression analysis. The goal of the present work is therefore to develop a hierarchical regression framework for cognitive models that allows researchers to directly relate model parameters to covariates. Specifically, we will use a hierarchical Bayesian approach as it allows for a principled
10. Using Bayesian Regression to Incorporate Covariates into Hierarchical Cognitive Models

Figure 10.1: Biases in analyzing dichotomized variables. (A) Error variance in t-test based on dichotomization compared to regression analysis. The scatterplot shows the relationship between a covariate and a model parameter (grey dots), the dashed line indicates the median of the covariate, horizontal black lines show the mean parameter values for each group obtained by dichotomization along the median, the grey arrow indicates the resulting error for one data point. The diagonal black line shows the least-squares regression line, the black arrow indicates the associated error. (B) Scatterplot showing the relationship between two correlated covariates (grey dots). The dashed line indicates the median of covariate 2, the dark grey squares show the mean value on both covariates of each group obtained by dichotomizing covariate 2 along the median.

quantification of the evidence for relationships between model parameters and covariates.

Within Bayesian statistics, the principled way of quantifying evidential support for scientific hypotheses is by computing Bayes factors. Bayes factors hold a number of advantages over conventional tests of statistical significance. First, significance tests can only ever reject but never accept the null hypothesis. Bayes factors, on the other hand, can express support for the null hypothesis as well as the alternative hypothesis (Rouder et al., 2009). Second, whilst significance tests force a binary choice upon researchers between rejecting the null hypothesis or remaining in a state of suspended disbelief, Bayes factors allow researchers a graded expression of the evidence for the competing hypotheses provided by their data. Third, conventional significance tests require researchers to commit to a sampling plan before data collection begins and to continue collecting data even if a hypothesis can be confidently rejected or accepted before the full sample has been acquired. Bayes factors, on the other hand, allow researchers to assess the support for competing hypotheses repeatedly during the sampling process and stop collecting data when a hypothesis is supported or rejected to a satisfying degree (Edwards et al., 1963; Kass & Raftery, 1995; Rouder, 2014).

In the next section we develop a Bayesian regression framework that can be easily attached to existing cognitive models and allows researchers to compute Bayes factors for the regression weights relating model parameters to covariates. We subsequently use a simulation study to illustrate the superiority of our Bayesian regression framework compared to a typical dichotomization-based analysis and point out the biases associated with the latter method.
10.1 Regression Framework for Relating Cognitive Model Parameters to Covariates

The regression framework we develop in this section can easily be applied to a wide range of cognitive models, such as multinomial processing trees (Batchelder & Riefer, 1999; Matzke, Dolan, et al., 2015; Riefer & Batchelder, 1988), reinforcement-learning models (Busemeyer & Stout, 2002; Sutton & Barto, 1998), or sequential sampling models (S. D. Brown & Heathcote, 2008; van Ravenzwaaij, Provost, & Brown, in press). As an illustrative example, we use the PVL-Delta model (Ahn et al., 2008; Fridberg et al., 2010; Steingroever, Wetzels, & Wagenmakers, 2013b; Steingroever et al., 2014)—a popular reinforcement-learning model for in the Iowa gambling task (IGT; Bechera et al., 1994). We will therefore first briefly outline the structure of the IGT and give a short summary of the PVL-Delta model and its hierarchical Bayesian implementation before explaining the regression extension of the model.

Iowa Gambling Task and Hierarchical PVL-Delta Model

The IGT is an economic decision-making task that aims to measure decision-making deficits in clinical populations. In the computerized version of the IGT, participants are given an initial credit of $2000 and are presented with four decks of cards, each of which is associated with a characteristic payoff structure. On each trial, participants pick a card and receive feedback about the wins and losses for that card, as well as the running tally. Participants are instructed to choose cards from the decks in a way that maximizes their long-term net outcomes (see Bechera et al., 1994, for more details on the task).

The PVL-Delta model aims to explain the cognitive processes that drive participants’ choices on the IGT. We will focus here on a conceptual description of the model, its parameters, and their interpretation. The PVL-Delta model conceptualizes risky decision making as a three-step process that is governed by four model parameters. First, a loss aversion parameter $w \in [0, 5]$ describes the weighting of net losses relative to net gains, where a value of $w > 1$ means that negative net outcomes impact the subjective utility more strongly than positive net outcomes. Values close to 1 imply equal weighting of net losses and net wins, and as $w$ approaches 0, net losses are increasingly neglected. Second, the outcome sensitivity parameter $A \in [0, 1]$ determines the shape of the utility function. As $A$ approaches 1, the utility function becomes more linear, meaning that the subjective utility of the decks increases proportionally with increasing net outcomes, whereas as $A$ approaches 0, the utility function approximates a step function, meaning that the subjective utility is determined only by the sign of the net outcomes but not their actual value. Third, an updating parameter $a \in [0, 1]$ determines how past expectancies influence the evaluation of the current outcome. A value of $a$ close to 1 indicates quick forgetting and strong recency effects whilst a value of $a$ close to 0 indicates slow forgetting and weak recency effects. Finally, the response consistency parameter $c \in [0, 5]$ determines the relative amount of exploitation vs. exploration, with values close to 0 leading to random choice behavior and larger values leading to more deterministic behavior.

Steingroever et al. (submitted) have presented a Bayesian hierarchical implementation of the PVL-Delta model (solid arrows in Figure 10.2; see also Steingroever, Wetzels, & Wagenmakers, 2013b; and see Wetzels, Vandekerckhove, et al., 2010, for a hierarchical implementation of the related Expectancy Valence model). In their hierarchical implementation of the model, trials $t$ of the IGT (inner plate) are nested within participants $i$ (outer plate). For each trial $t$ of participant $i$ the choice of a deck of cards on the subsequent trial $Ch_{i,t+1}$, and the wins $W_{i,t}$ and losses $L_{i,t}$ on the current trial are observed nodes (grey rectangles); the utility $U_{i,t}$, the expected
utility $E_{v_{k,i,t}}$, the sensitivity parameter $\theta_i$, and the probability of choosing deck $k$ on the next trial $P[S_k(t + 1)]$ are deterministic nodes (double-bordered circles) as they are fully determined by the model equations and parameters. Moreover, the individual-level model parameters $z_i \in \{A_i, v_i, a_i, c_i\}$ are modelled based on their probit-transformation, which means that the individual-level model parameters $z_i$ are treated as deterministic nodes. Their probit-transforms $z'_i$, on the other hand, are stochastic nodes (single-bordered circles) sampled from a group-level normal distribution with mean $\mu_{z'}$ and standard deviation $\sigma_{z'}$. The priors for the group-level parameters are a standard normal distribution $\mu_{z'} \sim N(0, 1)$ and a uniform distribution $\sigma_{z'} \sim U(0, 1.5)$.

**Regression Extension**

Our implementation of the regression model is based on the framework for Bayesian regression analysis suggested by [Liang, Rui, German, Clyde, and Berger (2008)](Liang2008), and [Rouder and Morey (2012)](Rouder2012). The group-level normal priors for the probit-transformed model parameters in Steingroever et al.’s model [submitted] are replaced by a regression extension that relates the individual-level model parameters $z'_i$ to the covariates $j = 1 \ldots P$. Specifically, each individual-level probit-transformed model parameter is sampled from the following normal distribution:

$$z'_i \sim N(\mu_{z'}, \beta_{z'}, \sigma_{z'}^2).$$  \hfill (10.1)

Here $\mu_{z'}$ is the intercept of the regression line, $\beta_{z'}$ is a transposed $P \times 1$ vector of $P$ centered covariate values for participant $i$ (i.e., $\beta_{z'} = [x_{i1} - \bar{x}_1, \ldots, x_{iP} - \bar{x}_P]$ with $\bar{x}_j$ the mean of covariate $j$), $\beta_{z'}$ is the $P \times 1$ vector of conventional regression weights for model parameter $z'$, and $\sigma_{z'}^2$ is the residual variance of the model parameter $z'$. The standardized effect size for covariate $j$ is a transformation of the conventional regression weight:

$$\alpha_{z'j} = \frac{\beta_{z'j} s_j}{\sigma_{z'}},$$  \hfill (10.2)

where $s_j$ is the standard deviation of the covariate. The advantage of this reparameterization is that the standardized effect sizes $\alpha$ can be assigned a (multivariate) Cauchy prior. Cauchy priors have a number of favorable theoretical properties, such as leading to consistent Bayes factors (see [Rouder & Morey, 2012] for a discussion), but are computationally inefficient. To improve computational efficiency, the Cauchy distribution can be expressed as a continuous mixture of normal distributions, which is known as the mixture of g-prior ([Zellner & Siow, 1980](Zellner1980)). Specifically, in the case of a single model parameter and a single covariate there is only one standardized effect size $\alpha_{z'}$ that is assigned a Cauchy prior. To express this Cauchy prior as a mixture of normal distributions, the standardized effect size is assigned a normal prior that depends on a random variable $g$:

$$\alpha_{z'} \mid g \sim N(0, g),$$  \hfill (10.3)

and $g$ is assigned an inverse gamma prior:

$$g \sim IG(1/2, s^2/2),$$  \hfill (10.4)

where the shape parameter of the inverse gamma distribution is set to 1/2 and the scale parameter is set to $s^2/2$. By integrating over $g$, one obtains a Cauchy prior with scale parameter $s$, where $s$ describes the interquartile range of plausible values for $\alpha_{z'}$.

In the more general case of multiple covariates, a multivariate Cauchy prior can be constructed by assigning the vector of regression weights for each model parameter a multivariate normal prior that depends on a random variable $g$ and the matrix of covariate values:

$$\beta_{z'} \mid g \sim N(0, g\sigma_{z'}^2(X^T X/N)^{-1}).$$  \hfill (10.5)
\( \mathcal{N} \) denotes the multivariate normal distribution, the mean of the distribution is set to 0 which is a \( P \times 1 \) vector of 0s, and the variance-covariance matrix is \( g\sigma_{z'}^2(XX'/N)^{-1} \). Moreover, \( X \) is the centered \( N \times P \) design matrix that contains the \( P \) centered covariate values for each of the \( N \) participants. \( X^T \) is the transpose of the centered design matrix, and \( (X^TX/N)^{-1} \) is the inverse of \( (X^XX/N) \). Finally, \( g \) is assigned an inverse gamma prior:

\[
g \sim IG(1/2, s^2/2), \quad (10.6)
\]

where the shape parameter of the inverse gamma is 1/2, and the scale parameter is \( s^2/2 \). The multivariate Cauchy prior that results from integrating over \( g \) has scale parameter \( s \), which again describes the interquartile range of plausible values for \( \alpha \) (see Appendix [H] for details on the prior distribution).

Figure 10.2 shows the graphical implementation of the PVL-Delta model with our regression extension. The model components we added to the Steingroever et al.’s (submitted) hierarchical implementation of the PVL-Delta model are indicated by dashed lines. As in the hierarchical PVL-Delta model, the probit-transformed model parameters are stochastic nodes that are nested within participants. However, in addition to the group-level stochastic entities \( \mu_{z'} \) and \( \sigma_{z'} \), the model parameters also depend on the vector \( \beta_{z'} \) and the observed vector of covariate values \( x_i \), that is nested within participants; the relationship between these entities is given by Equation 10.1. Moreover, the vector \( \beta_{z'} \) depends on the vector of covariate values \( x_i \) via Equation 10.5. In line with Steingroever et al.’s (submitted) implementation of the hierarchical PVL-Delta model, we assigned the intercept \( \mu_{z'} \) a standard normal prior \( \mu_{z'} \sim N(0, 1) \). We assigned the residual variance \( \sigma_{z'}^2 \) an inverse-gamma prior \( \sigma_{z'}^2 \sim IG(2, 1/2) \) with shape parameter 2 and scale parameter 1/2, instead of the uniform prior used in the hierarchical PVL-Delta model. Our choice of a relatively informative prior was mainly made to speed up model convergence (see below). Nevertheless, we also tried a uniform prior which yielded qualitatively identical results. Finally, we assigned the vector of regression weights \( \beta_{z'} \) the mixture of g prior described in Equations 10.5 – 10.6 and set the scale parameter \( s = 1 \). The Stan code for the model can be downloaded from our Open Science Framework folder at [osf.io/6tfz3](osf.io/6tfz3).

**Computing Bayes Factors**

Within the regression framework developed above, researchers can test for a relationship between a normally distributed model parameter, in our case the probit-transformed parameter \( z' \), and a covariate \( x_j \) by computing the Bayes factor for the standardized effect size \( \alpha_{z'j} \). Bayes factors are the standard way of testing hypotheses within the Bayesian framework. They express the relative likelihood of the observed data \( y \) under two competing hypotheses, \( H_0 \) and \( H_1 \) (Rouder et al., 2009):

\[
BF_{01} = \frac{p(y \mid H_1)}{p(y \mid H_0)}, \quad (10.7)
\]

A sensible null hypothesis might be that the standardized effect size for model parameter \( z' \) on the covariate \( x_j \) is 0, \( H_0 : \alpha_{z'j} = 0 \), and the alternative hypothesis might state that the standardized effect size is not 0, \( H_1 : \alpha_{z'j} \neq 0 \), that is, a point-null hypothesis that is nested under the alternative hypothesis. For such hypotheses, the Bayes factor for the parameter in question can conveniently be obtained using the Savage-Dickey density ratio (Dickey & Lientz, 1970; Wagenmakers et al., 2010). According to this test, the Bayes factor is the ratio of the alternative hypothesis’s prior density over its posterior density at the point-null \( BF_{01} = p(\alpha_{z'j} = 0 \mid H_1)/p(\alpha_{z'j} = 0 \mid y, H_1) \).
10. Using Bayesian Regression to Incorporate Covariates into Hierarchical Cognitive Models

\[ \mu_{z'} \sim \mathcal{N}(0, 1) \]
\[ \sigma_{z'}^2 \sim \mathcal{IG}(2, 1/2) \]
\[ \beta_{x'} | g \sim \mathcal{N}(0, \Sigma_{\beta_{x'}}) \]
\[ g \sim \mathcal{IG}(1/2, s^2/2) \]
\[ \Sigma_{\beta_{x'}} = (g\sigma_{z'}^2 X^T X / N)^{-1} \]
\[ z_i' \sim \mathcal{N}(\mu_{z'} + x_i' \beta_{y'}, \sigma_{z'}^2) \]
\[ A_i = \Phi(A'_i) \]
\[ w_i = 5\Phi(w'_i) \]
\[ a_i = \Phi(a'_i) \]
\[ c_i = 5\Phi(c'_i) \]

Figure 10.2: Hierarchical PVL-Delta model with regression extension. Solid lines indicate components of Steingroever et al.’s (submitted) hierarchical implementation of the PVL-Delta model; newly added regression components for relating model parameters to covariates are indicated by dashed lines.

10.2 Simulation Study

The goal of our simulation study is to demonstrate how dichotomizing continuous covariates biases Bayes factors and how these biases can be avoided using the regression framework developed above. To generate realistic data for our simulations, we first fitted the PVL-Delta model with the regression extension to a published data set (Steingroever et al., submitted; see also Steingroever, Davis, et al., 2015). We subsequently used the resulting parameter estimates to generate synthetic data for two scenarios, one in which covariates are not correlated with each other, and one in which covariates are correlated. To emulate a typical dichotomization-based analysis strategy, we applied the PVL-Delta model in combination with a median-split of the covariates to the simulated data. Finally, we compared the resulting Bayes factors from the dichotomization-based analysis to the Bayes factors obtained from the PVL-Delta model in combination with our regression extension.

Generating synthetic data

Data set

We based the setup for our simulated data on the data published in Steingroever et al., Submitted] because of the simple experimental design and the clear structure of the covariates measured. In
Steingroever et al.'s study 70 participants performed 100 trials of the IGT (Bechara et al., 1994). In addition, they completed Betsch and Ianello’s (2010) decision style questionnaire, which consists of 70 items that assess participants’ tendency to use an intuitive or deliberate decision style on a seven-point Likert scale. Steingroever et al. submitted participants’ responses to a principal component analysis and divided participants into two groups depending on their scores on the two factors, deliberation and intuition. In addition, they fitted the PVL-Delta model to the IGT choices from the deliberate and the intuitive decision makers, respectively, and related each participant’s factor scores to the estimated PVL-Delta parameters. A full description of the sample, IGT, and questionnaire data can be found in Steingroever et al. (submitted).

We fitted the PVL-Delta model with the regression extension to Steingroever et al.’s IGT data and used participants’ scores on the deliberation and intuition scales as covariates $x_1$ and $x_2$, respectively. In contrast to Steingroever et al., whose analysis only included the data of participants who scored high on one scale and low on the other, we included the data of all participants in our analysis. As Steingroever et al. reported relatively small effects of the covariates on the model parameters, we expected to also find relatively small standardized effect sizes $\alpha_{z_{ij}}$ and therefore set the scale parameter of the Cauchy prior to $s = 1/3$ (Equation 10.6). To fit the PVL-Delta model to the data, we implemented the model with the regression extension in Stan (Carpenter et al., in press; Stan Development Team, 2016a, 2016b) and obtained samples from the posterior distributions of the model parameters. For each model parameter we ran three MCMC chains and collected 50000 posterior samples per chain. We discarded the first 5000 samples of each chain as burn-in samples and furthermore thinned each chain, discarding 4 out of every 5 samples. Starting values for the population means $\mu_{z_{ij}}$ were randomly drawn from standard normal distributions, starting values for the population standard deviations $\sigma_{z_{ij}}$ were randomly drawn from exponential distributions with scale parameter 1, and starting values for the standardized effect sizes $\alpha_{z_{ij}}$ were randomly drawn from normal distributions with mean 0 and standard deviation 2. All chains were run until convergence (Gelman-Rubin diagnostic $\hat{R} \leq 1.004$, Gelman & Rubin, 1992).

Model fit and generating parameter values

The left half of Table 10.1 shows the estimated posterior means for our fit of Steingroever et al.’s data. As can be seen, the regression weights $\beta_{z_{ij}}$ for the regression of participants’ model parameters on their covariate values are relatively small; the strongest relationships are between the outcome sensitivity parameter $A$ and the deliberation scale, and between the loss aversion parameter $w$ and the intuition scale (i.e., $\beta_{A1} = 0.61$ and $\beta_{w2} = -0.51$, respectively). Nevertheless, compared to the ratio of the standard deviations of the model parameters and the covariates, the regression weights $\beta_{z_{ij}}$ are very small. However, to be able to demonstrate the adverse effects of dichotomizing covariates, we needed to generate data with clearly identifiable relationships between model parameters and covariates (recall that, in the case of uncorrelated covariates, dichotomizing covariates should result in statistical tests missing existent effects). We therefore set $\beta_{A1} = 1$ and $\beta_{A2} = 0$, which means that outcome sensitivity should be associated with deliberation, but not intuition. In addition, we set $\beta_{w1} = 0$ and $\beta_{w2} = -0.9$, which means that loss aversion should be negatively associated with intuition, but not deliberation. Because the regression weights $\beta_{A}$ and $\beta_{w}$ were now larger than the values estimated in our model fit, we needed to reduce the residual variance for the corresponding model parameters to maintain reasonable variance in the covariate scores between participants (compare Equation 10.2). We therefore set the residual variances $\sigma^2_A$ and $\sigma^2_w$ to 3/8 the values estimated in our model fit. The resulting parameter values used to generate data in our simulations are shown in the right half of Table 10.1.
Table 10.1: Posterior estimates of parameter values for Steingroever et al.’s (submitted) data and adjusted parameter values used to generate synthetic data. Subscript 1 indicates effect sizes for the deliberation scale, subscript 2 indicates effect sizes for the intuition scale.

<table>
<thead>
<tr>
<th></th>
<th>Estimated</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_{z'1}$</td>
<td>$\beta_{z'2}$</td>
</tr>
<tr>
<td>$A'$</td>
<td>0.61</td>
<td>0.31</td>
</tr>
<tr>
<td>$w'$</td>
<td>-0.04</td>
<td>-0.51</td>
</tr>
<tr>
<td>$a'$</td>
<td>-0.08</td>
<td>0.24</td>
</tr>
<tr>
<td>$c'$</td>
<td>-0.02</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

Note. Empty cells indicate that the parameter values for generating synthetic data equaled the values estimated for Steingroever et al.’s data.

Data generation

For our simulations we generated 50 synthetic data sets. Each simulated data set consisted of 150 synthetic participants, which should allow our regression analysis to reliably detect relationships between model parameters and covariates. For each participant we generated two covariate values, $x_{i1}$ and $x_{i2}$, as well as one value for each of the four PVL-Delta parameters. To obtain covariate values that were related to a specific model parameter, but not to others, we started by generating a $2 \times 1$ vector of covariate values for each participant from a multivariate normal distribution, $\mathbf{x}_i \sim \mathcal{N}(\mathbf{\mu}, \Sigma)$, with $2 \times 1$ mean vector $\mathbf{\mu} = \mathbf{0}$, and $2 \times 2$ covariance matrix $\Sigma$. In the case of covariates that did not correlate with each other the covariance matrix was the identity matrix. In the case of covariates that were correlated with each other the covariance matrix had diagonal entries 1 and off-diagonal entries 0.7. In a second step, we generated probit-transformed PVL-Delta parameters for each participant using Equation 10.1, $z_i' \sim \mathcal{N}(\mu_{z'}, x_i' \beta_{z'}, \sigma_{z'}^2)$. We set the data-generating group-level parameter values for the regression weights $\beta_{z'}$, mean group-level parameters $\mu_{z'}$, and residual variances $\sigma_{z'}^2$ to the values given in Table 10.1.

