Decision Making

Comparing Decision-Making Models of the Iowa Gambling Task

Authors:

[redacted]

[redacted] [redacted]

Lecturer:

Andreas Højlund

January 10th, 2023

Abstract

Decision-making is a complex cognitive process which when impaired greatly impacts quality of life. This is evident from subjects with damage to the vmPFC and subjects who suffer from substance-use disorders. Understanding decision-making and the latent psychological processes which composes it is thereby important. Prospect theory describes how people make decisions under risk. The decisions which prospect theory is based on, however, differ greatly from real-life decisions. Furthermore, some research has found that prospect theory might not be able to describe decisions which are based on experience. This has lead to the suggestion that decision-making should be differentiated into decision from description (i.e., with stated probabilities) and decision from experience, where the latter includes a learning component. This is explored by fitting and validating the PVL-Delta model and the recently introduced ORL model on Iowa Gambling Task data ($N = 232$). The PVL-Delta model is based on prospect theory while the ORL model is not. The two models are explored using hierarchical Bayesian analysis with Markov chain Monte Carlo simulations. We find that ORL fits the behavior of the data better, albeit the results should be interpreted with care as the Markov chains did not converge. Though the results do point in the direction of prospect theory not being necessary to model experience-based decision-making data, it is premature to conclude that two different models are needed as the ORL remains largely unexplored. We propose a number of future directions to further explore the relationship between decisions from description and decisions from experience.

Contents

1 Introduction

(K) Decision-making is an essential part of daily life. Consequently, deficits in decision-making as seen in people suffering from substance-use disorder (SUD) and brain injuries influence the quality of life greatly (Bechara et al., 2001; Salmond et al., 2005). To elucidate the latent psychological processes which facilitate decision-making, cognitive models have been applied to various decision-making tasks. Many studies have attempted to model experience-based decision-making using prospect theory (Steingroever et al., 2013b; D. Fridberg et al., 2010; Ahn et al., 2008). Some have, however, claimed that prospect theory cannot be used to model experience-based decision-making tasks, because it is based on decisions from description. The claim is corroborated by findings of Weber et al. (2004), Barron and Erev (2003), and Erev (1998), who found that the data differed vastly depending on the specific conditions of the study. This led to the suggestion of differentiating between two types of decision-making: decision from description and decision from experience (Hertwig et al., 2004).

In decision from description people are explicitly told the underlying probability distributions of the possible outcomes. Contrary, in decision from experience, the underlying probability distribution is not explained and the probabilities must be learned by sampling. Hertwig et al. (2004) called for two theories of decision-making, because they found that decision from description leads to an overweighting of probabilities of rare events as opposed to decision from experience which leads to an underweighting of these events (Weber et al., 2004; Barron and Erev, 2003; Erev, 1998). When reanalyzing the data, Fox and Hadar (2006) claimed that the finding of Hertwig et al. was due to a sampling error of the experiment. That is, the participants did not experience a representative sample of the underlying probability distribution during sampling. Fox and Hadar (2006) subsequently concludes that the behavioral pattern in the decision from experience condition can be modeled with prospect theory if the statistical sampling error is accounted for. Hence, Fox and Hadar do not recognize the need for two models of decision-making as suggested by Hertwig et al. (2004).

(S) Thus, opposing views currently govern the field of decision-making and it is not clear whether prospect theory can describe the behavior in both decision from description and decision from experience scenarios or if modeling the two phenomena separately is more meaningful. This controversy is investigated in our analysis which compares the Prospect Valence Learning with Delta-learning rule (PVL-Delta) (Ahn et al., 2008; Ahn et al., 2014; Steingroever et al., 2013b) with the Outcome-Representation Learning model (Haines et al., 2018). PVL-Delta is based on prospect theory, while the ORL model is not. To understand the basic assumption of the PVL-Delta, prospect theory is elaborated in the following section.