Finally, based on the four PVL-Delta parameters, we simulated 200 trials of the IGT for each participant. We doubled the number of trials per participant compared to the data in Steingroever et al.’s (submitted) study to reduce the impact of imprecise estimates of the PVL-Delta parameters on the estimation of the regression weights. To generate simulated IGT trials for each participant, we first spawned a set of payoffs for each deck of cards based on the payoff scheme used in Steingroever et al.’s (submitted) study. We then applied the cumulative distribution function of the standard normal distribution to the probit-scaled model parameters $z_i'$ generated previously to obtain the corresponding PVL-Delta parameter $z$. We furthermore initialized the expected utilities for all four decks of cards to be 0, meaning that the choice of the first card was entirely random for all simulated participants. After generating a random choice on the first trial for each participant, we evaluated the outcome, updated the expected utilities, and generated the participant’s choice on the next trial and the parameter values for that simulated participant. We continued this process iteratively until we had accumulated 200 simulated choices. Further details and the R code used to generate the simulated data can be found at osf.io/6tfz3.

Analysis Using Dichotomized Covariates

Dichotomization-based analysis strategies take several forms. One that is frequently seen in practice is the median-split. In a median-split analysis, participants are divided into two groups based on whether their value on the covariate lies above or below the median. To emulate this analysis
strategy in the context of our simulation study, we developed a version of the PVL-Delta model that estimates separate group-level means $\mu_{z'}$ for participants scoring above the median on one covariate and below the median on another covariate. Note that including these separate group-level means in the model constitutes a relatively sophisticated version of a dichotomization-based analysis; in practice, researchers are more likely engage in a two-step analysis approach, first fitting the cognitive model separately to the groups of participants scoring above and below the median, and subsequently testing the estimated model parameters for differences between groups. However, such a two-step procedure introduces additional biases beyond those introduced by dichotomization which is beyond the scope of the present work.

Our median-split model assumes the same hierarchical structure as the PVL-Delta model, with trials being nested within participants whose probit-transformed parameter values are sampled from a group-level normal distribution. The mean of the group-level distribution from which a participant’s probit-transformed parameter values are drawn depends on the person’s values on the covariates. Therefore, each participant $i$ is assigned a $P \times 1$ vector $d_i$ of dummy variables, where the $j$th entry of the vector is 1 if the person’s score on covariate $j$ is greater than the median, and 0 otherwise:

$$z' \sim N(\mu_{z'} + \delta_{z}^T d_i \sigma_z, \sigma_z^2).$$

Here $\mu_{z'}$ is the mean for a person who scores below-median on all covariates (i.e., the person’s dummy variables are all 0). Furthermore, $\delta_{z'}$ is the transposed $P \times 1$ vector of standardized effect sizes (i.e., $\delta_{z'} = [\delta_{z'1}, \ldots, \delta_{z'P}]'$) and $\delta_{z'1}$ is the standardized effect size indicating the difference, in standard deviations, between the group-level mean for a person with below-median values on all covariates, and the group-level mean for a person with an above-median value on covariate $j$ and below-median values on all other covariates. Finally, $\sigma_z^2$ is the variance of the model parameter $z'$.

As in the PVL-Delta model with the regression extension, we assigned the group-level means $\mu_{z'}$ a standard normal prior $\mu_{z'} \sim N(0, 1)$, and the group-level variance $\sigma_z^2$ an inverse-gamma prior $\sigma_z^2 \sim IG(2, 1/2)$ with shape parameter 2 and scale parameter 1/2. Finally, we assigned the standardized effect sizes $\delta_{z'j}$ independent Cauchy priors $\delta_{z'j} \sim C(1)$ with location parameter 0 and scale parameter 1.

### Data Analysis

We analyzed the simulated data using the PVL-Delta model with regression extension and the median-split version of the PVL-Delta model. For both models we computed Bayes factors contrasting the null hypothesis that there is no relationship between model parameters and covariates with the alternative hypothesis that there is such a relationship. More specifically, in the case of the regression model, the null hypothesis stated that the standardized effect size for a specific model parameter $z'$ on a specific covariate $x_j$ is 0, $H_0 : \alpha_{z'j} = 0$, and the alternative hypothesis stated that the standardized effect size is not 0, $H_1 : \alpha_{z'j} \neq 0$. In the case of the median-split model, the null hypothesis stated that the standardized difference in group means is 0, $H_0 : \delta_{z'j} = 0$, and the alternative hypothesis stated that the difference in group means is not 0, $H_1 : \delta_{z'j} \neq 0$.

We based our computation of the Bayes factors for both models on the Savage-Dickey density ratio [Dickey & Lientz, 1970] [Wagenmakers et al., 2010]. To obtain estimates of the posterior density for each model’s effect size parameters, we first implemented both models in Stan [Carpenter et al., in press] [Stan Development Team, 2016a, 2016b]. As we expected sizable effects in the simulated data, we set the scale parameter for the regression model’s Cauchy prior to $s = 1$. We subsequently ran MCMC chains until convergence (Gelman-Rubin diagnostic $\hat{R} \leq 1.001$, [Gelman & Rubin, 1992]). For each model parameter we ran two MCMC chains and collected 45000
posterior samples per chain. We discarded the first 5000 samples of each chain as burn-in samples and furthermore thinned each chain, discarding four out of every five samples, which left us with a total of 8000 samples per chain. As the prior distribution for the regression weights $\beta_{z'j}$ in the model with regression extension cannot easily be expressed in closed form, we approximated the density of the prior distribution for each regression weight using MCMC sampling with the same setup as for the model fit. We estimated the density of the posteriors and the priors for the $\beta_{z'j}$ using log-spline functions, and computed the exact value of the Cauchy priors for the $\delta_{z'j}$. Finally, we computed the Bayes factors by taking the ratio of posterior densities to the prior densities at 0.

Results

Figure 10.3 shows the log Bayes factors for the alternative hypothesis obtained in our simulations. We chose to plot the log of the Bayes factors here, rather than the Bayes factors, because the Bayes factors spanned up to five orders of magnitude, which means that, on the linear scale, large Bayes factors would obscure differences in Bayes factors at the low end of the scale. Moreover, because we generated our data in such a way that only the PVL-Delta parameters $A$ and $w$ had sizable relationships with the covariates, we will only present the results for these parameters here. The full results for all model parameters as well as details on the estimated effect sizes can be found in Appendix 11. The left panel of Figure 10.3 shows the log Bayes factors for our simulations with uncorrelated covariates. As can be seen, the Bayes factors obtained from the regression analysis showed strong evidence for an effect of the first covariate on the $A$ parameter (dark grey dots, left column in the top row), whereas the median-split analysis provided much weaker evidence for such an effect (light grey dots, left column in the top row). To quantify the difference between Bayes factors on the linear scale, we took the negative reciprocal of Bayes factors favoring the null hypothesis (i.e., $BF_1 < 1 \to -BF_1^{-1}$) and computed the relative difference defined as: $(BF_{MS} - BF_{RG})/|BF_{RG}|$, where $BF_{MS}$ are the Bayes factors from the median-split analysis and $BF_{RG}$ are the Bayes factors from the regression analysis. For the effect of the first covariate on the $A$ parameter, the median relative difference was -0.85, indicating a strong underestimation of the evidence in the median-split analysis. Similarly, the regression analysis strongly supported an effect of the second covariate on the $w$ parameter (dark grey dots, right column in the bottom row), whereas the median-split analysis provided much weaker evidence for such an effect (light grey dots, right column in the bottom row). The median relative difference in Bayes factors was -0.95, indicating a tremendous underestimation of the evidence in the median-split analysis. For the null-effects of the first covariate on the $w$ parameter (right column, top row) and of the second covariate on the $A$ parameter (left column, bottom panel), both analyses performed similarly, with median relative differences in Bayes factors of 0.48 and 0.40, respectively, indicating that the median-split analysis favored the alternative hypothesis slightly more strongly than the regression analysis.

The right panel for Figure 10.3 shows the log Bayes factors for our simulations with correlated covariates. The Bayes factors obtained from the regression analysis again showed stronger evidence for an effect of the first covariate on the $A$ parameter (dark grey dots, left column in the top row) than the median-split analysis (light grey dots, left column in the top row). However, the median relative difference in Bayes factors of -0.62 was much smaller than in the case of uncorrelated covariates. Similarly, the regression analysis provided stronger support for an effect of the second covariate on the $w$ parameter (dark grey dots, right column in the bottom row), than the median-split analysis (light grey dots, right column in the bottom row). The median relative difference in Bayes factors of -0.89 was slightly smaller than in the case of uncorrelated covariates. Unlike in the case of uncorrelated covariates, in the case of correlated covariates the median-split
10.2. Simulation Study

\[ \rho(X_1, X_2) = 0 \quad \rho(X_1, X_2) \neq 0 \]

Figure 10.3: Bayes factors from 50 simulated data sets for the regression and median-split analysis. Data points show the log Bayes factors for the alternative hypothesis (log(BF_{10})) obtained in the regression (RG, dark grey dots) and median-split (MS, light grey dots) analysis for the PVL-Delta model’s A and w parameters (columns) and two covariates (rows). The left subplot shows the results for the case of uncorrelated covariates, the right subplot shows the results for the case of correlated covariates. Lines indicate the mean log BF. Data points are jittered along the x-axis for improved visibility.
10. Using Bayesian Regression to Incorporate Covariates into Hierarchical Cognitive Models

\[ \rho(X_1, X_2) = 0 \]

Figure 10.4: Posteriors of effect sizes for the case of uncorrelated covariates. Shown are the posterior distributions quantile-averaged across 50 simulated data sets. The left subplot shows the results for the \( A \) parameter, the right subplot shows the results for the \( w \) parameter. Thick black lines are the posteriors of the standardized effect sizes \( \alpha \) (left column in each subplot), thick grey lines are the posteriors of the standardized mean differences \( \delta \) (right column in each subplot), thin grey lines show the priors. The top row shows the results for the first covariate \((X_1)\), the bottom row shows the results for the second covariate \((X_2)\).

The biases inherent in the median-split analysis are also clearly visible in the posterior distributions for the effect sizes. Figure 10.4 shows the quantile-averaged posterior distributions of the standardized differences in group means, \( \delta \), and the standardized effect size, \( \alpha \), for the case of uncorrelated covariates. The left column of the top left subplot shows the prior (thin grey line) and the posterior (thick black line) for the regression of the \( A \) parameter on the first covariate. Compared to the prior, which has considerable mass at the point null hypothesis \( \alpha_{A,1} = 0 \), the posterior has nearly no mass at the point null, resulting in Bayes factors that strongly favor the alternative hypothesis. The right column in the same subplot shows the prior (thin grey line) and posterior (thick grey line) for the standardized difference in the \( A \) parameter between participants.
10.2. Simulation Study

\[ \rho(X_1, X_2) \neq 0 \]

Figure 10.5: Posteriors of effect sizes for the case of correlated covariates. Shown are the posterior distributions quantile-averaged across 50 simulated data sets. The left subplot shows the results for the \( A \) parameter, the right subplot shows the results for the \( w \) parameter. Thick black lines are the posteriors of the standardized effect sizes \( \alpha \) (left column in each subplot), thick grey lines are the posteriors of the standardized mean differences \( \delta \) (right column in each subplot), thin grey lines show the priors. The top row shows the results for the first covariate \( (X_1) \), the bottom row shows the results for the second covariate \( (X_2) \).

who score above-median on the first covariate and participants who score below-median. As can be seen, the posterior has little mass at the point null hypothesis \( \delta_{A,1} = 0 \), resulting in Bayes factors favoring the alternative hypothesis. However, compared to the posterior under the regression model, the posterior under the median-split model is considerably wider and has more mass at the point null, which results in the underestimation of the evidence against the null observed above. A comparable pattern can be seen in the bottom right subplot; the posterior under the median-split model has more mass at the point null than the posterior under the regression model, resulting in a strong underestimation of the evidence against the null. Finally, the top right and bottom left subplots show the comparison for the true null-effects of the first covariate on the \( A \) parameter and of the second covariate on the \( w \) parameter, respectively. Although the posterior under the median-split model again has less mass at the point null than the posterior under the regression model, the differences are less pronounced and both models favor the null hypothesis.

Figure 10.5 shows the quantile-averaged posterior distributions of the standardized differences...
Using Bayesian Regression to Incorporate Covariates into Hierarchical Cognitive Models

10

in group means, $\delta$, and the standardized effect size, $\alpha$, for the case of correlated covariates. The top left and bottom right subplots show comparable patterns to the case of uncorrelated covariates; the posterior under the regression and the median-split model both have much less mass at the point null than the respective prior, resulting in Bayes factors favoring the null hypothesis. However, compared to the prior, the posterior under the regression model is much narrower than the posterior under the median-split model, which leads to smaller Bayes factors under the median-split model. Finally, the top right and bottom left subplots show the comparison for the true null-effects of the first covariate on the $A$ parameter and of the second covariate on the $w$ parameter, respectively. As can be seen, the posterior for the regression weights is centered at 0 and has considerably more mass at the point null than the prior. Therefore, Bayes factors under the regression model correctly favor the null hypothesis. However, the posterior under the median-split model lies to the left of the point null for the $A$ parameter and to the right of the point null for the $w$ parameter, and thus has considerably less mass at the point null than the posterior under the regression model. Consequently, Bayes factors under the median-split model understate the evidence for the null and in many instances even support the alternative hypothesis, suggesting spurious associations between the first covariate and the $w$ parameter and between the second covariate and the $A$ parameter.

10.3 Discussion

The goal of the present work was to develop a methodological framework that allows researchers to test hypotheses about associations between the cognitive processes and behavioral and physiological covariates in a principled way. To this end we showed how Bayesian linear regression can be used to relate the parameters of a cognitive model to covariates, and use Bayes factors to quantify the evidential support for specific associations between model parameters and covariates. As an example application, we chose the PVL-Delta model which aims to explain risky decision making as the result of a reinforcement-learning process. Adding a regression extension to the PVL-Delta model allowed us to simultaneously account for participants’ model parameters and measurements of participants’ preferred decision styles.

One major advantage of incorporating a Bayesian regression framework into cognitive models is that it obviates the need for dichotomization-based analysis strategies, circumventing the associated statistical biases. Despite repeated warnings against the use of dichotomization-based analyses (Austin & Brunner, 2004; Cohen, 1983; Maxwell & Delaney, 1993; MacCallum et al., 2002; Royston et al., 2006), a number of recent studies have relied on median splits (e.g., Beitz, Salthouse, & Hasker, 2014; Cooper et al., 2015; Kwak et al., 2014; Steingroever et al., submitted) to test for associations between the parameters of different cognitive models and covariates. We conducted a simulation study to illustrate the degree of bias introduced by such flawed analysis strategies. To this end, we generated simulated data under two scenarios. In one scenario covariates were not correlated with each other, and some of the covariates were correlated with some of the model parameters but not others. Our simulations showed that, in the first scenario, a median-split analysis leads to Bayes factors that underestimate the evidence for true effects compared to the Bayes factors obtained from a regression model. In the second scenario, Bayes factors from a median-split analysis again understated the evidence for true effects but, in addition, a median-split analysis also suggested spurious effects of covariates on model parameters that were, in fact, unrelated.

Interestingly, the Bayes factors for spurious effects suggested by the median-split analysis were...
10.3. Discussion

relatively small compared to the Bayes factors for true effects. This result is most likely due to the
fact that the median-split analysis generally leads to wider posteriors than the regression analysis,
resulting in overall smaller Bayes factors. Surprisingly, for the median-split as well as the regression
analysis, Bayes factors favoring the null hypothesis were very modest across simulations. There are
two probably causes for this result. First, due to the relatively large uncertainty in the estimates
of the model parameters, posteriors were relatively wide compared to the priors, putting relatively
little posterior mass at the point null. Second, the default scale parameter for the priors on the
effect sizes might have been too small compared to the effects in the simulated data, putting too
much prior mass at the point null, thus resulting in small Bayes factors for the null hypothesis.

With the goal of a general-purpose application in mind, one question that inevitably arises is
how easily the Bayesian regression framework used here can be adapted to other cognitive models.
The requirements with respect to the model are relatively modest. First, the cognitive model needs
to be implemented in a hierarchical way to allow researchers to relate individual participants’ model
parameters to measured covariates. For many popular models such hierarchical implementations
are readily available (Matzke, Dolan, et al., 2015; Steingroever et al., 2014; Wiecki, Sofer, & Frank,
2013; Ahn, Haines, and Zhang’s, 2016; R-package contains hierarchical implementations of several
popular models of decision making) or can be easily developed using MCMC software packages
such as JAGS (Plummer, 2003) or Stan (Carpenter et al., in press; Stan Development Team,
2016b). Second, the model parameters of interest need to be normally distributed. Although this
assumption is often reasonable and can be readily adopted, in other cases, specific bounds on the
parameter values are required due to the cognitive interpretation of the parameters or mathematical
constraints. However, such constraints can often be overcome by using transformed versions of the
model parameters, rather than the model parameters themselves, in the regression analysis. In the
case of the PVL-Delta model, for instance, all model parameters are restricted to closed intervals,
yet probit transforming the parameters allowed us to add the Bayesian regression extension to the
PVL-Delta model. These two conditions are all that is required for our regression extension to be
added to a cognitive model and are easily met by most existing models.

Although reinforcement-learning models, and the PVL-Delta model in particular, served
merely as an example for our Bayesian regression framework, we believe that our regression
extension can greatly facilitate research involving risky decision making. One potential application
beyond identifying relationships between model parameters and physiological measurements is the
statistical control of nuisance variables. A number of authors have suggested that performance on
the IGT might be subject to practice effects (Ernst et al., 2003; Lejuez et al., 2003; Verdejo-García
& Pérez-García, 2007), although no study to date has comprehensively addressed this problem
(Buelow & Suhr, 2009). Including time-on-task as a covariate in model-based analyses might allow
researchers not only to control for practice effects but also to pinpoint which cognitive processes are
affected by practice and which processes remain stable over time. Similar model-based analyses
in perceptual decision making, for example, have suggested that whilst participants’ processing
of stimuli remains unaffected by practice, their response mode can change over time although
considerable practice might be needed for participants to reach optimal performance (Hawkins,
Forstmann, Wagenmakers, Ratcliff, & Brown, 2015; Simen et al., 2009).

To conclude, in the present work we presented a hierarchical Bayesian regression extension
for cognitive models that relates participants’ estimated model parameters to behavioral and
physiological covariates. This regression framework allows researchers to test hypotheses about
relationships between model parameters and covariates in a principled manner using Bayes factors.
Moreover, our regression framework overcomes many of the biases associated with alternative
analysis strategies often seen in practice. In particular, we illustrated in our simulation study
the superiority of our regression framework to the often-practiced median-split analysis that can
lead researchers to either miss existing relationships between model parameters and covariates, or suggest spurious associations between model parameters and covariates, depending on whether the covariates are correlated with each other or not. Due to its relatively modest requirements for the cognitive model under consideration, our regression extension can be easily applied to a vast number of existing models and promises to obviate the need for inappropriate, cumbersome, and biased analysis strategies such as the median-split.

**Acknowledgements**

We thank Thorsten Pachur for providing the data used in this article. This research was supported by a Netherlands Organisation for Scientific Research (NWO) grant to UB (406-12-125) and to HS (404-10-086), and a European Research Council (ERC) grant to EJW (283876).
Chapter 11

Summary and Future Directions

This thesis, entitled “Safe Models for Risky Decisions”, scrutinized assumptions about the performance of healthy participants on the Iowa gambling task (IGT; Bechara et al., 1994) and challenged the trustworthiness of conclusions typically obtained from fitting reinforcement-learning models to IGT data. We argued that the risk of drawing premature conclusions from behavioral analyses and computational modeling can be minimized if, in future applications of the IGT and RL models, researchers follow a number of crucial steps. These steps concern behavioral data analyses, model selection, model fitting, and assessment of absolute model fit. In particular, we advocate Bayesian techniques involving Bayesian repeated measures ANOVA for behavioral data analyses, the Bayes factor for model selection, the Bayesian hierarchical framework for model fitting, and posterior predictives to assess the absolute account of the models for the data at hand. Before taking some distance to place the results in broader perspective, I first summarize and discuss the main conclusions of this dissertation.

11.1 Summary of Results

Chapter 2 gave an overview of the performance of healthy participants on the IGT, and pointed to behavioral findings that question key assumptions about IGT performance of healthy participants. Specifically, we showed that (1) healthy participants often fail to develop a pronounced preference for both good decks: Instead, participants often prefer the decks with infrequent losses (i.e., frequency-of-losses effect); (2) healthy participants show idiosyncratic choice behavior (see Figure 11.1); and (3) healthy participants do not show a systematic transition from an initial exploration phase of the decks to an exploitation phase. These findings question the prevailing interpretation of IGT data and suggest that, in future applications of the IGT, key assumptions about performance of healthy participants warrant closer scrutiny.