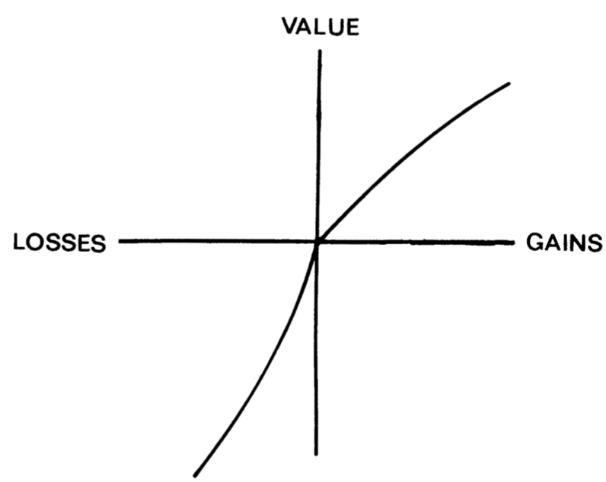
1.1 Prospect theory

(S) In 1979, Daniel Kahneman and Amos Tversky proposed prospect theory to describe decision-making under risk (Kahneman and Tversky, 1979). The theory builds upon the discovery that people do not make decisions based solely on the probabilities of receiving a specific outcome. Rather, they found that decisions are dependent on a subjective value assigned to the potential losses and gains of the decision, which is not always equal to the objective outcome. They formulated this as a subjective utility function (Kahneman and Tversky, 1979; Barberis, 2013).

A hypothetical subjective utility function can be seen in Figure 1. The x-axis depicts the monetary losses and gains and the y-axis depicts the subjective value which one relates to a specific loss or gain. Kahneman and Tversky (1979) found that in general, the slope is concave for gains and convex for losses, with a steeper slope for losses. The latter indicates that people are in general specifically sensitive to losses, and the former that the value of a gain or loss depends on the change, rather than the absolute magnitude of the gain or loss. That is, the difference in e.g., gaining 100$ or 200$ has a greater difference in subjective value than the difference in gaining 1100$ or 1200$. The difference in subjective value depends on a reference point which depends, among other things, on the wealth of the individual (Kahneman and Tversky, 1979).

Figure 1

A hypothetical subjective utility function according to prospect theory



Note. Reprinted from Kahneman and Tversky (1979) Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47(2), 263-291, https://doi.org/10.2307/1914185.

Prospect theory is based on hypothetical problems with stated probabilities (Kahneman and Tversky, 1979). The decisions which prospect theory was developed to predict are thereby made from *description.*

Prospect theory has mainly been applied in experimental settings, and it has proven difficult to apply the theory to economic analysis besides the explicit setting in which it was originally tested (Barberis, 2013). Yet, prospect theory is recognized as the best available descriptive theory of evaluating decisions under risk (Barberis, 2013). Now, it is clear the most choices made in real life are not accompanied by a stated probability. Thus, the question is whether prospect theory can be applied to experimental paradigms of decision-making where the hypothetical problem is not presented with explicit probabilities but rather the probabilities are learned through *experience*.

1.2 The Iowa Gambling Task

(K) The Iowa Gambling Task (IGT) was introduced to investigate decision-making abilities in subjects with damage to the ventromedial prefrontal cortex (vmPFC) (Bechara et al., 1994). It has since been used to differentiate subjects with various SUDs (Bechara et al., 2001; Ahn et al., 2014).

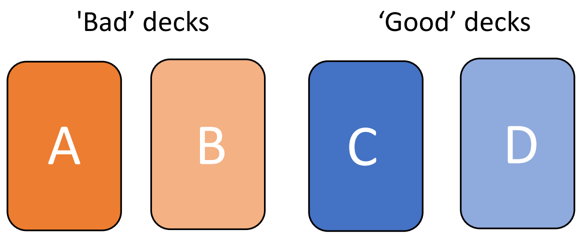
In the IGT, subjects are presented with four card decks (A, B, C, or D), which are alike in appearance. Each round, they are asked to choose a card from a deck which yields a subsequent monetary gain which sometimes is accompanied by a monetary penalty. The penalty might exceed the gain and thereby result in a loss. Participants begin the task with a loan of 2000$ and are instructed to choose cards to maximize their outcome. The participants are stopped when they reach 100 trials (Bechara et al., 1994).