Chapter 3 used the behavioral findings of Chapter 2, and investigated whether they are in line with the data-fitting potential of three popular reinforcement-learning (RL) models—the Expectancy Valence model (EV; Busemeyer & Stout, 2002), the Prospect Valence Learning model (PVL: Ahn et al., 2008), and a combination of these models, the EV-PU model. However, parameter space partitioning (PSP) revealed important discrepancies between the model-specific popularity and empirical popularity of several choice patterns. In particular, all three models fail to generate pronounced deck preferences often observed in experiments, and the EV model—the first model proposed for the IGT—fails to generate a frequency-of-losses effect that is prominent in healthy participants. This suggests that the EV model should not be used if the data display a
Figure 11.1: Deck selection profiles of five participants express the high within-group variability in the performance of healthy participants. The filled circles indicate the occurrence of rewards and losses together; the empty circles indicate the occurrence of only rewards. The data are published in Steingroever, Wetzels, Horstmann, et al. (2013).
11.1. Summary of Results

Figure 11.2: Simulation performance of the three RL models with respect to the two stylized data sets used in Chapter 5. The first column presents the observed mean proportions of choices from each deck within 10 blocks. Each block contains 10 trials. The second, third, and fourth column present the mean probabilities of choosing each deck on each trial as generated with the EV, PVL, and PVL-Delta model, respectively.

frequency-of-losses effect; in such a case, the EV model will probably provide a poor fit to the data and thus an inaccurate description of the relevant psychological processes. Taken together, our results of Chapter 3 suggest that the three models under consideration have sufficient data-fitting potential for only a restricted number of choice patterns, and a model that can be universally used is still lacking.

Chapter 4 illustrates three important methods for model validation – parameter recovery, parameter space partitioning, and test of specific influence – as applied to the PVL-Delta model, yet another combination of the EV and PVL models. Our results suggest that the PVL-Delta model (1) recovers parameter accurately; (2) accounts for empirical choice patterns featuring a preference for the good decks or the decks with infrequent losses; (3) fails to account for empirical choice patterns featuring a preference for the bad decks; and (4) performs moderately on the test of selective influence that investigates the effectiveness of experimental manipulations designed to target only a single model parameter. In particular, the test of specific influence showed that the manipulations were successful for all but one parameter. To conclude, despite a few shortcomings, the PVL-Delta model seems to be a better IGT model than the popular EV and PVL models.

Chapter 5 exposes the current practice to draw inferences from parameters of RL models without first having sufficiently assessed the absolute fit of the models for the data. We illustrated the importance of assessing absolute model performance using the EV, PVL, and PVL-Delta models and data from two stylized data sets and five published data sets. Our results showed that all models provided an acceptable fit to the data sets (i.e., post hoc fit); however, when the model parameters were used to generate choices, only the PVL-Delta model captured the qualitative patterns in the
11. Summary and Future Directions

data (i.e., simulation performance; see Figure 11.2). Our results highlight that a model’s ability
to fit a particular choice pattern does not guarantee that the model can also generate that same
choice pattern. However, such ability is crucial to ensure meaningful conclusions from the model
parameters. In future applications of the RL models absolute model performance should therefore
carefully be assessed to avoid premature conclusions from the model parameters.

Chapter 6 is a rejoinder to Konstantinidis et al. (2014)’s reply on Chapter 5. In this chapter,
we stressed that the initial goal of Chapter 5 was not to conduct a model selection exercise, but to
illustrate why applied researchers should carefully assess absolute model performance before they
draw conclusions from the estimated parameters. In addition, we elaborated on the advantages
and drawbacks of both the post hoc absolute fit method and the simulation method. A crucial
drawback of the post hoc absolute fit method is that in a strict sense this method does not predict,
but postdicts because it uses the data twice; once to fit the model, and once to generate
“predictions” which are really postdictions. In addition, we pointed out that model selection
criteria based on the post hoc fit method (e.g., the BIC and $G^2$ criteria) do not fully account for
model complexity because they only consider one dimension of model complexity (i.e., the number
of model parameters). Finally, we highlighted the distinction between statistical aspects of model
adequacy (e.g., good fit to the observed data and good predictions for new data) and psychological
relevance of parameter estimates (e.g., good parameter recovery and good performance on tests of
selective influence).

Chapter 7 showed how importance sampling can be used to obtain Bayes factors to compare
Bayesian individual-level implementations of four RL models using data of 771 healthy participants.
In contrast to the BIC and $G^2$ criteria discussed in Chapter 6, the Bayes factor is a model
selection tool that coherently and completely discounts model complexity. Our results provide
strong evidence for the Value-Plus-Persevation (VPP; Worthy, Pang, & Byrne, 2013) model
and moderate evidence for the PVL model, but little evidence for the EV and PVL-Delta
models (see Figure 11.3). However, we pointed out that it is crucial to interpret our results
in combination with results obtained from other model comparison studies in order to obtain a
balanced and comprehensive assessment of model adequacy. For instance, poor parameter recovery
and simulation performance of the VPP model (Ahn et al., 2014) –an eight-parameter model–
suggest that the parameters might have little psychological value, and that a more thorough analysis
of the validity of the VPP model is required.

Chapter 8 built on the insights from Chapter 7 and showed how bridge sampling can be
used to obtain Bayes factors for both an individual-level Bayesian implementation of the EV
model (Figure 11.4) and a hierarchical implementation. We showed that bridge sampling is
not only reliable, accurate, and efficient, but also relatively straightforward to implement. In
addition, bridge sampling is suitable for models in mathematical psychology that are often complex,
non-nested, and hierarchical.

Chapter 9 proposed a suite of three complementary model-based methods for assessing the
cognitive variables and processes underlying IGT performance: (1) Bayesian hierarchical parameter
estimation; (2) Bayes factor model comparison; and (3) Bayesian latent-mixture modeling. To
illustrate these Bayesian analysis techniques, we tested the extent to which differences in decision
style (i.e., intuitive, affective vs. deliberate, planned) explain differences in IGT performance. Our
results suggest that, on a behavioral level, intuitive and deliberate decision makers behave similarly
on the IGT, and the modeling analyses consistently showed that these behavioral preferences are
driven by similar cognitive processes. This chapter offered two major methodological advances: (1)
the use of the product space method to obtain Bayes factors for comparing two groups of decision
makers in a hierarchical Bayesian framework; and (2) the use of a hierarchical mixture model to
infer group memberships (Figure 11.5).
11.2 Future Directions

Figure 11.3: 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, and the boxes indicate the interquartile range (i.e., the distance between the .25 and .75 quantiles).

Chapter 10 proposed a Bayesian regression framework that can be used to extend existing Bayesian hierarchical implementations of RL models. This framework allows researchers to obtain Bayes factors in order to quantify the evidential support for relationships between covariates (e.g., decision style) and model parameters. The advantage of this method is that it avoids the multi-step procedure from Chapter 9 where we divided all participants into two groups depending on their score on the decision style covariate, and subsequently tested the groups of participants for differences in their estimated model parameters.

11.2 Future Directions

Several challenges confront researchers who wish to understand the psychological processes that influence risky decision making. These challenges concern (1) the experimental paradigm that is used to investigate how risky decisions are made; (2) the cognitive models that are used to disentangle the driving processes; and (3) the methods that are used to fit and compare the models. Therefore, to advance our understanding of risky decision making several toeholds are possible.

First, the results of Chapter 2 suggest that we need a better paradigm to investigate how risky decisions are made in a controlled, experimental setting (see also Chiu & Lin, 2007; Dunn et al., 2006; Lin et al., 2007). On the one hand, this goal can be achieved by improving the
Figure 11.4: Comparison of the log marginal likelihoods of 30 participants of Busemeyer and Stout (2002) obtained with bridge sampling (x-axis) and importance sampling (y-axis) reported in Chapter 7. The solid line expresses perfect correspondence of the two methods.

Figure 11.5: Posterior classification as belonging to the group of intuitive decision makers. According to the decision style inventories, participants 1-19 were classified as deliberate decision makers (i.e., white bars), whereas participants 20-38 were classified as intuitive decision makers (i.e., grey bars). The horizontal line represents a posterior classification of .5.
11.2. Future Directions

IGT to become more diagnostic and informative, and a valid measure of risky decision making (Buelow & Suhr, 2009). Possibilities are to (1) use at least 150 trials instead of 100 trials (e.g., Wetzels, Vandekerckhove, et al., 2010); (2) avoid the concurrent presentation of wins and losses, but immediately present the net outcome; (3) change the IGT payoff scheme (for several attempts see Appendix B that presents three different IGT payoff schemes, and Chiu & Lin, 2007); and (4) obtain data from additional measures, such as response time (RT) measures, post-decision wagering, confidence ratings for the choices, or quality ratings for the decks. Another approach is to propose a new risky decision-making task that retains the essential characteristics of the IGT (i.e., incorporation of the tradeoff between immediate rewards, but long-term negative consequences vs. safe options that are profitable in the long-run), and at the same time also accounts for the shortcomings illustrated in Chapter 2. Previous research has already pursued this approach, and yielded, for example, the Columbia card task (Figner et al., 2009) and the Soochow gambling task (Lin et al., 2009). Such advances towards a better experimental task to investigate risky decision making—be it a modification of the IGT or a complete new paradigm—are essential, especially given the goal to use such a task for individual diagnoses (Bechara, 2007).

Second, despite the meticulous model development and comparison efforts of the last 15 years, this dissertation underlines that the search for an appropriate IGT model is still subject to ongoing research. Past research focused on developing new models that are a combined version of existing RL models (e.g., the PVL-Delta and EV-PU models) or a slightly adapted variant of existing RL models (e.g., Dai et al., 2015; Pang, Blanco, Maddox, & Worthy, 2016). In addition, past research focused on models that include additional processes, such as perseveration (Worthy, Pang, & Byrne, 2013). Alternatively, to advance the search for an appropriate IGT model, it might be promising to explore models that differ from the traditional RL framework more strongly, such as the instance-based learning model (Gonzalez & Dutt, 2011). Yet, another approach is to enrich current modeling approaches by combining them with models that focus on other processes, such as RT or wagering (Konstantinidis & Shanks, 2014; Persaud, McLeod, & Cowey, 2007).

Third, the state-of-the-art framework for parameter inference can be improved even further. We generally advocated the Bayesian hierarchical framework; however, the idiosyncratic choice behavior of healthy participants reported in Chapter 2 suggests that it may be a mistake to assume a single group of healthy participants. Instead, it might be more realistic to use a Bayesian hierarchical framework that contains a mixture component (e.g., Bartlema, Lee, Wetzels, & Vanpaemel, 2014; Huizenga et al., 2007, and Chapter 9). In addition, parameter inference can be improved by developing more informative prior distributions based on published empirical findings (e.g., Gershman, 2016).

Moreover, it is important to note that the results of this dissertation, but also many other model comparison studies over the last 15 years, are mainly based on data of healthy participants. It is therefore informative and importing to investigate whether our results generalize to clinical populations. The RL models are often applied to IGT data of clinical groups to better understand their decision-making deficits, and therefore it is crucial that conclusions from model parameters are trustworthy. Hence we recommend that absolute model performance is assessed also for clinical datasets. In addition, we recommend that sophisticated model comparison tools are used to identify a model that can jointly capture psychological processes driving IGT performance of clinical and healthy participants. Moreover, it is desirable to test the assumptions underlying IGT performance of patients with lesions to the ventromedial prefrontal cortex (vmPFC) analogous to the procedure reported in Chapter 2 focusing on healthy participants. Such research could reveal whether Bechara et al. (1994)’s key assumptions at least hold for the clinical population of vmPFC patients. Yet another insightful project would be to try and replicate Yechiam et al. (2005)’s mapping of the model parameters of 10 different clinical groups using state-of-the-art methods.
11. Summary and Future Directions

Last but not least, it is important to be aware that modeling IGT data is of interest to a diverse group of researchers and practitioners (e.g., clinical psychologists, statisticians, mathematical psychologists, and computer scientists). To advance our understanding of risky decision making, it is therefore important to further collaborations between these different groups (Ahn, Dai, Vassileva, Busemeyer, & Stout, 2016). In addition, it is important to make model-fitting routines easier available and applicable (e.g., by adding them to JASP; JASP Team, 2015; or by providing an R package; Ahn, Haines, & Zhang, 2016). Clear guidelines and support are needed to facilitate the choice of an appropriate IGT model out of the large available pool of different models, to explain how the to fit the models, how to assess the account of the models for the data, and how to interpret the results. Exemplary steps towards this ambition have already been initiated by Ahn, Haines, and Zhang (2016).

11.3 Concluding Remarks

The development of the Iowa gambling task (Bechara et al., 1994) and the application of reinforcement-learning models greatly advanced research efforts about risky decision making. On the one hand, these research effort led to many achievements as pointed out by Ahn and Busemeyer (2016): “computational modeling has greatly contributed to understanding cognitive processes underlying our decision-making” (p. 1; see also the special issue “Iowa Gambling Task (IGT): Twenty Years After” in Frontiers in Psychology). Another summary of the achievements is given by Ahn, Dai, et al. (2016, p. 62):

The past 10 years have seen the development of cognitive models for the IGT and SGT, with adequate model fits, and parameters that appear to have good utility for distinguishing between various clinical samples, and that relate to significant individual characteristics such as personality measures and severity of clinical symptoms. These have deepened the understanding of the variety and nature of differences between various substance abuse and other clinical groups, opening a potential window into the way basic psychological processes such as learning from experience or feedback, and sensitivity to reward and punishment, may be affected by substance abuse or may create vulnerability factors for developing substance use disorders. Furthermore, these models have provided a possible way in which individual characteristics can be assessed and targeted in individually tailored treatments.

On the other hand, there is still much progress ahead, as pointed out in this dissertation and also by several other researchers: Ahn and Busemeyer (2016), for example, admit that “Despite the growing enthusiasm, no computational assays or methods have influenced clinical practice yet” (p. 1). In addition, Gershman (2016) point to problems of RL models by stating that “fitting the parameters of these models can be challenging: the parameters are not identifiable, estimates are unreliable, and the fitted models may not have good predictive validity” (p. 1).

Discussing a large variety of models and methods to compare the models, this dissertation illustrated that research efforts about risky decision making greatly advanced during the last years. On the other hand, this dissertation also illustrated the major challenges by pointing to problems with respect to behavioral analyses and cognitive modeling. In particular, I tried to point out how we can achieve more meaningful conclusions about the psychological processes driving risky decision making. First of all, I argued that it is crucial to change the way how IGT data are typically analyzed. I proposed using Bayesian repeated measures ANOVA, analyses of deck selection profiles, and consideration of group-choices as a function of blocks for each deck
separately to avoid hiding a frequency-of-losses effect and to better reveal the individual deck preferences. Second, I argued that it is crucial to assess absolute model performance carefully, using post hoc fit and simulation performance in order to avoid premature conclusions from the model parameters. Third, I illustrated a number of different methods that can be used to compare RL models or the model parameters of different groups of decision makers. Pursuing these suggestions will hopefully advance us towards the ambitious goal of using the IGT—or an alternative risky decision-making paradigm—and cognitive models to reliably measure risky decision making and to understand the underlying psychological processes.
Appendices
Appendix A

Performance and Awareness in the Iowa Gambling Task

This chapter has been published as:
Performance and awareness in the Iowa gambling task.
Behavioral and Brain Sciences, 37, 41–42.

Abstract

Newell & Shanks (Newell & Shanks, 2014; N&S) conclude that healthy participants learn to differentiate between the good and bad decks of the Iowa Gambling Task, and that healthy participants even have conscious knowledge about the task’s payoff structure. Improved methods of analysis and new behavioral findings suggest that this conclusion is premature.

Newell & Shanks (N&S) convincingly argue that past research has severely overstated the importance of conscious processes in decision making. We agree with N&S on many counts, but here we focus on what is perhaps our sole source of dissention. N&S conclude that healthy participants who perform the Iowa Gambling Task (IGT) learn to differentiate between the good and bad decks, and that this behavioral differentiation is even reflected in conscious knowledge about the payoff structure. We believe this conclusion may be premature: Several pitfalls in IGT data analysis methods frustrate a fair interpretation of IGT data, and several behavioral findings go against the authors’ conclusion.

The first pitfall is that the traditional way of analyzing IGT data is incomplete and potentially misleading because it collapses choice behavior over the two good decks and over the two bad decks. This procedure hides the impact of the frequency of losses (bad deck B and good deck D yield rare losses, whereas bad deck A and good deck C yield frequent losses) and potentially obscures diagnostic information. For example, consider the data of Fridberg et al. (2010) healthy participants. Fridberg et al. plot the mean proportion of choices from the good and bad decks as

\footnote{1The final publication is available at \texttt{doi:10.1017/S0140525X13000861}.}
a function of trial number, replotted here in the left panel of Figure 1. This panel suggests that participants learn to prefer the good decks. However, Fridberg et al. also plot the mean proportion of choices from each deck separately, replotted here in the right panel of Figure A.1. This panel shows that, across all trials, participants prefer the decks with infrequent losses (B & D).

A similar problem is evident in work that assesses conscious knowledge about the IGT either with subjective experience ratings \((C+D) - (A+B)\) (Bowman et al., 2005; Cella, Dymond, Cooper, Turnbull, et al., 2007), or by determining whether participants have conscious knowledge that would encourage them to choose one of the two best decks (Maia & McClelland, 2004). However, participants who consider “one of the best decks as the best deck” do not necessarily understand that there are two best decks and that both bad decks should be avoided. To investigate whether participants understand that there are two good decks, participants should identify the best and second-best deck on each trial.

The final pitfall concerns the way in which IGT studies typically assess the learning process, namely by applying an analysis of variance to assess whether participants’ preference for the good decks (i.e., \((C+D) - (A+B)\)) increases over blocks of trials (main effect of block). A significant effect of block is typically taken as evidence that participants learned to discriminate between the good and bad decks. However, when the main effect of block is significant, this does not imply it is also substantial. For example, consider the data of Bowman et al. (2005), who tested three groups of healthy participants that differed in whether they obtained a manual or computerized IGT combined with or without a 6-second delay. The only significant effect was a main effect of block. However, even in the last block (i.e., the final 20 trials), the three groups showed at most a weak preference for the good decks, as \((C+D) - (A+B)\) ranged from about 3 to about 6.5. A value of 3 corresponds to an average of 11.5 out of 20 choices from the good decks, and a value of 6.5 corresponds to an average of 13.25 out of 20 choices from the good decks. Similar unconvincing results were evident from subjective ratings of how positive each deck was experienced. These findings suggest that neither participants’ behavioral preference for the good decks nor their conscious preference for the good decks is substantial. Cella et al. (2007) reported similar findings.

Next to the above mentioned pitfalls, several behavioral findings contradict the conclusion
from N&S. First, a detailed reanalysis of eight data sets showed that healthy participants learn to prefer the good decks in only one data set (see Steingroever, Wetzels, Horstmann, et al., 2013, and references therein). In the remaining seven data sets, participants either only learn to avoid bad deck A (frequent losses) or prefer the decks with infrequent losses (decks B & D). Such a preference for the decks with infrequent losses—the frequency-of-losses effect—has been reported by many studies. The empirical evidence for the frequency-of-losses effect contradicts the assumption that healthy participants learn to prefer the good decks.

Second, Steingroever, Wetzels, Horstmann, et al. (2013) showed that participants have a tendency to switch frequently throughout the entire task. This is counterintuitive because one expects a strong decrease in the mean number of switches once participants learned to prefer the good decks. The frequent switches suggest that participants do not learn to systematically differentiate between the good and bad decks, a suggestion that is illustrated by deck selection profiles of 394 participants (Steingroever, Wetzels, Horstmann, et al., 2013; see https://dl.dropbox.com/u/12798592/DeckSelectionProfiles.zip for the deck selection profiles); each participant has a highly idiosyncratic choice pattern, and for most participants it is impossible to identify a point where they realized that the good decks should be preferred.

In sum, detailed analyses of IGT data have shown that even healthy participants are unable to discriminate the good decks from the bad decks, a finding that suggests a lack of both conscious and unconscious knowledge in this task.
Appendix B

Data from 617 Healthy Participants Performing the Iowa Gambling Task: A “Many Labs” Collaboration

This chapter has been published as:

Data from 617 healthy participants performing the Iowa gambling task: A “many labs” collaboration.

Journal of Open Psychology Data, 3:e5

Abstract

This data pool (N = 617) comes from 10 studies assessing performance of healthy participants (i.e., no known neurological impairments) on the Iowa gambling task (IGT)—a task measuring decision making under uncertainty in an experimental context. Participants completed a computerized version of the IGT consisting of 95 – 150 trials. The data consist of the choices of each participant on each trial, and the resulting rewards and losses. The data are stored as .rdata, .csv, and .txt files, and can be reused to (1) analyze IGT performance of healthy participants; (2) create a “super control group”; or (3) facilitate model-comparison efforts.

\footnote{The final publication is available at \url{http://openpsychologydata.metajnl.com/article/10.5334/jopd.ak/}}
B. Data from 617 Healthy Participants Performing the IGT

B.1 Overview

Context

Collection Date(s)

2000 - 2013

Background

This data pool comes from eight independent published studies (Fridberg et al., 2010; Horstmann; Kjome et al., 2010; Maia & McClelland, 2004; Premkumar et al., 2008; Wetzels, Vandekerckhove, et al., 2010; Wood et al., 2005; Worthy, Pang, & Byrne, 2013), one submitted study (Steingroever et al., submitted), and one unpublished study (Steingroever, Wetzels, & Wagenmakers, 2011). These studies report the performance of a total of 617 healthy participants on the Iowa gambling task (IGT; Bechara et al., 1994). The IGT is arguably the most popular neuropsychological paradigm to measure decision-making deficits in an experimental context. Part of the data was already reanalyzed elsewhere (i.e., Steingroever et al., submitted; Steingroever, Wetzels, Horstmann, et al., 2013; Steingroever, Wetzels, & Wagenmakers, 2013a, 2013b; Steingroever et al., 2014, 2016) in order to assess basic assumptions underlying the performance of healthy participants on the IGT, and to compare reinforcement-learning models that try to disentangle psychological processes underlying performance on the IGT.

B.2 Methods

Sample

Table B.1 describes the data pool. All included studies used (a variant of) the traditional IGT payoff scheme (Bechara et al., 1994) or the payoff scheme introduced by Bechara and Damasio (2002). A detailed description of the payoff schemes can be found in the Supporting Text 1.

In the traditional payoff scheme, the net outcome of 10 cards from the bad decks (i.e., decks A and B) is -250, and +250 in the case of the good decks (i.e., decks C and D). In addition, there are two decks with frequent losses (decks A and C), and two decks with infrequent losses (decks B and D). In the traditional payoff scheme, there is a variable loss in deck C (i.e., either -25, -50, or -75; classified here as payoff scheme 1). However, some of the included studies used a variant of this payoff scheme in which the loss in deck C was held constant (i.e., -50; classified here as payoff scheme 2). A second difference between payoff scheme 1 and 2 is that payoff scheme 1 uses a fixed sequence of rewards and losses, whereas payoff scheme 2 uses a randomly shuffled sequence.

The payoff scheme introduced by Bechara and Damasio (2002; classified here as payoff scheme 3) also consists of two good decks (decks C and D), and two bad decks (decks A and B), and two decks with frequent losses (decks A and C) and two decks with infrequent losses (decks B and D). However, in contrast to payoff schemes 1 and 2, the schedules of rewards and losses in payoff scheme 3 are structured in such a way that the discrepancy between rewards and losses in the bad decks (decks A and B) changes such that the net outcome decreases by 150 every block of 10 cards (i.e., in the first block, the net outcome is -250, but in the sixth block, it is -1000). By contrast, the net outcome of the good decks (decks C and D) increases by 25 every block of 10 cards (i.e., in the first block, the net outcome is 250, but in the sixth block, it is 375). Thus, the good decks become gradually better, whereas the bad decks become gradually worse. In addition, in contrast
Table B.1: Overview of the studies included in the data pool. See text for a description of the different payoff schemes.

<table>
<thead>
<tr>
<th>Study</th>
<th>Number of participants</th>
<th>Number of trials</th>
<th>Payoff</th>
<th>Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fridberg et al. (2010)</td>
<td>15</td>
<td>95</td>
<td>1</td>
<td>M = 29.6 years (SD = 7.6)</td>
</tr>
<tr>
<td>Horstmann b</td>
<td>162</td>
<td>100</td>
<td>2</td>
<td>M = 25.6 years (SD = 4.9), 82 female</td>
</tr>
<tr>
<td>Kjome et al. (2010)</td>
<td>19</td>
<td>100</td>
<td>3</td>
<td>M = 33.9 years (SD = 11.2), 6 female</td>
</tr>
<tr>
<td>Maia and McClelland (2004)</td>
<td>40</td>
<td>100</td>
<td>1</td>
<td>Undergraduate students</td>
</tr>
<tr>
<td>Premkumar et al. (2008)</td>
<td>25</td>
<td>100</td>
<td>3</td>
<td>M = 35.4 years (SD = 11.9), 9 female</td>
</tr>
<tr>
<td>Steingroever et al. (submitted)</td>
<td>70</td>
<td>100</td>
<td>2</td>
<td>M = 24.9 years (SD = 5.8), 49 female</td>
</tr>
<tr>
<td>Steingroever et al. (2011)</td>
<td>57</td>
<td>150</td>
<td>2</td>
<td>M = 19.9 years (SD = 2.7), 42 female</td>
</tr>
<tr>
<td>Wetzels, Vandekerckhove et al. (2010)</td>
<td>41</td>
<td>150</td>
<td>2</td>
<td>Students</td>
</tr>
<tr>
<td>Wood et al. (2005)</td>
<td>153</td>
<td>100</td>
<td>3</td>
<td>M = 45.25 years (SD = 27.21) d</td>
</tr>
<tr>
<td>Worthy, Pang, and Byrne (2013)</td>
<td>35</td>
<td>100</td>
<td>1</td>
<td>Undergraduate students, 22 female</td>
</tr>
</tbody>
</table>

Information that was provided in the original articles. This information consists of the mean age and the standard deviation in brackets, or alternatively the occupation of the participants. In addition, the number of female participants is provided for most datasets.

Data collected by Annette Horstmann. These data were first published in Steingroever, Wetzels, Horstmann, et al. (2013). A subset of this dataset is published in Horstmann, Villringer, and Neumann (2012).

Data of the standard condition. Data of three other conditions can be downloaded here: [http://www.ruudwetzels.com/data/EV$_$data.zip](http://www.ruudwetzels.com/data/EV$_$data.zip).

The first 90 participants of this dataset are between 18-40 years old (M = 23.04, SD = 5.88), and participants 91-153 are between 61 and 88 years old (M = 76.98, SD = 5.20).

to payoff schemes 1 and 2, the wins differ within each deck in payoff scheme 3. Just as payoff scheme 1, payoff scheme 3 uses a fixed sequence of wins and losses.

B.3 Dataset Description

Object name

IGTdataSteingroever2014.zip. This zip archive contains the following files:

- IGTdata.rdata
- choice_95.csv, choice_100.csv, choice_150.csv, wi_95.csv, wi_100.csv, wi_150.csv, lo_95.csv, lo_100.csv, lo_150.csv, index_95.csv, index_100.csv, index_150.csv.
- choice_95.txt, choice_100.txt, choice_150.txt, wi_95.txt, wi_100.txt, wi_150.txt, lo_95.txt, lo_100.txt, lo_150.txt, index_95.txt, index_100.txt, index_150.txt.
Data Type
Processed data

Format names and versions
The data are provided in three different formats: .rdata (R), .csv (Excel), and .txt. The .rdata file is called “IGTdata.rdata” and it contains the following 12 matrices:

- **choice_95**, **choice_100**, and **choice_150**: These matrices contain the choices of all studies that used a 95-trial, 100-trial, and 150-trial IGT, respectively. The dimension of each matrix corresponds to the number of subjects x number of trials. For example, choice_95 (Figure B.1) is a 15 x 95 matrix, and the entry of the third row and fifth column corresponds to the choice that the third participant made on the fifth trial. The entries of the three choice matrices are either 1, 2, 3, or 4, where 1, 2, 3, and 4 stand for deck A, B, C, and D, respectively.

- **wi_95**, **wi_100**, and **wi_150**: These matrices contain the rewards of all studies that used a 95-trial, 100-trial, and 150-trial IGT, respectively. The dimension of each matrix corresponds to the number of subjects x number of trials. For example, wi_100 is a 504 x 100 matrix, and the entry of the third row and fifth column corresponds to the reward that the third participant received on the fifth trial. The entries of the three reward matrices vary between 40 and 170.

- **lo_95**, **lo_100**, and **lo_150**: These matrices contain the losses of all studies that used a 95-trial, 100-trial, and 150-trial IGT, respectively. The dimension of each matrix corresponds to the number of subjects x number of trials. For example, lo_150 is a 98 x 150 matrix, and the entry of the third row and fifth column corresponds to the loss that the third participant received on the fifth trial. The entries of the three loss matrices vary between $-2500$ and 0. Thus, the losses are saved as negative numbers.

- **index_95**, **index_100**, and **index_150**: These matrices contain the name of the first author of the study that reports the data of the corresponding participant. For example, the third entry of index_95 (Figure B.2) can be used to identify who collected the choices saved in the third row of choice_95, wi_95, and lo_95.

These 12 matrices are saved altogether in the “IGTdata.rdata” file. In addition, we saved the 12 matrices as separate .csv and .txt files. For example, the matrix choice_95 (Figure B.1) is contained in the “IGTdata.rdata” file, but also in “choice_95.csv” and “choice_95.txt”. The file names of the .csv and .txt files indicate which matrix they contain.

Data Collectors
Language
N/A

License
Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)
B.3. Dataset Description

Figure B.1: Screenshot of a subset of the choice_95 matrix. Each row contains the sequence of choices from a specific participant. For example, the entry of the third row and fifth column corresponds to the choice that the third participant made on the fifth trial (i.e., “2” – deck B). To determine who collected the data of this particular participant, one needs to refer to the third row of index_95 (cf. Figure B.2).

<table>
<thead>
<tr>
<th>Subj</th>
<th>Choice_1</th>
<th>Choice_2</th>
<th>Choice_3</th>
<th>Choice_4</th>
<th>Choice_5</th>
<th>Choice_6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subj_1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Subj_2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Subj_3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Subj_4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Subj_5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Subj_6</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Subj_7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Subj_8</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Subj_9</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Subj_10</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure B.2: Screenshot of a subset of the index_95 matrix. Each row can be used to identify who collected the data of a specific participant. The screenshot shows that the data of subjects 1 – 10 who completed a 95-trial IGT were collected by Fridberg et al. (2010). The data of these subjects are saved in the corresponding rows of the choice_95, reward_95, and loss_95 matrices.
B. Data from 617 Healthy Participants Performing the IGT

Embargo
N/A

Repository location
https://osf.io/8t7rm

Publication date
05/11/2014

B.4 Reuse Potential

This data pool has several reuse potentials: First, it could be used to more thoroughly investigate healthy participants’ performance on the IGT. Second, it could be reused as a “super control group”. This means that performance of an experimental group can be assessed relative to the performance of healthy participants included in this data pool. Third, the data pool could be reused to compare computational models for the IGT. However, it should be noted that the 10 datasets were collected in different environments, and that the performance of the participants on the IGT may possibly be affected by factors that varied across the included studies (e.g., the use and type of incentives, questions about the IGT during the performance to assess participants’ awareness, randomly shuffled payoff or fixed payoff sequence, the type of task instruction).

Acknowledgements

We would like to thank Kala Battistelli, Andreas Below, Katie Chamberlain, Courtni East, Noénie J. Eichhorn, Lindsey Ferris, Monica Gamboa, Kaitlynn Goldman, Karolin Gohlke, Karla Gomez, Alexis Gregg, Jordan Hall, Christy Ho, Jonas Klinkenberg, Lauren Laserna, Katja Macher, Samantha Mallec, Megan McDermott, Ramona Menger, Gerard Moeller, Michael Pang, Anthony Schmidt, Candice Tharp, Lucas Weatherall, Michael Wesley, and Christopher Whitlow for their help in collecting the data.

Funding Statement

This work was supported by BMBF [IFB Adiposity Diseases (FKZ:01EO1001, ADI-K50F, ADI-K753, ADI-K754) to Annette Horstmann], the DFG (Mind and Brain; CRC 1052 “Obesity Mechanisms” A5 to Annette Horstmann), the Einstein-Stiftung (Mind and Brain Institute), a fellowship from the Foundation for Science and Technology of Portugal (to Tiago V. Maia) and by National Institute of Mental Health Grant P50-MH64445 (to James L. McClelland), funds from the Wellcome Trust, UK (067427/z/02/z), the National Institute on Drug Abuse (NIDA; PI F.G. Moeller, contract/grant number: K02DA00403, P50DA009262), and CCTS/CRU grant number UL1 RR024148.
B.5 Supporting Text 1

This document provides details on the three different payoff schemes employed by the included studies. In addition, it explains how the data can be read in R, and presents an example of an analysis.

Payoff Schemes

Table B.2: Rewards and losses from 40 choices of each deck as used in the traditional payoff scheme with variable loss in deck C (see Bechara et al., 1994; classified here as payoff scheme 1). Within each deck, the presented payoff sequence is repeated after participants have made 40 choices from the corresponding deck.

<table>
<thead>
<tr>
<th>Trial</th>
<th>win A</th>
<th>win B</th>
<th>win C</th>
<th>win D</th>
<th>loss A</th>
<th>loss B</th>
<th>loss C</th>
<th>loss D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-150</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-300</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-200</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-250</td>
<td>-1250</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-350</td>
<td>0</td>
<td>-50</td>
<td>-250</td>
</tr>
<tr>
<td>11</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-350</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-250</td>
<td>-1250</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-200</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-300</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-150</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-50</td>
<td>-250</td>
</tr>
<tr>
<td>21</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>-1250</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-300</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-350</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-200</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-250</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-150</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-75</td>
<td>-250</td>
</tr>
<tr>
<td>30</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
</tbody>
</table>
Table B.2 – Continued from previous page

<table>
<thead>
<tr>
<th>Trial</th>
<th>win A</th>
<th>win B</th>
<th>win C</th>
<th>win D</th>
<th>loss A</th>
<th>loss B</th>
<th>loss C</th>
<th>loss D</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-350</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-200</td>
<td>-1250</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-250</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-25</td>
<td>-250</td>
</tr>
<tr>
<td>36</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-150</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>38</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>-1250</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
</tbody>
</table>

Table B.3: A possible sequence of rewards and losses from 10 choices of each deck based on the traditional payoff scheme with constant loss in deck C (see Bechara et al., 1994; classified here as payoff scheme 2). A payoff sequence with the presented characteristics is randomly generated for each block of 10 trials.

<table>
<thead>
<tr>
<th>Trial</th>
<th>win A</th>
<th>win B</th>
<th>win C</th>
<th>win D</th>
<th>loss A</th>
<th>loss B</th>
<th>loss C</th>
<th>loss D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-300</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-150</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-350</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>-1250</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-250</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>-50</td>
<td>-250</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>-200</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
</tbody>
</table>

Tables B.2–B.4 illustrate the payoff schemes classified as payoff scheme 1, 2, and 3, respectively. The difference between payoff scheme 1 and 2 is that the first one uses variable losses in deck C and has a fixed sequence of wins and losses (Table B.2), whereas the second one uses a constant loss in deck C and, within each deck, a random repetition of the payoffs presented in Table B.3. The third payoff scheme changes rewards and losses in such a way that the net outcome of the bad decks becomes increasingly negative every 10 trials, whereas the net outcome of the good decks becomes increasingly positive. Thus, the difference in the net outcomes of the good and bad decks increases every 10 trials. In addition, in contrast to payoff schemes 1 and 2, the wins differ within each deck in payoff scheme 3. Just as payoff scheme 1, payoff scheme 3 uses a fixed sequence of wins and losses.

Table B.4: Rewards and losses from 60 choices of each deck as used in the payoff scheme introduced by Bechara and Damasio, 2002; classified here as payoff scheme 3).

<table>
<thead>
<tr>
<th>Trial</th>
<th>win A</th>
<th>win B</th>
<th>win C</th>
<th>win D</th>
<th>loss A</th>
<th>loss B</th>
<th>loss C</th>
<th>loss D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Continued on next page
Table B.4 – Continued from previous page

<table>
<thead>
<tr>
<th>Trial</th>
<th>win A</th>
<th>win B</th>
<th>win C</th>
<th>win D</th>
<th>loss A</th>
<th>loss B</th>
<th>loss C</th>
<th>loss D</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>120</td>
<td>80</td>
<td>60</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>80</td>
<td>110</td>
<td>40</td>
<td>45</td>
<td>-150</td>
<td>-50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>90</td>
<td>120</td>
<td>55</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>110</td>
<td>90</td>
<td>55</td>
<td>55</td>
<td>-300</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>100</td>
<td>45</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
<td>90</td>
<td>50</td>
<td>40</td>
<td>-200</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>120</td>
<td>120</td>
<td>45</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>110</td>
<td>110</td>
<td>60</td>
<td>50</td>
<td>-250</td>
<td>-1250</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>90</td>
<td>80</td>
<td>40</td>
<td>60</td>
<td>-350</td>
<td>0</td>
<td>-50</td>
<td>-250</td>
</tr>
<tr>
<td>11</td>
<td>110</td>
<td>110</td>
<td>55</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>130</td>
<td>100</td>
<td>55</td>
<td>40</td>
<td>-350</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>90</td>
<td>90</td>
<td>65</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>100</td>
<td>130</td>
<td>45</td>
<td>40</td>
<td>-250</td>
<td>-1500</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>120</td>
<td>120</td>
<td>70</td>
<td>45</td>
<td>-200</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>110</td>
<td>130</td>
<td>40</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>90</td>
<td>110</td>
<td>50</td>
<td>65</td>
<td>-300</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>130</td>
<td>90</td>
<td>60</td>
<td>70</td>
<td>-150</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>120</td>
<td>100</td>
<td>70</td>
<td>50</td>
<td>-250</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>120</td>
<td>40</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>-50</td>
<td>-275</td>
</tr>
<tr>
<td>21</td>
<td>120</td>
<td>120</td>
<td>60</td>
<td>60</td>
<td>-250</td>
<td>-1750</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>140</td>
<td>110</td>
<td>65</td>
<td>55</td>
<td>-300</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>110</td>
<td>140</td>
<td>55</td>
<td>65</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>110</td>
<td>130</td>
<td>80</td>
<td>80</td>
<td>-350</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>100</td>
<td>100</td>
<td>40</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>120</td>
<td>110</td>
<td>60</td>
<td>80</td>
<td>-200</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>130</td>
<td>120</td>
<td>55</td>
<td>40</td>
<td>-250</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>110</td>
<td>120</td>
<td>65</td>
<td>65</td>
<td>-150</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>140</td>
<td>140</td>
<td>40</td>
<td>55</td>
<td>-250</td>
<td>0</td>
<td>-75</td>
<td>-300</td>
</tr>
<tr>
<td>30</td>
<td>120</td>
<td>110</td>
<td>80</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>130</td>
<td>130</td>
<td>65</td>
<td>65</td>
<td>-350</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>120</td>
<td>140</td>
<td>75</td>
<td>75</td>
<td>-200</td>
<td>-2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>140</td>
<td>120</td>
<td>55</td>
<td>60</td>
<td>-250</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>130</td>
<td>110</td>
<td>60</td>
<td>65</td>
<td>-250</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>110</td>
<td>130</td>
<td>70</td>
<td>75</td>
<td>-150</td>
<td>0</td>
<td>-25</td>
<td>-325</td>
</tr>
<tr>
<td>36</td>
<td>150</td>
<td>150</td>
<td>65</td>
<td>85</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>140</td>
<td>110</td>
<td>55</td>
<td>45</td>
<td>-150</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>38</td>
<td>120</td>
<td>150</td>
<td>75</td>
<td>55</td>
<td>-300</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>150</td>
<td>120</td>
<td>45</td>
<td>70</td>
<td>-350</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>110</td>
<td>140</td>
<td>85</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>41</td>
<td>140</td>
<td>140</td>
<td>70</td>
<td>70</td>
<td>-350</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>42</td>
<td>130</td>
<td>150</td>
<td>80</td>
<td>80</td>
<td>-200</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Continued on next page
Table B.4 – Continued from previous page

<table>
<thead>
<tr>
<th>Trial</th>
<th>win A</th>
<th>win B</th>
<th>win C</th>
<th>win D</th>
<th>loss A</th>
<th>loss B</th>
<th>loss C</th>
<th>loss D</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>150</td>
<td>130</td>
<td>60</td>
<td>65</td>
<td>-250</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>140</td>
<td>120</td>
<td>65</td>
<td>70</td>
<td>-250</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>120</td>
<td>140</td>
<td>75</td>
<td>80</td>
<td>-150</td>
<td>0</td>
<td>-25</td>
<td>-350</td>
</tr>
<tr>
<td>46</td>
<td>160</td>
<td>160</td>
<td>70</td>
<td>90</td>
<td>0</td>
<td>-2250</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>47</td>
<td>150</td>
<td>120</td>
<td>60</td>
<td>50</td>
<td>-150</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>48</td>
<td>130</td>
<td>160</td>
<td>80</td>
<td>60</td>
<td>-300</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>49</td>
<td>160</td>
<td>130</td>
<td>50</td>
<td>75</td>
<td>-350</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>120</td>
<td>150</td>
<td>90</td>
<td>60</td>
<td>-250</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>51</td>
<td>150</td>
<td>150</td>
<td>75</td>
<td>75</td>
<td>-350</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>52</td>
<td>140</td>
<td>160</td>
<td>85</td>
<td>85</td>
<td>-200</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>53</td>
<td>160</td>
<td>140</td>
<td>65</td>
<td>70</td>
<td>-250</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>54</td>
<td>150</td>
<td>130</td>
<td>70</td>
<td>75</td>
<td>-250</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>55</td>
<td>130</td>
<td>150</td>
<td>80</td>
<td>85</td>
<td>-150</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>56</td>
<td>170</td>
<td>170</td>
<td>75</td>
<td>95</td>
<td>-250</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>57</td>
<td>160</td>
<td>130</td>
<td>65</td>
<td>55</td>
<td>-150</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
<tr>
<td>58</td>
<td>140</td>
<td>170</td>
<td>85</td>
<td>65</td>
<td>-300</td>
<td>-2500</td>
<td>-25</td>
<td>-375</td>
</tr>
<tr>
<td>59</td>
<td>170</td>
<td>140</td>
<td>55</td>
<td>80</td>
<td>-350</td>
<td>0</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>60</td>
<td>130</td>
<td>160</td>
<td>95</td>
<td>65</td>
<td>-250</td>
<td>0</td>
<td>-75</td>
<td>0</td>
</tr>
</tbody>
</table>
Appendix C

Appendix to Chapter 4: “Validating the PVL-Delta Model for the Iowa Gambling Task”

In this appendix, we present how we visually assessed convergence, additional absolute model performance checks and the results of two parameter-recovery studies that confirm that we correctly implemented the PVL-Delta model. For the parameter-recovery studies, we used two synthetic data sets that were generated with the PVL-Delta model. The data-generating parameters correspond to the medians of the individual-level joint posteriors that were obtained by fitting the PVL-Delta model to two real data sets.

Figure C.1 shows the HMC chains of one individual-level parameter. From the figure it is evident that the chains have converged successfully from their starting values to their stationary distribution, looking like “hairy caterpillars” that are randomly intermixed. We inspected this type of plot for every parameter to visually assess convergence in addition to the formal diagnostic measure of convergence $\hat{R}$.

Figure C.2 presents the fit performance and simulation performance of the PVL-Delta model that was obtained with random draws from the joint posterior distributions over the group-level parameters (hereafter group-level joint posteriors). Note that Figure 4.9 presents the fit performance and simulation performance based on the individual-level joint posteriors. A comparison of both figures reveals that the fit performance based on the group-level joint posteriors (Figure C.2) closely matches the fit performance based on the individual-level joint posteriors (Figure 4.9). However, there are a few discrepancies in the case of the simulation performance: From Figures 4.9 and C.2 it is evident that the simulation performance based on the group-level joint posteriors is more extreme, that is, the most preferred deck is preferred even stronger, whereas the least preferred deck is avoided even stronger. However, it is evident that in general Figure C.2 mirrors the conclusion drawn from Figure 4.9.

Figure C.3 presents the results of the first recovery study. This data set contains 18 synthetic participants. The figure contains four panels; each panel illustrates the recovery of one of the four model parameters. In each panel, the mode of the group-level posterior is represented by the dotted line, whereas the solid line represents the true group-level parameter. In addition, the panels can also be used to assess the individual-level recovery: The unfilled dots represent the modes of the
C. Appendix to Chapter 4: “Validating the PVL-Delta Model for the Iowa Gambling Task”

Figure C.1: HCM chains of the individual-level consistency parameter $c$ of the third participant in the consistency condition. In addition to the formal diagnostic measure of convergence $\hat{R}$, we inspected this type of plot for every parameter to visually assess convergence.

![HCM chains of the individual-level consistency parameter c](image)

individual-level posteriors, whereas the filled dots represent the true individual-level parameters.

Note that the individual-level posterior distributions are not sorted by the subject ID; in order to visualize the degree of individual differences in each model parameter, we sorted the individual-level posterior distributions by the true individual-level parameters.

From Figure C.3 it is evident that the group-level updating parameter is slightly underestimated, but the remaining group-level parameters are recovered very accurately. However, the recovery of the individual-level parameters is less accurate. Especially in the case of the shape parameter, most of the individual-level modes differ from the true individual-level parameters by regressing to the mode of the group-level parameter (i.e., shrinkage); small deviations are noticeable in the case of the individual-level loss aversion parameters and the individual-level updating parameter. Yet, in the case of the consistency parameter, most individual-level parameters are recovered very accurately.

Figure C.4 presents the results of the second recovery study. This data set contains 30 synthetic participants. It is evident that all group-level parameters are recovered very accurately. However, the recovery of the individual-level parameters is less accurate. Especially in the case of the individual-level shape parameters and the individual-level loss aversion parameters, it is evident that the individual-level modes differ from the true individual-level parameters. Yet, the recovery of the individual-level updating parameters and the individual-level consistency parameters is adequate.
Figure C.2: Observed choice behavior and assessment of absolute model performance. The first column shows the mean proportion of choices from each deck within 15 blocks as observed in the four experimental conditions reported by [Wetzels, Vandekerckhove, et al. (2010)]. Each block contains 10 trials. The second and third column show the fit performance and simulation performance, respectively, for each of the four conditions. Fit performance and simulation performance are based on random draws from the Fit performance and simulation performance are based -level joint posteriors.
Figure C.3: Recovery of individual and group-level parameters of the PVL-Delta model. Data of 18 participants completing a 100-trial IGT. The dotted lines and the unfilled dots represent the modes of the group-level posteriors and of the individual-level posteriors, respectively. The black lines and black dots represent the true group-level and true individual-level parameters, respectively.
Figure C.4: Recovery of individual and group-level parameters of the PVL-Delta model. Data of 30 participants completing a 100-trial IGT. The dotted lines and the unfilled dots represent the modes of the group-level posteriors and of the individual-level posteriors, respectively. The black lines and black dots represent the true group-level and true individual-level parameters, respectively.
Appendix D

Appendix to Chapter 5: “Absolute Performance of Reinforcement-Learning Models for the Iowa Gambling Task”

D.1 Recipe for Obtaining Choice Probabilities According to the Post Hoc Absolute Fit Method

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

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

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

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

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

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

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

8. Repeat steps 1 – 7 for each subject 100 times to account for the posterior uncertainty. This step is omitted if maximum likelihood estimates were used.
D.2 Recipe for Obtaining Choice Probabilities According to the Simulation Method

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

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

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

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

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

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

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

8. Repeat steps 1 – 7 for each subject 100 times to account for the posterior uncertainty. This step is omitted if maximum likelihood estimates were used.
In this appendix, we present a recipe on how to obtain Bayes factors with importance sampling, and 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, we present the results of a robustness analysis showing that our conclusions are unaffected by the choice of the priors on the model parameters. Finally, we present a model comparison study using BIC for the same models and data pool as used in the article.

### E.1 Recipe for Importance Sampling

In this section, we present a recipe that describes how we obtained Bayes factors with importance sampling. We use $M$ to refer to specific model that can either be the EV, PVL, PVL-Delta, or VPP model.

1. Fit model $M$ to the data of participant $s = 1$.
2. Find the beta distributions (i.e., Beta($\alpha, \beta$)) with the best fit to the posterior distributions of $\theta^1$. Save the corresponding $\alpha$ and $\beta$ parameters.
3. Draw a set of parameters from the Beta importance densities, and compute the associated likelihood. Save the likelihood.
4. Repeat the previous step $D - 1$ times (with $D$ the number of draws).
5. Compute the marginal likelihood using Equation \[E.1\]

\[
m(y \mid M_{(i)}) \approx \frac{1}{N} \sum_{i=1}^{N} \frac{p(y \mid \theta_i, M_{(i)})p(\theta_i \mid M_{(i)})}{g(\theta_i \mid M_{(i)})}, \quad \theta_i \sim g(\theta \mid M_{(i)}) \quad (E.1)
\]

\[\text{Note that} \ \theta \ \text{represents a subject and model-specific parameter vector (see Table 2 in the main article for each model's parameters). This means that we obtain one beta distribution for each of the parameters contained in} \ \theta.\]
E.2 Model-Recovery Studies

In this section, we present the results of the model-recovery study. The purpose of this study was to confirm that the Bayes factor tends to favor the data-generating model. This study is based on eight generated data sets: We generated 25 synthetic participants completing a 100-trial IGT using each of the four models. As data-generating parameters we used the median parameter values obtained from fitting the models to a subset of the data used in the article.

We fit each of the four models to the four data sets, and then applied importance sampling to derive Bayes factors for all possible model comparisons. Analogous to the analyses reported in the main text, we present histograms showing the distribution of the log Bayes factors. In addition, we calculated the median posterior model probability for each model, and the proportion of participants for whom each model has the highest posterior model probability. The latter two should be high whenever the data-generating model is the same as the model that was used to fit the data (see also Pitt & Myung, 2002).

Figure E.1 shows the distribution of the log Bayes factors of 25 synthetic participants completing a 100-trial IGT. It is evident that in the case of all models, the majority of the synthetic participants provides evidence for the data-generating model. This finding is corroborated by Table E.1. The median posterior model probability and the percentage of participants for whom each model has the highest posterior model probability are highest for the data-generating model. Thus, these results suggest that our implementation of importance sampling is correct and that the Bayes factor is a useful model comparison tool.

Table E.1: Median posterior model probabilities (MPMP) and percentage of participants for whom the corresponding model has the largest posterior model probability. The data were generated with either the EV, PVL, PVL-Delta, or VPP model, and describe the performance of 25 synthetic participants on a 100-trial IGT (i.e., first model-recovery study).

<table>
<thead>
<tr>
<th>Data-generating model</th>
<th>EV</th>
<th>PVL</th>
<th>PVL-Delta</th>
<th>VPP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MPMP</td>
<td>%</td>
<td>MPMP</td>
<td>MPMP</td>
</tr>
<tr>
<td>EV</td>
<td>.95</td>
<td>100</td>
<td>.00</td>
<td>.07</td>
</tr>
<tr>
<td>PVL</td>
<td>.00</td>
<td>0</td>
<td>.69</td>
<td>56</td>
</tr>
<tr>
<td>PVL-Delta</td>
<td>.03</td>
<td>0</td>
<td>.02</td>
<td>12</td>
</tr>
<tr>
<td>VPP</td>
<td>.02</td>
<td>0</td>
<td>.12</td>
<td>20</td>
</tr>
</tbody>
</table>

E.3 Savage-Dickey Density Ratio Tests

An alternative way to check our implementation of importance sampling is to investigate whether Bayes factors obtained with our implementation of importance sampling are in line with Bayes factors obtained with the Savage-Dickey density ratio test (Dickey & Lientz, 1970; Dickey, 1971). The Savage-Dickey density ratio offers a method to compute Bayes factors for nested models. In order to be able to compare Bayes factors obtained with these two different methods, we thus needed to create nested RL models. This was done by fixing an arbitrary parameter of each model. We decided to fix the $a$ parameter of each model to a predefined value $a_0$, and indicate nested
Figure E.1: Histograms of the log(BF) of 25 synthetic participants completing a 100-trial IGT. Data of the first to fourth row were generated with the EV, PVL, PVL-Delta, and VPP model, respectively. A positive log(BF$_{12}$) indicates that the data are more likely to occur under the first model (i.e., the data-generating 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 (i.e., the model that did not generate the data).
models by $M_i^*$. Thus, the idea is to compare each of the four RL models to its nested version using both importance sampling and Savage-Dickey.

The Savage-Dickey method is explained in detail in [Lee and Wagenmakers, 2013], [Vandekerckhove et al., 2015], and [Wagenmakers et al., 2010]; here, we only provide the main idea: To obtain a Bayes factor comparing a RL model $M_i$ (where $i \in \{EV, PVL, PVL$-Delta, VPP$\}$) to its nested version $M_i^*$, we need to divide the prior ordinate at a fixed value of parameter $a$ (i.e., $a_0$) by the posterior ordinate at that same fixed parameter value. The Bayes factor according to the Savage-Dickey method is then defined as:

$$BF_{M_i, M_i^*} = \frac{p(y \mid M_i)}{p(y \mid M_i^*)} = \frac{p(a = a_0 \mid M_i)}{p(a = a_0 \mid y, M_i)},$$  \hspace{0.5cm} (E.2)

where $y$ is the data, and $a = a_0$ indicates that the parameter $a$ is fixed to a predefined value $a_0$.

The Bayes factor that we wish to approximate with importance sampling is the ratio of the marginal likelihood of the complete RL model and its nested version, that is:

$$BF_{M_i, M_i^*} = \frac{m(y \mid M_i)}{m(y \mid M_i^*)}.$$  \hspace{0.5cm} (E.3)

We applied the Savage-Dickey density ratio test and importance sampling to the same synthetic data set as used in the last section (i.e., 25 synthetic participants completing a 100-trial IGT). In Figure E.2 we present the Savage-Dickey density ratio test for the four models and four synthetic subjects; the results for the remaining participants are similar. The header of each plot shows the Bayes factor obtained with importance sampling (i.e., BF_IS), and the Bayes factor obtained with the Savage-Dickey method (i.e., BF_SD). The dashed and solid lines represent the prior and posterior distribution, respectively. The black dots indicate the height of the prior and posterior distributions at $a = a_0$. From the figure it is evident that there is a close correspondence between Bayes factors obtained with the Savage-Dickey density ratio test and importance sampling suggesting that we correctly implemented importance sampling.
Figure E.2: Illustration of the Savage-Dickey density ratio test for all models. Data of the first to fourth row were generated and fit with the EV, PVL, PVL-Delta, and VPP model, respectively. The header of each plot shows the BF obtained with importance sampling (i.e., BF-IS), and the Bayes factor obtained with the Savage-Dickey method (i.e., BF_SD). The dashed and solid lines represent the prior and posterior distribution, respectively. The black dots indicate the height of the prior and posterior distributions at $a = a_0$. 
E.4 Robustness Analyses

In this section, we present the results of a robustness analysis. The aim of this analysis is to investigate the extent to which our conclusions are altered by the choice of the priors on the model parameters. Whereas we used uniform priors on the model parameters in the analyses presented in the main article (i.e., Beta(1, 1)), we repeat here the analyses with two different priors: either a Beta(1, 2) or a Beta(2, 1) distribution. The different prior distributions are visualized in Figure E.3.

It is evident that the Beta(1, 1) distribution puts equal mass on all parameter values, the Beta(1, 2) distribution favors smaller parameter values, whereas the Beta(2, 1) distribution favors larger parameter values.

![Figure E.3: Visualization of the different priors. The prior distribution shown in the left panel is used in the analyses presented in the main article, whereas the prior distributions present in the middle and right panel are used in the sensitivity analyses.](image)

Figures E.4 – E.6 show, separately for the three different prior distributions, 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 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 $M_1$ than model $M_2$.

Figures E.4 – E.6 show that there are some quantitative differences depending on which prior distribution is used. For example, the EV model is stronger supported when a Beta(2, 1) prior is used compared to the two other prior distributions. However, the qualitative conclusions are the same irrespective of the prior distribution; all three figures show that 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.

The findings from Figures E.4 – E.6 are corroborated by Table E.2. The second, fourth, and sixth column of Table E.2 show the median posterior model probabilities, and the third, fifth, and seventh column show the percentage of participants for whom the corresponding model has the largest posterior model probability, separately for the three different prior distributions. It is evident that the VPP model is supported the most; that is, the data from 52-59% 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., 25-30%). It is also evident that the EV model is stronger supported than the PVL-Delta model when a Beta(2, 1) prior is used—a finding that is reversed for the two other prior distributions.
E.4. Robustness Analyses

Figure E.4: Beta(1, 1) prior: Histograms of the log(BF) for pairwise comparison of four RL models applied to the IGT data from each of 771 participants (cf. Figure 2 in the main article). A positive log(BF) indicates that the data are more likely to occur under the first model than under the second model, whereas a negative log(BF) 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.
E. Appendix to Chapter 7: “Bayes Factors for Reinforcement-Learning Models of the Iowa Gambling Task”

Figure E.5: Beta(1, 2) prior: Histograms of the log(BF) for pairwise comparison of four RL models applied to the IGT data from each of 771 participants. A positive \( \log(BF) \) indicates that the data are more likely to occur under the first model than under the second model, whereas a negative \( \log(BF) \) 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)](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.
E.4. Robustness Analyses

Figure E.6: Beta(2, 1) prior: 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) indicates that the data are more likely to occur under the first model than under the second model, whereas a negative log(BF) 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 E.2: 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, for the three different prior distributions separately. Grey shaded cells refer to the best model.

<table>
<thead>
<tr>
<th></th>
<th>Beta(1, 1)</th>
<th></th>
<th>Beta(1, 2)</th>
<th></th>
<th>Beta(2, 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MPMP</td>
<td>%</td>
<td>MPMP</td>
<td>%</td>
<td>MPMP</td>
</tr>
<tr>
<td>EV</td>
<td>.00</td>
<td>7</td>
<td>.00</td>
<td>6</td>
<td>.00</td>
</tr>
<tr>
<td>PVL</td>
<td>.04</td>
<td>25</td>
<td>.04</td>
<td>30</td>
<td>.02</td>
</tr>
<tr>
<td>PVL-Delta</td>
<td>.00</td>
<td>9</td>
<td>.00</td>
<td>11</td>
<td>.00</td>
</tr>
<tr>
<td>VPP</td>
<td>.64</td>
<td>59</td>
<td>.49</td>
<td>52</td>
<td>.66</td>
</tr>
</tbody>
</table>

Figure E.7: Beta(1, 1) prior: 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 E.2), and the boxes indicate the interquartile range (i.e., the distance between the .25 and .75 quantiles).

The distributions of individual posterior model probabilities are visualized in Figures E.7 – E.9, which presents violin plots of the 771 posterior model probabilities for each of the four RL models, for the three different prior distributions separately. The dots indicate the median posterior model probability (cf. second, fourth, and sixth column of Table E.2), and the boxes indicate the interquartile range (i.e., the distance between the .25 and .75 quantiles). From Figures E.7 – E.9 it is evident that in the case of the EV, PVL, and PVL-Delta models, the individual posterior model probabilities follow a right skewed distribution suggesting that the data of most participants...
Figure E.8: Beta(1, 2) prior: 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. fourth column of Table E.2), and the boxes indicate the interquartile range (i.e., the distance between the .25 and .75 quantiles).

provide little evidence for these models. 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. Altogether Figures E.7 – E.9 suggest that there are only minor difference in the distributions of individual posterior model probabilities. To conclude, this robustness analysis suggests that our main conclusions are unaffected by the choice of the prior distribution.
E. Appendix to Chapter 7: “Bayes Factors for Reinforcement-Learning Models of the Iowa Gambling Task”

E.5 Comparison to BIC

In this section, we present the results of a model comparison study based on BIC for the same models and data pool as used in the article. The BIC is called post hoc fit criterion in the context of RL models for the IGT. Therefore, we call it here “BIC post hoc fit criterion”. The advantage of the BIC post hoc fit criterion is that it is easier to compute than the importance sampling Bayes factors. However, it should be kept in mind that the BIC post hoc fit criterion considers only one dimension of model complexity, that is, the number of parameters, and that the BIC post hoc fit criterion is derived as an asymptotic approximation of Bayesian model selection using Bayes factors (J. I. Myung, Cavagnaro, & Pitt, in press). Another popular measure is the Watanabe-Akaike information criterion (WAIC; Watanabe, 2010, 2013). However, WAIC is not suitable for our predictive goal, that is, to predict the next choice given all previous choices (Aki Vehtari, personal communication, 16.07.2014; see also a discussion on Andrew Gelman’s blog http://andrewgelman.com/2014/09/25/waic-time-series/, and Vehtari & Ojanen, 2012).
Table E.3: Median posterior model probabilities (MPMP), and percentage of participants for whom the corresponding model has the largest posterior model probability obtained from three different methods: (1) Importance sampling, and (2) BIC. Grey shaded cells refer to the best model.

<table>
<thead>
<tr>
<th>Model</th>
<th>Importance Sampling</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MPMP</td>
<td>%</td>
</tr>
<tr>
<td>EV</td>
<td>.00</td>
<td>7</td>
</tr>
<tr>
<td>PVL</td>
<td>.04</td>
<td>25</td>
</tr>
<tr>
<td>PVL-Delta</td>
<td>.00</td>
<td>9</td>
</tr>
<tr>
<td>VPP</td>
<td>.64</td>
<td>59</td>
</tr>
</tbody>
</table>

Computations of BIC

The BIC for model $M_\ell$ is defined as follows (Schwarz, 1978):

$$
BIC_{M_\ell} = -2 \log L_\ell + k_\ell \log n,
$$

(E.4)

where $L_\ell$ is the maximum likelihood of model $M_\ell$, $k_\ell$ is the number of free parameters of model $M_\ell$, and $n$ is the number of IGT trials (Wagenmakers, 2007; Worthy, Pang, & Byrne, 2013; but see also for example Ahn et al., 2008; Fridberg et al., 2010; and Yechiam et al., 2010, where the BIC post hoc fit criterion is computed for RL models relative to a baseline model). Thus, the first term in Equation E.4 (i.e., the log maximum likelihood) quantifies the goodness-of-fit, whereas the second term penalizes a model for its complexity. Note that for the sake of clarity we omitted the notation that indexes a specific participant.

Approximations of the Bayes Factor

The BIC score can be used to approximate the Bayes factor using the following equation (e.g., Wagenmakers, 2007):

$$
BF_{12} \approx \exp \left( \frac{BIC_{M_2} - BIC_{M_1}}{2} \right),
$$

(E.5)

Equation E.5 allows us to investigate whether the approximations of the Bayes factors are in line with Bayes factors obtained from importance sampling.

Results

Table E.3 shows the median posterior model probabilities (MPMP), and percentage of participants for whom the corresponding model has the largest posterior model probability obtained from two different methods: (1) Importance sampling, and (2) BIC post hoc fit criterion. Just as the Bayes factors obtained from importance sampling, the Bayes factors approximated with the BIC post hoc fit criterion suggest that the data of only a minority of participants provide strong evidence for the EV and PVL-Delta models. However, it is evident that in contrast to the Bayes factors obtained from importance sampling, Bayes factors approximated with the BIC post hoc fit criterion


---

2Since we did not use maximum likelihood to estimate the parameters, the fitting routine did not automatically provide us with $L_\ell$—the maximum likelihood of model $M_\ell$. However, we obtained $L_\ell$ by computing the likelihood of the parameter combination that corresponds to the maximum log posterior. The log posterior is automatically returned by Stan (i.e., called “lp__”). The BIC computation was confirmed by comparing our results obtained for the dataset of Worthy, Pang, and Byrne (2013) to the ones reported in the original article.
suggest that the data provide the most evidence for the PVL model and relatively little evidence for the VPP model. These findings are corroborated by Figure E.10 showing the distributions of the posterior model probabilities of all participants derived with the BIC post hoc fit criterion. This analysis illustrates the critique that the BIC prefers simple models that underfit the data (Burnham & Anderson, 2002). In this particular case, the VPP model is punished 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.
Appendix F

Appendix to Chapter 8: “A Tutorial on Bridge Sampling”

F.1 The Bridge Sampling Estimator as a General Case of Methods 1 – 3

In this section we show that the naive Monte Carlo, the importance sampling, and the generalized harmonic mean estimators are special cases of the bridge sampling estimator under specific choices of the bridge function $h(\theta)$ and the proposal distribution $g(\theta)$.

An overview is provided in Table F.1.

To prove that the bridge sampling estimator reduces to the naive Monte Carlo estimator, consider bridge sampling, choose the prior distribution as the proposal distribution (i.e., $g(\theta) = p(\theta)$), and specify the bridge function as $h(\theta) = 1/g(\theta)$. Inserting these specifications into Equation 8.12 yields:

$$
\hat{p}_4(y \mid h(\theta) = \frac{1}{g(\theta)}, g(\theta) = p(\theta)) = \frac{1}{N_2} \sum_{i=1}^{N_2} \frac{1}{p(\theta_i)} p(y \mid \tilde{\theta}_i) p(\tilde{\theta}_i) \frac{1}{N_1} \sum_{j=1}^{N_1} \frac{1}{p(\theta_j^*)} p(\theta_j^*)
$$

which is equivalent to the naive Monte Carlo estimator shown in Equation 8.6.

To prove that the bridge sampling estimator reduces to the importance sampling estimator, consider bridge sampling, choose the importance density as the proposal distribution (i.e., $g(\theta) = 1/g(\theta)$), and specify the bridge function as $h(\theta) = 1/g(\theta)$. Inserting these specifications into Equation 8.12 yields:

$$
\hat{p}_4(y \mid h(\theta) = \frac{1}{g(\theta)}, g(\theta) = p(\theta)) = \frac{1}{N_2} \sum_{i=1}^{N_2} p(y \mid \tilde{\theta}_i) = \frac{1}{N_2} \sum_{i=1}^{N_2} p(y \mid \tilde{\theta}_i), \quad \tilde{\theta}_i \sim p(\theta)
$$

Note that bridge sampling is also a general case of the Chib and Jeliazkov (2001) method of estimating the marginal likelihood using the Metropolis-Hastings acceptance probability (Meng & Schilling, 2002; Mira & Nicholls, 2004).
Table F.1: Summary of the Bridge Sampling Estimators for the Marginal Likelihood, and Its Special Cases: the Naive Monte Carlo, Importance Sampling, and Generalized Harmonic Mean Estimator

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimator</th>
<th>Samples</th>
<th>Bridge Function $h(\theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bridge sampling</td>
<td>$\frac{1}{N} \sum_{i=1}^{N} p(y \mid \hat{\theta}_i) \frac{p(\theta_i)}{g(\theta_i)}$</td>
<td>$\hat{\theta}_i \sim g(\theta)$, $\theta^*_j \sim p(\theta \mid y)$</td>
<td>$C = \frac{N_1}{N_2 + N_1} p(y \mid \theta)p(\theta) + \frac{N_2}{N_2 + N_1} p(y)g(\theta)$</td>
</tr>
<tr>
<td>Naive Monte Carlo</td>
<td>$\frac{1}{N} \sum_{i=1}^{N} p(y \mid \hat{\theta}_i)$</td>
<td>$\hat{\theta}_i \sim p(\theta)$</td>
<td>$\frac{1}{g(\theta)}$ and $g(\theta) = p(\theta)$</td>
</tr>
<tr>
<td>Importance sampling</td>
<td>$\frac{1}{N} \sum_{i=1}^{N} \frac{p(y \mid \hat{\theta}<em>i) \frac{p(\theta_i)}{g</em>{IS}(\theta_i)}}{g_{IS}(\theta_i)}$</td>
<td>$\hat{\theta}<em>i \sim g</em>{IS}(\theta)$</td>
<td>$\frac{1}{g_{IS}(\theta)}$ and $g(\theta) = g_{IS}(\theta)$</td>
</tr>
<tr>
<td>Generalized harmonic mean</td>
<td>$\left( \frac{1}{N} \sum_{i=1}^{N} \frac{g_{IS}(\theta^<em>_i)}{p(y \mid \theta^</em><em>i) \frac{p(\theta_i)}{g</em>{IS}(\theta^*_i)}} \right)^{-1}$</td>
<td>$\theta^*_i \sim p(\theta \mid y)$</td>
<td>$\frac{1}{p(y \mid \theta)p(\theta)}$ and $g(\theta) = g_{IS}(\theta)$</td>
</tr>
</tbody>
</table>

Note. $p(\theta)$ is the prior distribution, $g_{IS}(\theta)$ is the importance density, $p(\theta \mid y)$ is the posterior distribution, $g(\theta)$ is the proposal distribution, $h(\theta)$ is the bridge function, and $C$ is a constant. The last column shows the bridge function needed to obtain the special cases.

$g_{IS}(\theta)$, and specify the bridge function as $h(\theta) = 1/g_{IS}(\theta)$. Inserting these specifications into Equation [8.12] yields:

$$
\hat{p}_4(y \mid h(\theta)) = \frac{1}{g_{IS}(\theta)} g(\theta) = \frac{1}{g_{IS}(\theta)} \frac{1}{N_2} \sum_{i=1}^{N_2} \frac{p(y \mid \hat{\theta}_i) \frac{p(\theta_i)}{g_{IS}(\theta_i)}}{g_{IS}(\theta_i)}, \quad \hat{\theta}_i \sim g_{IS}(\theta), \quad \theta^*_j \sim p(\theta \mid y)
$$

$$
= \frac{1}{N_2} \sum_{i=1}^{N_2} \frac{p(y \mid \hat{\theta}_i) \frac{p(\theta_i)}{g_{IS}(\theta_i)}}{1 \cdot N_1} = \frac{1}{N_2} \sum_{i=1}^{N_2} \frac{p(y \mid \hat{\theta}_i) \frac{p(\theta_i)}{g_{IS}(\theta_i)}}{g_{IS}(\theta_i)}, \quad \hat{\theta}_i \sim g_{IS}(\theta),
$$

which is equivalent to the importance sampling estimator shown in Equation [8.7].

To prove that the bridge sampling estimator reduces to the generalized harmonic mean estimator, consider bridge sampling, choose the importance density as the proposal distribution (i.e., $g(\theta) = g_{IS}(\theta)$), and specify the bridge function as $h(\theta) = 1/(p(y \mid \theta)p(\theta))$. Inserting these specifications into Equation [8.12] yields:
\[
\hat{p}_4(y \mid h(\theta)) = \frac{1}{p(y \mid \theta) p(\theta)} g(\theta) = g_{IS}(\theta)
\]

\[
= \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} 1}{p(y \mid \tilde{\theta}_i) p(\theta_i) p(y \mid \tilde{\theta}_i)} \frac{1}{p(y \mid \theta^*_i) p(\theta^*_i) g_{IS}(\theta^*_i)} \frac{1}{N_1} \sum_{j=1}^{N_1} \frac{1}{p(y \mid \theta^*_j) p(\theta^*_j) g_{IS}(\theta^*_j)}
\]

\[
= \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} 1}{p(y \mid \theta^*_i)} \frac{1}{N_1} \sum_{j=1}^{N_1} \frac{1}{p(y \mid \theta^*_j) p(\theta^*_j) g_{IS}(\theta^*_j)}
\]

which is equivalent to the generalized harmonic mean estimator shown in Equation 8.8

F.2 Bridge Sampling Implementation: Avoiding Numerical Issues

In order to avoid numerical issues, we can rewrite Equation 8.15 in the following way:

\[
\hat{p}_4(y^{(t+1)}) = \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} \frac{l_{2,i}}{s_1 t_{2,i} + s_2 \hat{p}_4(y^{(t)})}}{\frac{1}{N_1} \sum_{j=1}^{N_1} \frac{1}{s_1 t_{1,j} + s_2 \hat{p}_4(y^{(t)})}}
\]

\[
= \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} \exp \left( \log(t_{2,i}) \right)}{\frac{1}{N_1} \sum_{j=1}^{N_1} \exp \left( \log(t_{1,j}) \right) + s_2 \hat{p}_4(y^{(t)})}
\]

\[
= \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} \exp \left( \log(t_{2,i}) \right) \exp \left(-t^*\right)}{\frac{1}{N_1} \sum_{j=1}^{N_1} \exp \left( \log(t_{1,j}) \right) \exp \left(-t^*\right) + s_2 \hat{p}_4(y^{(t)}) \exp \left(-t^*\right)}
\]

\[
= \frac{1}{\exp \left(-t^*\right)} \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} \exp \left( \log(t_{2,i}) - t^*\right)}{\frac{1}{N_1} \sum_{j=1}^{N_1} \exp \left( \log(t_{1,j}) - t^*\right) + s_2 \hat{p}_4(y^{(t)}) \exp \left(-t^*\right)}
\]
\[
\hat{r}(t) = \hat{p}_4(y)^{(t)} \exp(-l^*)
\]

so that
\[
\hat{p}_4(y)^{(t)} = \hat{r}(t) \exp(l^*)
\]

Then we obtain
\[
\hat{p}_4(y)^{(t+1)} = \exp(l^*) \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} \frac{\exp \left( \log(l_{2,i}) - l^* \right)}{s_1 \exp \left( \log(l_{2,i}) - l^* + s_2 \hat{p}_4^{(t)} \right)}}{\frac{1}{N_1} \sum_{j=1}^{N_1} \frac{1}{s_1 \exp \left( \log(l_{1,j}) - l^* + s_2 \hat{p}_4^{(t)} \right)}}
\]

\[
\hat{p}_4(y)^{(t+1)} \exp(-l^*) = \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} \frac{\exp \left( \log(l_{2,i}) - l^* \right)}{s_1 \exp \left( \log(l_{2,i}) - l^* + s_2 \hat{p}_4^{(t)} \right)}}{\frac{1}{N_1} \sum_{j=1}^{N_1} \frac{1}{s_1 \exp \left( \log(l_{1,j}) - l^* + s_2 \hat{p}_4^{(t)} \right)}}
\]

\[
\hat{r}(t+1) = \frac{\frac{1}{N_2} \sum_{i=1}^{N_2} \frac{\exp \left( \log(l_{2,i}) - l^* \right)}{s_1 \exp \left( \log(l_{2,i}) - l^* + s_2 \hat{p}_4^{(t)} \right)}}{\frac{1}{N_1} \sum_{j=1}^{N_1} \frac{1}{s_1 \exp \left( \log(l_{1,j}) - l^* + s_2 \hat{p}_4^{(t)} \right)}}
\]

Hence, we can run the iterative scheme with respect to \( \hat{r} \) which is more convenient because it keeps the terms in the sums manageable and multiply the result by \( \exp(l^*) \) to obtain the estimate of the marginal likelihood or, equivalently, we can take the logarithm of the result and add \( l^* \) to obtain an estimate of the logarithm of the marginal likelihood.

**F.3 Correction for the Probit Transformation**

In this section we describe how the probit transformation affects our expression of the generalized harmonic mean estimator (Equation 8.8) to yield Equation 8.9. Recall that we derived the generalized harmonic mean estimator using the following equality:

\[
\frac{1}{p(y)} = \int \frac{g_{ts}(\theta)}{p(y | \theta) p(\theta)} p(\theta | y) \, d\theta. \tag{F.1}
\]

For practical reasons, in the running example, we used a normal distribution on \( \xi \) as importance density. \( \xi \) was obtained by probit-transforming \( \theta \) (i.e., \( \xi = \Phi^{-1}(\theta) \)). In particular, the normal importance density was given by \( \frac{1}{\sigma} \phi \left( \frac{\xi - \mu}{\sigma} \right) \). Note that this importance density is a function of \( \xi \),
whereas the general importance density $g_{IS}$ in Equation F.1 is specified in terms of $\theta$. Therefore, to include our specific importance density to Equation F.1, we need to write it in terms of $\theta$. This yields $\frac{1}{\sigma} \phi \left( \frac{\Phi^{-1}(\theta) - \hat{\mu}}{\sigma} \right) \frac{1}{\phi(\Phi^{-1}(\theta))}$, where the latter factor comes from applying the change-of-variable method. Replacing $g_{IS}(\theta)$ in Equation F.1 by this expression, results in:

$$\frac{1}{p(y)} = \int \frac{1}{\sigma} \phi \left( \frac{\Phi^{-1}(\theta) - \hat{\mu}}{\sigma} \right) \frac{1}{\phi(\Phi^{-1}(\theta))} p(\theta \mid y) \, d\theta$$

\[= \mathbb{E}_{\text{post}} \left( \frac{1}{\sigma} \phi \left( \frac{\Phi^{-1}(\theta) - \hat{\mu}}{\sigma} \right) \frac{1}{\phi(\Phi^{-1}(\theta))} \right), \tag{F.2}\]

Rewriting results in:

$$p(y) = \left( \mathbb{E}_{\text{post}} \left( \frac{1}{\sigma} \phi \left( \frac{\Phi^{-1}(\theta) - \hat{\mu}}{\sigma} \right) \frac{1}{\phi(\Phi^{-1}(\theta))} \right) \right)^{-1},$$

which can be approximated as:

$$\hat{p}_3(y) = \left( \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \left( \frac{1}{\sigma} \phi \left( \frac{\Phi^{-1}(\theta_j^*) - \hat{\mu}}{\sigma} \right) \frac{1}{\phi(\Phi^{-1}(\theta_j^*))} \right) \right)^{-1}, \quad \theta_j^* \sim p(\theta \mid y),$$

samples from the posterior distribution

\[= \left( \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \left( \frac{1}{\sigma} \phi \left( \frac{\xi_j - \hat{\mu}}{\sigma} \right) \right) \right)^{-1}, \quad \xi_j = \Phi^{-1}(\theta_j^*) \text{ and } \theta_j^* \sim p(\theta \mid y), \tag{F.3}\]

which shows that the generalized harmonic estimate can be obtained using the samples from the posterior distribution, or the probit-transformed ones. In the online-provided code, we use the latter approach (see also Overstall & Forster, 2010).

### F.4 Details on the Application of Bridge Sampling to the Individual-Level EV Model

In this section, we provide more details on how we obtained the unnormalized marginal likelihood for a specific participant $s$, $s \in \{1, 2, \ldots, 30\}$, with choices $Ch_s(T)$ and corresponding payoffs
Since we focus on one specific participant, we drop the subscript \( s \) in the remainder of this section. As explained in Appendix F.2, we run the iterative scheme with respect to \( \hat{r} \) to avoid numerical issues. Consequently, we have to compute \( \log(l_1) \) and \( \log(l_2) \). Using \( \tilde{\zeta}_i = (\tilde{\omega}_i, \tilde{\alpha}_i, \tilde{\gamma}_i) \) for the \( i^{th} \) sample from the proposal distribution, we get for \( \log(l_2) \) (\( \log(l_1) \) works analogously):

\[
\log(l_2, i) = \log \left( \frac{p(Ch(T) \mid \Phi(\tilde{\zeta}_i), X(T)) p(\Phi(\tilde{\zeta}_i)) \phi(\tilde{\zeta}_i)}{g(\tilde{\zeta}_i)} \right).
\]

Therefore, instead of computing the unnormalized logarithm of the unnormalized posterior distribution directly, we compute the logarithm of the unnormalized posterior distribution:

\[
\log(p(Ch(T) \mid \Phi(\tilde{\zeta}_i), X(T)) p(\Phi(\tilde{\zeta}_i)) \phi(\tilde{\zeta}_i)) = \log(p(Ch(T) \mid \Phi(\tilde{\zeta}_i), X(T))) + \log(\phi(\tilde{\omega}_i)) + \log(\phi(\tilde{\alpha}_i)) + \log(\phi(\tilde{\gamma}_i)),
\]

because we assumed independent priors on each model parameter \( w, a, c \). \( \log(p(\Phi(\tilde{\zeta}_i))) = 0 \) because \( p \) refers to the uniform prior on \([0, 1]\).

### F.5 Details on the Application of Bridge Sampling to the Hierarchical EV Model

Analogous to the last section, we explain here how we obtained the logarithm of the unnormalized posterior for the hierarchical implementation of the EV model. Using \( Ch_s(T) \) to refer to all choices of subject \( s \), \( X_s(T) \) for the corresponding net outcomes, \( \tilde{\zeta}_{s,i} = (\tilde{\omega}_{s,i}, \tilde{\alpha}_{s,i}, \tilde{\gamma}_{s,i}) \) for the \( i^{th} \) sample from the proposal distribution for the individual-level parameters of subject \( s \), and \( \tilde{\xi}_i \) for the \( i^{th} \) sample from the proposal distribution for all group-level parameters (e.g., \( \tilde{\xi}_i = (\tilde{\mu}_{\omega,i}, \tilde{\tau}_{\omega,i}, \tilde{\mu}_{\alpha,i}, \tilde{\tau}_{\alpha,i}, \tilde{\mu}_{\gamma,i}, \tilde{\tau}_{\gamma,i}) \)), we get:

\[
\log \left( \prod_{s=1}^{30} p(Ch_s(T) \mid \Phi(\tilde{\zeta}_{s,i}), X_s(T)) p(\tilde{\zeta}_{s,i} \mid \tilde{\xi}_i) \right) p(\tilde{\xi}_i)
\]

\[
= \sum_{s=1}^{N} \log(p(Ch_s(T) \mid \Phi(\tilde{\zeta}_{s,i}), X_s(T)])[+ \log \left( \frac{1}{1.5\Phi(\tilde{\omega}_{s,i})} \phi \left( \frac{\tilde{\omega}_{s,i} - \tilde{\mu}_{\omega,i}}{1.5\Phi(\tilde{\omega}_{s,i})} \right) \right) + \log \left( \frac{1}{1.5\Phi(\tilde{\alpha}_{s,i})} \phi \left( \frac{\tilde{\alpha}_{s,i} - \tilde{\mu}_{\alpha,i}}{1.5\Phi(\tilde{\alpha}_{s,i})} \right) \right) + \log \left( \frac{1}{1.5\Phi(\tilde{\gamma}_{s,i})} \phi \left( \frac{\tilde{\gamma}_{s,i} - \tilde{\mu}_{\gamma,i}}{1.5\Phi(\tilde{\gamma}_{s,i})} \right) \right) + \log(\phi(\tilde{\mu}_{\omega,i})) + \log(\phi(\tilde{\mu}_{\alpha,i})) + \log(\phi(\tilde{\mu}_{\gamma,i})) + \log(\phi(\tilde{\tau}_{\omega,i})) + \log(\phi(\tilde{\tau}_{\alpha,i})) + \log(\phi(\tilde{\tau}_{\gamma,i})).
\]
Appendix to Chapter 9: “Bayesian Techniques for Analyzing Group Differences in the Iowa Gambling Task: A Case Study of Intuitive and Deliberate Decision Makers”

In this appendix we describe in detail the experiment that was conducted to obtain the data reported in the main article, and the product space method that was used for the Bayes factor model comparison.

G.1 Experiment

Material

Iowa Gambling Task

The IGT was presented as a computerized task, based on the original version of Bechara et al. (1994). On the computer screen, four decks of cards were presented, labeled “A”, “B”, “C”, and “D”. Participants were initially given a (hypothetical) loan of +2000 Swiss Francs (CHF) and instructed to consecutively choose among the decks, by clicking on one of the decks at each trial, resulting in a draw of a card from the deck, so as to maximize their long-term net outcome (cf. Bechara et al. 1994, 1997). After each selection of a deck, participants received feedback on the gains as well as the losses (if any) associated with the card, and the running tally. The trials were self-paced.

\footnote{In contrast to the payoff scheme introduced by Bechara et al. (1994), we used a stable loss of $-50 in deck C.}
Measurement of decision styles

To measure individual participants' decision style, we used an inventory compiled by [Betsch and Iannello (2010)](https://example.com), whose subscales are taken from five different questionnaires: the Rational-Experiential Inventory (REI; [Pacini & Epstein](https://example.com) [1999]), the Preference for Intuition and Deliberate Scale (PID; [Betsch](https://example.com) [2004]), the General Decision Making Style (GDMS; [Scott & Bruce](https://example.com) [1995]) questionnaire, the Cognitive Style Indicator (CoSI; [Cools & van den Broeck](https://example.com) [2007]), and the Perceived Modes of Processing Inventory (PMPI; [Burns & D’Zurilla](https://example.com) [1999]). All of these questionnaires measure a person's tendency to rely on an intuitive and a deliberate decision mode on two separate bipolar subscales. For instance, participants are presented with statements such as “My feelings play an important role in my decisions.” (intuition subscale of the PID), or “Before making decisions, I first think them through.” (deliberation subscale of the PID). At each item, participants are asked to indicate the extent to which the statement represents their opinion (on a scale from 1 = very much disagree to 7 = very much agree). The original versions of the REI, PID and the GDMS contain items that include the term “intuition”. [Betsch and Iannello (2010)](https://example.com) argued that this might activate different concepts across people. These items were therefore excluded from [Betsch and Iannello (2010)](https://example.com)'s compiled inventory. Altogether, the questionnaire consisted of 70 items from 12 subscales. As described in more detail below, we distinguished intuitive and deliberate decision makers based on their total scores on these subscales.

### Procedure

Participants completed the experiment individually. They signed an informed consent form and started the experiment with the IGT, followed by demographic questions and a computerized version of the decision-style inventory. Then they were thanked, debriefed, and received course credits or a flat fee of 7.50 CHF—a decision that had to be made before the experiment—as well as a performance-contingent bonus from their IGT performance (specifically, final IGT score/1000 * 1.5 CHF).

Table G.1: Three-factor solution of the principal component analysis. Also reported are Cronbach’s α as a measure of the reliability of each subscale.

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Deliberation</th>
<th>Intuition</th>
<th>Spontaneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deliberation (GDMS) (α = .80)</td>
<td>.890</td>
<td>.009</td>
<td>-.184</td>
</tr>
<tr>
<td>Deliberation (PID) (α = .74)</td>
<td>.875</td>
<td>-.170</td>
<td>-.094</td>
</tr>
<tr>
<td>Knowing (CoSI) (α = .79)</td>
<td>.873</td>
<td>-.067</td>
<td>-.126</td>
</tr>
<tr>
<td>Rational ability (REI) (α = .77)</td>
<td>.820</td>
<td>-.148</td>
<td>.176</td>
</tr>
<tr>
<td>Rational engagement (REI) (α = .81)</td>
<td>.659</td>
<td>-.105</td>
<td>-.145</td>
</tr>
<tr>
<td>Planning (CoSI) (α = .70)</td>
<td>.602</td>
<td>-.325</td>
<td>.130</td>
</tr>
<tr>
<td>Experiental ability (REI) (α = .84)</td>
<td>.058</td>
<td>.870</td>
<td>.165</td>
</tr>
<tr>
<td>Intuition (GDMS) (α = .69)</td>
<td>-.122</td>
<td>.850</td>
<td>.056</td>
</tr>
<tr>
<td>Intuition (PID) (α = .79)</td>
<td>-.151</td>
<td>.809</td>
<td>.018</td>
</tr>
<tr>
<td>Experiental engagement (REI) (α = .41)</td>
<td>-.321</td>
<td>.604</td>
<td>.033</td>
</tr>
<tr>
<td>Automatic (PMPI) (α = .77)</td>
<td>.114</td>
<td>.126</td>
<td>.897</td>
</tr>
<tr>
<td>Spontaneous (GDMS) (α = .73)</td>
<td>-.354</td>
<td>.075</td>
<td>.805</td>
</tr>
</tbody>
</table>
Decision Styles

We first determined for each participant the mean score on each of the 12 subscales of the decision-style inventory compiled by Betsch and Iannello (2010). Table G.1 shows that all subscales had acceptable levels of internal reliability, except for the experiential engagement subscale of the REI. However, we decided to keep the experiential engagement subscale in our analyses because excluding it did not lead to different conclusions in the subsequent analyses. Based on each participant’s mean score on each of the 12 subscales, we then conducted a principal component analysis with rotation based on the varimax method. The Kaiser criterion suggested a three-factor solution, which accounted for 70% of the total variance. Table G.1 reports the factor loadings of the 12 subscales on these three factors. On the first factor the subscales capturing a deliberate, rational, and planned decision style showed consistently high loadings (deliberation factor). The second factor had consistently high loading for the subscales capturing an intuitive and experiential decision style (intuition factor). The third factor had high loadings of subscales measuring spontaneous decision making (spontaneity factor). Individually, the three factors accounted for 39%, 18.5%, and 12.5% of the variance, respectively.

In the following, we distinguish participants as deliberate or intuitive decisions makers depending on their factor score on the deliberation factor and the intuition factor. Following previous research (Betsch & Kunz, 2008), we classified participants as intuitive if they had a factor

---

2Although conceptually similar to “intuition”, based on the factor solution the spontaneity factor seems to
score above the median of the intuition factor and, at the same time, a factor score below the median of the deliberation factor; participants with the opposite pattern were classified as deliberate. This classification scheme yielded 19 participants in the intuitive group and 19 participants in the deliberate group. Thirty-two participants thus remained unclassified. Figure G.1 shows the distribution of scores on the 12 subscales of the questionnaire compiled by Betsch and Iannello (2010), separately for the intuitive and deliberate group. As can be seen, the groups have strongly different profiles on the scales and cover different value ranges.

G.2 Obtaining Bayes Factors with the Product Space Method

In this section we describe how we obtained the Bayes factor with the product space method (Carlin & Chib 1995; Lodewyckx et al. 2011). The Bayes factor $BF_{12}$ is defined as the change from prior model odds $\frac{p(M_1)}{p(M_2)}$ of two models, $M_1$ and $M_2$, to posterior model odds $\frac{p(M_1 | D)}{p(M_2 | D)}$ brought about by the data $D$:

$$\frac{p(M_1 | D)}{p(M_2 | D)} = \frac{p(M_1)}{p(M_2)} \times \frac{m(D | M_1)}{m(D | M_2)}$$  \hspace{1cm} \text{(G.1)}$$

For all but the simplest models the Bayes factor cannot be derived analytically. We therefore need a method to approximate the Bayes factor. One such method is the product space method (for alternative methods such as reversible jump, see Green 2003; Sisson 2005, and for importance sampling, see Hammersley & Handscomb 1964; Steingroever et al. 2016; Vandekerckhove et al. 2015). The product space method is a transdimensional Markov chain Monte Carlo (MCMC) method, that is, a method that aims to estimate the posterior model odds for chosen prior model odds (see Equation [G.1]). This method requires the construction of a “supermodel” encompassing the models to be compared. This “supermodel” is a hierarchical combination of the models to be compared. The hierarchical combination is achieved by a model index that measures the proportion of times that either model is visited to account for the observed data. The prior of the model index corresponds to the prior model odds (i.e., specified before the analysis), and the posterior of the model index corresponds to the posterior model odds. The posterior model index can be estimated by MCMC posterior sampling methods. We can therefore estimate the posterior probability of model $M_k$ using:

$$\hat{p}(M_k | D) = \frac{\text{Number of occurrences of } M_k}{\text{Total number of iterations}}.$$  \hspace{1cm} \text{(G.2)}$$

The posterior model probability quantifies the evidence that the data $D$ provide for model $M_k$ relative to all other models under consideration (Berger & Molina 2005). Given the estimated posterior model probabilities of two different models, we can estimate the Bayes factor using Equation [G.1] because the prior model odds are known (i.e., specified before the analysis).
Appendix $H$

Appendix to Chapter 10: “Using Bayesian Regression to Incorporate Covariates into Hierarchical Cognitive Models”

H.1 Complete Results of the Simulation Study

In this appendix we provide the results of our simulation study for all four parameters of the PVL-Delta model. Figure H.1 gives the log Bayes factors for all PVL-Delta parameters from our simulations with uncorrelated covariates. Dark grey dots show the Bayes factors obtained in the regression analysis, light grey dots show the Bayes factors obtained in the median-split analysis. Recall that our simulated data were generated so that the first covariate would be positively correlated with the $A$ parameter and the second covariate would be negatively correlated with the $w$ parameter. The correlations between $A$ and the second covariate, and between $w$ and the first covariate were set to 0, and the relationships between the remaining model parameters and the covariates were set to the values estimated from Steingroever et al.'s (submitted) data, and were negligible. As described in the main text, the Bayes factors from the regression analysis showed strong evidence for an effect of the first covariate on the $A$ parameter (dark grey dots, left column in the top row), whereas the median-split analysis provided much weaker evidence for such an effect (light grey dots, left column in the top row). Similarly, the regression analysis strongly supported an effect of the second covariate on the $w$ parameter (dark grey dots, second column in the bottom row), whereas the median-split analysis provided weaker evidence for such an effect (light grey dots, second column in the bottom row). For the null-effects of the first covariate on the $w$ parameter (second column, top row) and of the second covariate on the $A$ parameter (left column, bottom panel), both analyses performed similarly without any appreciable differences in Bayes factors. Finally, both analyses provided only weak support if any for an effect the covariates on the $a$ and $c$ parameters and there were no sizable differences in Bayes factors.

Figure H.2 gives the log Bayes factors for all PVL-Delta parameters from our simulations with correlated covariates. As described in the main text, the Bayes factors obtained from the regression analysis again showed stronger evidence for an effect of the first covariate on the $A$ parameter (dark
\[ \rho(X_1, X_2) = 0 \]

Figure H.1: Bayes factors from 50 simulated data sets for the regression and median-split analysis with uncorrelated covariates. Data points show the log Bayes factors for the alternative hypothesis (log(BF$_{10}$)) obtained in the regression (RG, dark grey dots) and median-split (MS, light grey dots) analysis for the PVL-Delta model’s A and w parameters (columns) and two covariates (rows). Lines indicate the mean log BF. Data points are jittered along the x-axis for improved visibility.
grey dots, left column in the top row) than the median-split analysis (light grey dots, left column in the top row). Similarly, the regression analysis provided stronger support for an effect of the second covariate on the $w$ parameter (dark grey dots, second column in the bottom row) than the median-split analysis (light grey dots, second column in the bottom row). However, unlike in the case of uncorrelated covariates, in the case of correlated covariates the median-split analysis now suggested spurious effects of the first covariate on the $w$ parameter (second column, top row) and of the second covariate on the $A$ parameter (left column, bottom row). Finally, the regression as well as the median-split analysis did not provide strong evidence for any effects of the covariates on the $a$ and $c$ parameters, and there were no clear differences in Bayes factors visible between the two analyses.

Taken together, these results illustrate that, in the case of uncorrelated covariates, a median-split analysis tends to understate the evidence for existent effects. In the case of correlated covariates, a median-split analysis also understates the evidence for existent effects but in addition suggests spurious effects of covariates on model parameters that are in fact unrelated. Furthermore, our results show that in cases where model parameters do not have any appreciable relationships with any of the covariates, as was the case for the $a$ and $c$ parameters, regression and median-split analyses perform similarly and there are no appreciable biases associated with a dichotomization-based analysis.

Figure H.3 shows the posterior means of the standardized effect sizes estimated in the regression analysis (RG, left panels in each subplot) and the posterior means of the standardized effect sizes estimated in the median-split analysis (MS, right panels in each subplot). The left subplot shows results for the case of uncorrelated covariates, the right subplot shows the results for the case of correlated covariates. The top row shows the results for the first covariate, the bottom row for the second covariate. The results are complementary to the results for the Bayes factors. In the case of uncorrelated covariates (left subplot), the estimated effects in both models are largest for effects that we created to be non-zero (i.e., the effect of the first covariate on $A$ and the effect of the second covariate on $w$, leftmost column of the panels in the top row and second-from-left column of the panels in the bottom row, respectively). Moreover, both models correctly estimate the direction of the effect of the first covariate on the $A$ parameter to be positive (leftmost column of the panels in the top row), and the direction of the effect of the second covariate on $w$ to be negative (second-from-left column of the panels in the bottom row). Both models also correctly estimate the effects of the first and second covariate on $a$ and $c$ to be close to 0 (second-from-right and rightmost columns of each panel).

In the case of correlated covariates (right subplot), both analyses again correctly estimate the size and direction of the strong effects of the first covariate on the $A$ parameter (leftmost column of the panels in the top row) and of the second covariate on the $w$ parameter (second-from-left column of the panels in the bottom row). However, whilst the regression analysis correctly estimates the relationships between the first covariate and the $w$ parameter (left panel, second-from-left column in the top row) and between second covariate and the $A$ parameter (left panel, leftmost column in the bottom row) to be approximately 0, the median-split analysis suggests a weakly negative association between the first covariate and $w$ (right panel, second-from-left column in the top row) and between the second covariate and $A$ (right panel, leftmost column in the bottom row). Finally, both models correctly estimate the effects of the covariates on the $a$ and $c$ parameters to be close to 0.

These results align well with the results for the Bayes factors. In the case of uncorrelated covariates, the regression analysis as well as the median-split analysis correctly indicate the direction and size of the relationships between covariates and model parameters. However, in the case of correlated covariates, the median-split analysis tends to suggest spurious relationships
Figure H.2: Bayes factors from 50 simulated data sets for the regression and median-split analysis with correlated covariates. Data points show the log Bayes factors for the alternative hypothesis (log(BF_{10})) obtained in the regression (RG, dark grey dots) and median-split (MS, light grey dots) analysis for the PVL-Delta model's $A$ and $w$ parameters (columns) and two covariates (rows). Lines indicate the mean log BF. Data points are jittered along the x-axis for improved visibility.
H.2 Priors for the Regression Coefficients

In this appendix we provide detailed information about the priors for the regression coefficients in our model. The overarching goal in regression analysis as used in the present work is to select the set of predictors that best accounts for the observed data. This selection process can be understood as a model comparison exercise. Consider, for example, a situation where a researcher has a criterion variable $y$ and a single predictor variable $x$ and wants to know whether $x$ has any predictive value for $y$. To answer this question, the researcher constructs two models, a null model $\mathcal{M}_0$ that only includes an intercept term, and an alternative model $\mathcal{M}_1$ that includes an intercept term and the predictor variable $x$, and compares the adequacy of the two models. Following the Bayesian approach, such model comparisons can be carried out by computing the
Bayes factor, that is, the ratio of the marginal likelihood of the observed data under the two models, 
\( \text{BF}_{10} = \frac{p(y \mid M_1)}{p(y \mid M_0)} \).

Bayes factors need to fulfill a number of theoretical desiderata (Bayarri, Berger, Forte, & García-Donato, 2012; Rouder & Morey, 2012). First, Bayes factors should be location and scale invariant. In the case of regression models, this means that the scale on which the criterion and predictor variables are measured (e.g., kilograms, grams, milligrams) and the location of the zero-point of the scale (e.g., temperature in Celsius vs. in Kelvin) should not influence the Bayes factor. Second, Bayes factors should be consistent, which means that as sample size approaches infinity, the Bayes factor should converge to the correct bound (i.e., \( \text{BF}_{10} \to 0 \) if \( M_0 \) is correct and \( \text{BF}_{10} \to \infty \) if \( M_1 \) is correct). Third, Bayes factors should be consistent in information, which means that the Bayes factor should not approach a finite value as the information provided by the data in favor of the alternative model approaches infinity. In the case of regression models this means that, as the coefficient of determination, \( R^2 \), in \( M_1 \) approaches 1, the Bayes factor should go to infinity.

One factor that critically influences the behavior of the Bayes factors is the choice of the priors for the model parameters. Assigning improper priors to model-specific parameters, for instance, leads to indeterminate Bayes factors (Jeffreys, 1961). In our example with a single predictor \( x \), the corresponding regression weight \( \beta_x \) is included in \( M_1 \) but not in \( M_0 \). If \( \beta_x \) is assigned an improper prior that is only determined up to a multiplicative constant, this constant will appear in the numerator of the Bayes factor but not in the denominator, which means that it will not cancel and the Bayes factor will depend on the multiplicative constant. Consequently, researchers need to choose the prior distribution for the model parameters in such a way that model comparisons yield Bayes factors with the desired theoretical properties.

An additional criterion in selecting priors for the model parameters is the degree to which priors are noninformative. In many situations, researchers have little information about the range in which the model parameters, that is, the regression weights, should fall. Therefore, the weights should be assigned a prior that puts little constraint on the possible values. One prior that has regularly been used in regression problems is Zellner’s g-prior (Zellner, 1986). In the case of \( P \) predictor variables and \( N \) observations for each variable, this prior takes the form:

\[
\beta | g \sim \mathcal{N}(0, g \sigma^2 (X^T X)^{-1})
\]

where \( \beta \) is the \( P \times 1 \) vector of regression weights, \( g \) is a scaling factor, \( \sigma^2 \) is the residual variance of the criterion variable, \( 0 \) is \( P \times 1 \) vector of zeros, \( X \) is the \( N \times P \) design matrix containing the predictor variables, and \( \mathcal{N} \) denotes the multivariate normal distribution. The degree to which this prior is informative is controlled by its variance-covariance matrix, which in turn depends on \( g \), \( \sigma^2 \), and \( X \). In Zellner’s conceptualization of this prior, the design matrix should be treated as a constant; the prior can then be interpreted as the prior on the regression weights arising from a repetition of the experiment with the same design matrix. The reciprocal of the residual variance, that is, the precision \( \phi \) should be assigned Jeffreys’ prior \( p(\phi) \propto 1/\phi \). Finally, the scaling factor \( g \) controls the weight given to the prior relative to the data. For example, \( g = 10 \) means that the data are given 10 times as much weight as the prior. The scaling factor thus controls how peaked or how informative the prior is.

Another way to understand the effect of the scaling parameter is to consider the shrinkage factor \( g/(1 + g) \) (Liang et al., 2008; Wetzels, Grasman, & Wagenmakers, 2012). Using this shrinkage factor, the posterior mean for \( \beta \) can be estimated as the product of the shrinkage factor and the least-squares estimate of the regression weights, \( \hat{\beta}_{OLS} \). Consequently, if \( g \) is set to a small value, the posterior estimate of \( \beta \) will be pulled towards 0 whereas a high value of \( g \) leads to a posterior mean that is similar to the least-squares estimate.
The question that remains is how \( g \) should be set. One popular choice is to set \( g = N \), which yields a unit information prior (Kass & Raftery, 1995). Specifically, the term \( \sigma^2(X^TX)^{-1} \) in the expression for the variance-covariance matrix of the prior equals the variance of the maximum-likelihood estimator of the regression weights, \( \text{var}(\beta_{\text{OLS}}) \). This estimate is based on the design matrix with \( N \) rows, which conveys the information of \( N \) observations. Therefore, by setting \( g \) to \( N \), the influence of the design matrix on the prior can be made equivalent to the information contained in a single observation.

However, as shown by Liang et al. (2008), the Zellner prior in its general form suffers from two shortcomings. First, if \( g \) is set to a fixed value, the resulting Bayes factors will suffer from the “information paradox”. This means that when a model \( M_1 \) is compared to the null model \( M_0 \), and the coefficient of determination \( R^2 \) under \( M_1 \) goes to 1 (i.e., there is infinite support for \( M_1 \)), the Bayes factor will tend to a finite value, thus violating the theoretical desideratum of consistency in information. Second, if \( g \) is set to a very large value to make the prior noninformative, Bayes factors will suffer from the Jeffreys-Lindley-Bartlett paradox. This means that \( M_0 \) will unduly be favored. In the limiting case when \( g \rightarrow \infty \), the Bayes factor will go to zero, irrespective of the information provided by the data, thus violating the theoretical property of consistency.

The problems of the Zellner prior are resolved by Jeffreys-Zellner-Siow prior (JZS prior; Nuijten, Wetzels, Matzke, Dolan, & Wagenmakers, 2015). In the approach suggested by Zellner and Siow (1980), the regression coefficients are assigned a multivariate Cauchy prior, which satisfies the consistency requirements on the Bayes factors (Liang et al., 2008; Wetzels et al., 2012). One slight drawback, however, of the multivariate Cauchy prior is that the marginal likelihood of the data under a model cannot be computed in closed form and numerical approximations require the computation of the \( P \)-dimensional integrals, which become unstable for models with large numbers of predictors. As pointed out by Liang et al. (2008), a remedy to this problem is to express the multivariate Cauchy distribution as a mixture of \( g \)-priors. In this approach, the scaling factor \( g \) in the Zellner prior is treated as a random variable:

\[
\beta | g \sim \mathcal{N}(0, g\sigma^2(X^TX/N)^{-1}),
\]

and \( g \) is assigned an inverse-gamma prior:

\[
g \sim \mathcal{IG}(1/2, s^2/2)
\]

with shape parameter 1/2 and scale parameter \( s^2/2 \). Note that the scale parameter \( s \) of the inverse gamma distribution is equal to the scale parameter of the multivariate Cauchy distribution. This form of the JZS prior combines the favorable theoretical properties of the multivariate Cauchy prior with the computational advantages of the mixture representation. Specifically, using the mixture representation of the JZS prior reduces the computation of a Bayes factor to a one-dimensional integral that can be computed numerically with great precision (Rouder & Morey, 2012).

The above discussion shows that using the JZS prior yields Bayes factors that are consistent and consistent in information, and thus satisfy two of the three desiderata. The third desideratum, location and scale invariance, can be achieved by assigning the JZS prior to the vector of standardized effect sizes \( \alpha \), rather than the vector of regression weights \( \beta \). The elements of the vector of standardized effect sizes are given by:

\[
\alpha_i = \beta_i \frac{sx_i}{\sigma},
\]

\footnote{Note that because the design matrix appears in the expression for the prior in the inverse of the matrix \((X^TX)^{-1}\), multiplying \((X^TX)^{-1}\) by \( N \) is equivalent to diving \((X^TX)\) by \( N \).}
where $\beta_i$ is the regression weight for the $i$th predictor variable, $s_{x_i}$ is the standard deviation of the $i$th predictor variable and $\sigma$ is the residual standard deviation of the criterion variable. Returning to our initial example with a single predictor variable $x$ and criterion variable $y$, the standardized effect size is given by $\alpha_x = \beta_x \frac{s_x}{\sigma}$. The researcher could now assign $\alpha_x$ a JZS prior and compute the marginal likelihood of the data under model $M_1$, which includes an intercept term and $x$ as a predictor, and under model $M_0$, which only includes an intercept term. The resulting Bayes factor will then satisfy all three desiderata, being invariant to linear transformations of $x$ and $y$, favoring the correct model as the sample size goes to infinity, and not approaching a finite asymptote as evidence for $M_1$ approaches infinity.
References


References


Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W., & Nathan, P. E. (2001). Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed...
References


References


References


quantitative and qualitative changes in decision making ability. *Behavioural Neurology*, 19, 59 - 63.

JASP Team. (2015). JASP (version 0.7) [computer software].


impaired and negative symptoms in schizophrenia. Psychiatry Research, 152, 121 - 128.


Stan Development Team. (2016b). Stan (Version 2.9.0) [Computer software]. 183 185 191


Steingroever, H., Pachur, T., Šmíra, M., & Lee, M. D. (submitted). Bayesian techniques for analyzing group differences in the Iowa gambling task: A case study of intuitive and deliberate decision makers. 121 149 150 176 179 180 181 182 183 184 190 210 211 251


Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2013a). A comparison of reinforcement-learning models for the Iowa gambling task using parameter space partitioning. The Journal of Problem Solving, 5, Article 2. 56 59 60 61 62 64 69 76 77 81 82 84 86 87 100 104 112 113 114 149 150 160 161 210


Steingroever, H., Wetzels, R., & Wagenmakers, E.-J. (2015). \( \hat{w} = 2, \hat{a} = .8, \hat{c} = .6 \): So what? On the meaning of parameter estimates from reinforcement-learning models. Decision, 2, 228 - 235. 112


272
References


273


Nederlandse Samenvatting

Hoe kiezen mensen tussen twee opties die verschillen in hun onmiddellijke en langdurige consequenties? Zou je bijvoorbeeld voor een heerlijk stukje taart gaan of eerder voor een appel? Het stukje taart vind je nu misschien lekkerder, maar als je daar vaker voor kiest, zou dat op lange termijn slecht voor je gezondheid kunnen zijn. In dit proefschrift, getiteld “Risicovolle Beslissingen Veilig Modelleren”, heb ik onderzocht hoe mensen zogenoemde risicovolle beslissingen nemen. Hiervoor heb ik gebruik gemaakt van data van de Iowa gambling taak (IGT; Bechara et al., 1994). Deze taak bestaat uit vier speelstapels (Figuur 12.1). Deelnemers worden gevraagd om herhaald kaarten te kiezen van de vier verschillende stapels met als doel een zo hoog mogelijke netto-uitkomst te realiseren. De taak is zodanig ontwikkeld dat dit doel alleen maar bereikt kan worden als je de goede stapels de voorkeur geeft en de slechte stapels vermijdt. Wat karakteriseert de goede en slechte stapels? Om de verschillen tussen de vier stapels te verduidelijken heb ik in Tabel 12.1 de winsten en verliezen van tien voorbeeldkeuzes van elke stapel gegeven. Het wordt duidelijk dat de slechte stapels elke keer een hoge winst geven, maar af en toe een heel hoog verlies; daarom zijn deze stapels op lange termijn ongunstig. De twee andere stapels – de goede stapels – geven elke keer een redelijk lage winst, maar ze zijn op lange termijn gunstig omdat hun onregelmatige verliezen heel laag zijn.

In dit proefschrift stelde ik dat descriptieve analyses van IGT data ons alleen maar kunnen

![Figuur 12.1: Voorbeeld van een schermafbeelding van de IGT.](image)
vertellen welke keuzes mensen hebben gemaakt en hoe hun keuzegedrag verandert naarmate de taak voortschrijdt. Echter, om erachter te komen hoe mensen de keuzes hebben gemaakt, d.w.z. welke psychologische processen hun keuzes hebben bepaald, moeten we cognitieve modellen gebruiken. Deze modellen maken aannames over de relevante psychologische processen, zoals motivatie (bijv. winsten als net zo belangrijk waarnemen als verliezen), geheugen (bijv. het vermogen eerdere uitkomsten van de stapels te herinneren), en responsconsistentie (bijv. de neiging beslissingen te baseren op verwachtingen van de stapels) en hoe deze processen interacteren om tot uiting te komen in het geobserveerde keuzegedrag. In de context van de IGT zijn zogenaamde reinforcement-learning (RL) modellen heel geschikt (Sutton & Barto, 1998). Deze modellen verklaren hoe mensen door trial and error hun beslissingen vormen.

De reden waarom de modellen toepasbaar zijn op verschillende patronen in de data is dat de modellen parameters hebben die kunnen variëren. Deze modellparameters staan voor de psychologische processen die verondersteld worden belangrijk te zijn voor risicovolle besluitvorming. Een voorbeeld is een modellparameter die beschrijft hoe belangrijk mensen onmiddellijke winsten achten. Een lage waarde van de parameter betekent dat mensen onmiddellijke winsten niet belangrijk vinden, terwijl een hoge parameterwaarde betekent dat mensen onmiddellijke winsten heel belangrijk vinden. Op deze manier kunnen de modellen gebruikt worden om erachter te komen waarom verschillende groepen verschillende keuzes maken. Hier hebben we echter parameterschattingen nodig voor de verschillende groepen. Om deze te verkrijgen moeten we de modellen passen op de data (zgn. model fitting).

In de klinische psychologie worden de IGT samen met RL modellen vaak gebruikt om erachter te komen of een klinische groep suboptimale beslissingen neemt, en zo ja, hoe dit verklaard kan worden. Voor dat soort toepassingen is het dus van groot belang dat de IGT daadwerkelijk risicovolle besluitvorming meet en dat de modellen de relevante processen correct beschrijven. In dit proefschrift beweer ik dat de huidige manier om IGT data te analyseren, de huidige manier om RL modellen toe te passen en de huidige manier om een geschikt RL model te vinden (zgn. modelselectie) gekenmerkt worden door veel problemen. Deze problemen leiden ertoe dat er vaak conclusies worden getrokken die voorbarig zijn. Om dit in de toekomst te voorkomen en om conclusies te trekken die sterker gebaseerd zijn op de data stel ik een aantal maatregels voor. Deze regels hebben betrekking op de volgende gebieden: (1) data-analyse; (2) modelselectie; (3) model fitting; en (4) beoordeling van de model fit.

In dit proefschrift beweer ik dat elk van de vier zojuist genoemde gebieden enorm kan profiteren
van het gebruik van de Bayesiaanse statistiek in plaats van de klassieke, of de zogenoemde frequentistische statistiek. De Bayesiaanse statistiek heeft de laatste jaren sterk aan populariteit gewonnen en is zeer geschikt voor problemen in de mathematische psychologie (Andrews & Baguley, 2013; Bayarri et al., 2016; Poirier, 2006; Vanpaemel, 2016; Verhagen et al., 2015; Wetzels et al., 2016). In de rest van deze samenvatting zal ik de belangrijkste resultaten van mijn proefschrift samenvatten.

Het doel van hoofdstuk 2 was te verduidelijken dat de huidige manier waarop IGT data wordt geanalyseerd het zogenoemde frequency-of-losses effect verbergt. Het frequency-of-losses effect staat voor de bevinding dat het gezonde proefpersonen vaak niet lukt een voorkeur te ontwikkelen voor de goede stapels (d.w.z. stapels C en D). Ze leren uiteindelijk enkel de stapels met regelmatige verliezen (d.w.z. stapels A en C) te vermijden. Dit keuzegedrag zorgt echter niet voor maximalisatie van de netto-uitkomst en duidt volgens de oorspronkelijke interpretatie van de IGT op suboptimale besluitvorming. In hoofdstuk 2 heb ik nog twee andere bevindingen gerapporteerd die aantonen dat belangrijke veronderstellingen over het keuzegedrag van gezonde proefpersonen niet door de data ondersteund kunnen worden. Ik heb namelijk aangetoond dat gezonde mensen (1) idiosyncratisch keuzegedrag vertonen (zie Figuur 12.2); en (2) niet een bepaalde systematiek aanhouden waarbij ze eerst de verschillende stapels verkennen en daarna de beste stapels plunderen. Deze bevindingen stellen de heersende interpretatie van IGT data in twijfel en suggereren dat er bij toekomstige toepassingen van de IGT bijzondere aandacht besteed moet worden aan de belangrijkste veronderstellingen over het keuzegedrag van gezonde mensen. In het bijzonder is het belangrijk dat de data van elke stapel apart geanalyseerd worden en opgesplitst in meerdere blokken van trials. In hoofdstuk 9 heb ik bovendien voorgesteld een Bayesiaanse herhaalde metingen variantieanalyse (ANOVA) te gebruiken. Met zo’n ANOVA kan je bijvoorbeeld onderzoeken of proefpersonen meer goede keuzes maken naarmate de taak voortschrijdt, en of er verschillen zijn in de voorkeuren voor de goede stapels tussen twee verschillende groepen. Een groot voordeel van de Bayesiaanse ANOVA is dat deze ook gebruikt kan worden om de bewijskracht van de data te bepalen voor de nulhypothese die ervan uitgaat dat er geen effect is (bijv. geen groepseffect of geen blokeffect).

Op grond van de empirische relevantie van het frequency-of-losses effect (d.w.z. een voorkeur hebben voor stapels B en D) in IGT data van gezonde proefpersonen wilde ik vervolgens in hoofdstuk 3 onderzoeken of RL modellen dit datapatroon überhaupt kunnen genereren. Dit is belangrijk omdat we, als we modelparameters interpreteren als representatie van de psychologische processen, zeker moeten weten dat het model goed past op de data. Maar als het model het keuzegedrag van de data überhaupt niet kan genereren, kunnen we ervan uitgaan dat het model niet op de data kan passen. In zo’n situatie kunnen we de parameterschattingen beter niet gebruiken om uitspraken te doen over de onderliggende psychologische processen. De methode die ik hiervoor gebruikt heb heet parameter space partitioning (PSP).

De resultaten van hoofdstuk 3 suggereren dat drie bekende RL modellen – het Expectancy Valence model (EV; Busemeyer & Stout, 2002), het Prospect Valence Learning model (PVL; Ahn et al., 2008), en een combinatie van deze twee modellen, het EV-PU model – geen duidelijke stapelvoorkeuren kunnen genereren, en dat het EV model helemaal geen frequency-of-losses effect kan genereren. Samengenomen suggereren deze resultaten dat de drie beschouwde modellen alleen een voldoende data-passend potentieel hebben voor een beperkt aantal keuzepatronen, en dat er nog een model ontbreekt dat universeel gebruikt kan worden.

Daarna heb ik in hoofdstuk 4 laten zien dat het PVL-Delta model – een RL model dat recentelijk veel aandacht heeft gekregen en ook een combinatie is van het EV en PVL model (Ahn et al., 2008; Fridberg et al., 2010; Steingroever et al., 2014) – een beter data-passend potentieel heeft dan de modellen uit hoofdstuk 3. Maar ook dit model kan niet universeel gebruikt worden omdat het geen
duidelijke voorkeur voor de slechte stapels kan genereren—een keuzepatroon dat voorondersteld wordt voor mensen met suboptimale besluitvorming.

Naast de PSP methode hebben ik een aantal andere methodes besproken die geschikt zijn voor modelselectie. Ten eerste is het belangrijk dat modellen parameters accuraat terugschatten (zie hoofdstukken 4 en 6). Deze eigenschap kan onderzocht worden met zogenoemde parameter recovery studies. Voor dit soort studies genereren we data met eigen gekozen parameterwaardes en met het model onder beschouwing. Vervolgens passen we het model op de gegenereerde dataset om parameterschattingen te verkrijgen. Als de parameterschattingen overeenkomen met de waardes die we voor de datageneratie hebben gebruikt, spreken we van een accurate parameter recovery. Ten tweede, om te garanderen dat de modelparame

Om modelparameters ook daadwerkelijk voor de veronderstelde psychologische processen staan, is het belangrijk dat het model goed scoort op een zogenoemde test van selectieve invloed (zie hoofdstuk 4). Zo'n test onderzoekt of experimentele manipulaties die bedoeld zijn om alleen maar één modelparame te bevloeden, zich weerspiegelen in de parameterschattingen. Ten derde heb ik in hoofdstukken 7, 8, en 9 modellen vergeleken met behulp van de Bayes factor. De Bayes factor is in de Bayesiaanse statistiek het standaard middel om modellen te vergelijken. Maar deze was nog niet eerder toegepast in RL modellen voor de IGT omdat de Bayes factor wiskundig gezien moeilijk te verkrijgen is. De relevante code is online beschikbaar waardoor ik hoop dat de Bayes factor in de toekomst vaker gebruikt wordt om modellen te vergelijken.

Om modellen te passen hebben ik gebruik gemaakt van het Bayesiaanse hierarchische raamwerk (hoofdstukken 4, 5, 8, en 10). Dit raamwerk leidt tot nauwkeurigere en informatievere schattingen dan alternatieve methoden omdat het raamwerk gebruik maakt van zowel de verschillen als gemeenschappelijkheden van proefpersonen van een groep (Ahn et al., 2011; Horn et al., 2015; Lejarraga et al., 2016; Navarro et al., 2006; Rouder & Lu, 2005; Rouder et al., 2005, 2008; Scheibehenne & Pachur, 2015; Shiffrin et al., 2008; Wetzels, Vandekerckhove, et al., 2010).

Om tenslotte te beoordelen hoe goed een model op de data past, heb ik laten zien dat het belangrijk is de absolute model fit te beoordelen (hoofdstukken 4, 6, en 9). Dit is in strijd met de huidige manier waarin alleen maar de relatieve model fit wordt beschouwd—een methodo die bepaalt hoeveel beter het RL model onder beschouwing op de data past dan een eenvoudig referentimodel. Dit geeft een getal dat helaas moeilijk te interpreteren is. Dit getal zegt alleen of het RL model beter past, maar niet hoeveel beter en ook niet of het model überhaupt voldoende op de data past (bijv. of het model kwalitatief hetzelfde datapatron voorspelt als door de data vertoond wordt). Het zou dus kunnen dat het RL model onder beschouwing volgens de relatieve index beter past op de data dan het referentimodel, maar eigenlijk passen beide modellen redelijk slecht. Om te bepalen of het RL model daadwerkelijk op de data past, is het daarom beter om te kijken of data gegenereerd met de parameterschattingen hetzelfde keuzegedrag opleveren als de data zelf. Alleen in deze situatie kunnen we ervan uitgaan dat de parameterschattingen voor de relevante psychologische processen staan.

Dat het zo belangrijk is de absolute model fit te bekijken heb ik in hoofdstuk 5 verduidelijkt met behulp van de EV, PVL, en PVL-Delta modellen. Hiervoor hebben ik gebruik gemaakt van twee gestileerde datasets en vijf gepubliceerde datasets. Ik heb eerst de modellen op de data gepast, en vervolgens heb ik met de parameterschattingen data gegenereerd. De geobserveerde data van de twee gestileerde datasets staan in de eerste kolom van Figuur 12.3 en de voorspellingen van de drie modellen in kolommen 2 – 4. Dit figuur verduidelijkt dat alleen het PVL-Delta model beide datapatronen kwalitatief correct voorspelt. In het geval van het EV model en de gestileerde dataset met een voorkeur voor de stapels met incidentele verliezen, zien we bijvoorbeeld dat het EV model een voorkeur voor de slechte stapels voorspelt, dus een kwalitatief verschillend datapatron. Dit suggereert dat we de parameterschattingen beter niet kunnen gebruiken om conclusies te

trekken over de psychologische processen. Om voorbarige conclusies te voorkomen in toekomstige toepassingen van RL modellen op IGT data is het daarom van groot belang dat de absolute fit van de modellen zorgvuldig wordt bekeken.

Als we de hierboven genoemde belangrijke stappen nastreven komen we hopelijk dichter bij het ambitieuze doel de IGT en cognitieve modellen te gebruiken om suboptimale besluitvorming te diagnosticeren en beter te begrijpen.
Dit proefschrift is tot stand gekomen mede dankzij de betrokkenheid en inbreng van een groot aantal personen die ik hier graag wil bedanken. Ik wil graag beginnen met mijn begeleider Eric-Jan. Dat ik bij Eric-Jan (en überhaupt in Amsterdam) terecht ben gekomen heb ik aan Stefan Schmukle te danken. Als derdejaars Bachelorstudent vertelde ik hem dat ik op zoek was naar een masterprogramma waarin ik mijn interesse in psychologie en wiskunde kon combineren. Voor hem was het meteen duidelijk dat ik met Eric-Jan moest samenwerken. In Amsterdam aangekomen, was ik snel overtuigd dat dit advies klopte. Gelukkig was Eric-Jan het hier ook mee eens, en ben ik snel zijn AiO geworden. Dat deze tijd heel speciaal is geworden heb ik grotendeels aan Eric-Jan te danken. Eric-Jan is een tomeloze ideeën-generator, ontzettend gedreven, positief en heel enthousiast, wat ongelofelijk motiverend en inspirerend voor mij was. Maar Eric-Jan is ook ontzettend kritisch waardoor ik heel veel heb geleerd! Bovendien gaf hij mij ook de vrijheid om mijzelf verder te ontwikkelen. Ik ben ontzettend dankbaar dat Eric-Jan mij heeft ondersteund om naast mijn promotieonderzoek eindelijk toe te kunnen geven aan mijn wens om wiskunde te studeren. Tenslotte ben ik Eric-Jan ontzettend dankbaar voor zijn gezelligheid, zijn gulheid, zijn brede interesse (ook buiten de wetenschap), zijn bereikbaarheid en alle karakterbuilding.

De leden van mijn promotiecommissie bedankt ik voor het lezen en de zorgvuldige beoordeling van mijn proefschrift: bedankt Raoul Grasman, Hilde Huizenga, Michael Lee, Tim Pleskac, Han van der Maas en Ingmar Visser.

I would also like to thank my co-authors for the interesting and inspiring discussions and all their contributions. In addition, I am very grateful for all the shared data that I was allowed to use for my research. This dissertation could not have been realized without this support.

Grote dank gaat ook naar mijn UvA-collega’s die voor een heel fijne (werk)sfeer zorgden zowel binnen als buiten de geweldige vakgroep methodenleer. Het was ontzettend fijn met jullie van gedachten te wisselen en met jullie samen te werken. Fijn dat we ook altijd een reden voor een borrel of teamuitje wisten te verzinnen zodat er wekelijks een drank after-work programma was, en we zelfs samen op vakantie of weekend retreats gingen!

Mijn lieve paranimfen Dora en Sanne. Ik ben jullie ontzettend dankbaar dat jullie deze taak op jullie willen nemen en dat jullie er altijd voor mij waren de afgelopen jaren. Dorika, wat was het fijn met jou een kamer te delen! Je kritische blik en je zorgvuldigheid maakten vaak indruk op mij. Het was ontzettend fijn dat je steeds mee wilde denken en ook steeds een oplossing wist of, zo nodig, afleiding kon bieden. Maar nog fier vind ik dat het niet bij het delen van werkgerelateerde vragen is gebleven, maar we ook hebben genoten van het delen van dagelijkse verhalen en veel witte wijntjes met spa rood! En dan mijn andere paranimf Sanne. Inhoudelijk hebben Sanne en ik het nooit over ons promotieonderzoek gehad, maar we weten wel dat we allebei dezelfde
ondersozeksmentaliteit hebben. Daardoor en door het feit dat we van dezelfde dingen enthousiast worden, hebben we in onze vrije tijd met veel plezier naar diverse onderwerpen grondig onderzoek gedaan en zouden we zomaar een aantal nieuwe proefschriften kunnen schrijven. Sanne, ik ben je ontzettend dankbaar voor je vriendschap, je steun, en al de gekke momenten die we de laatste jaren hebben samen meegemaakt. Hierdoor heb je er ontzettend veel aan bijgedragen dat ik mij in Amsterdam helemaal thuis voel en indirect dat mijn promotietijd zo speciaal is geworden.

Ik heb me de afgelopen jaren regelmatig afgevraagd of mijn proefschrift had afgekregen zonder mijn zwemmers! Soms zat ik na mijn werk met een hoofd vol gedachten op mijn fiets en kostte het mij enige overtuiging toch voor een duik in het koude water te kiezen. Maar eenmaal ingezwommen was het elke keer weer een feest. Lieve zwemmers, ik ben zo ontzettend blij dat ik jullie heb leren kennen en dat we zulke goede vrienden zijn geworden. Trots ben ik ook op ons douche-record van 45 minuten wat wel duidelijk maakt dat de gezelligheid niet tekort kwam. Na zo’n zwemtraining zat ik weer helemaal fluitend op mijn fiets, genietend van het mooie Amsterdam op weg naar mijn fijne PhD huis.

Ja, mijn PhD huis! Wat was het een luxe om in zo’n mooie omgeving te wonen samen met zulke lieve mensen die allemaal heel verschillend waren, maar die toch ook allemaal iets gemeen hadden: ze kwamen uit het buitenland naar Amsterdam om hier te promoveren. Inmiddels ben ik verhuisd, maar kijk ik graag en vaak terug op de gezellige tijd aan de Montelbaanstraat. Special thanks to Gorka, Kathrin, Hugh and Ymke, thanks so much for your gezelligheid, great company, for sharing PhD life frustrations, just for being there! You made our PhD house a place where I always felt so much at home!

Publications