Unbeknownst to the participants, the gain and/or loss of each deck on each trial is predetermined from a payoff structure, which is the same across participants. Choosing cards from deck A and B yields higher gains, but the losses are also greater, resulting in a net loss. Vice versa, choosing cards from deck C and D yields smaller gains, but the losses are smaller, resulting in a net gain. Thus, deck A and B are 'bad' decks, while deck C and D are 'good' decks. The net gain and the net loss of the 'good' and 'bad' decks, respectively, are the same, but the frequency of the losses differ (Bechara et al., 1994). The IGT assumes that the those with intact decision-making skills learn to choose the 'good' decks over the 'bad'.

The traditional IGT has a stationary payoff structure while the modified IGT increases the net gain or loss of 'good' and 'bad' decks, respectively, every ten trials. As choices are made, it should thereby be more apparent which decks are 'good' and which are 'bad' (Bechara et al., 2001) (see Figure 2).

Figure 2

The magnitude and frequency of gains and losses in the IGT



Traditional IGT

	A	B	C	D
Gain amount	100$	100$	50$	50$
Loss amount	150$ - 350$	1250$	25$ - 75$	250$
Likelihood of gain	50%	90%	50%	90%
Net sum after 100 trials	-250$	-250$	+250$	+250$

Traditional IGT

trial/deck	A	B	C	D
1	+100	+100	+50	+50
2	+100	+100	+50	+50
3	-50	+100	0	+50
4	+100	+100	+50	+50
5	-200	+100	0	+50
6	+100	+100	+50	+50
7	-100	+100	0	+50
8	+100	+100	+50	+50
9	-150	-1150	0	+50
10	-250	+100	0	-200
Net sum	*-250*	*-250*	*+250*	*+250*

Modified IGT

trial/deck	A	B	C	D
1	+100	+100	+50	+50
2	+120	+80	+60	+40
3	-70	+110	-10	+45
4	+90	+120	+55	+45
5	-190	+90	+5	+55
6	+100	+100	+45	+60
7	-120	+90	0	+40
8	+120	+120	+45	+55
9	-140	-1140	+10	+50
10	-260	+80	-10	-190
Net sum	*-250*	*-250*	*+250*	*+250*

Note. Top: Figure modified from (Haines et al., 2018). The color shades of the decks indicate the frequency of losses (darker = more frequent, lighter = less frequent). In deck A, the value of losses vary from 150$ to 350$ with 50$ increments. In deck C, losses vary from 25$ to 75$ with 25$ increments. Gains are fixed values. Bottom: The net outcome of the first ten trials in the traditional (left), and the modified (right) IGT. In the modified IGT, the net loss of 'bad' decks increases in increments of 150$ until it reaches 1000$ and the net gain of 'good' decks increases in increments of 50$ until it reaches 375$. The gains of respective decks are centered around a mean value (Bechara et al., 2001). Tables are based on the supplementary material of Steingroever et al. (2015).

1.3 Cognitive modeling of IGT data

(S) With the objective of understanding decision-making, cognitive models have been applied to decision-making tasks such as the IGT. A model which sufficiently describes IGT is interesting is because it might be generalizable to other experience-based tasks. In the following section, different cognitive strategies which might affect participants' performance in the IGT are elaborated. Their implementation in the PVL-Delta and ORL, respectively, is briefly described.

1.3.1 Expected value

(K) As participants draw cards from the different decks in the IGT, they develop an expectation to the value which a given deck will generate (Haines et al., 2018). Those with normal decision-making abilities will thereby learn to use the expected value of decks to avoid the 'bad' decks and draw more cards from the 'good' decks. Thus, the IGT facilitate decision-making based on participants' processing of expected value (Bechara et al., 1994; Beitz et al., 2014). The expected value is thereby an essential parameter to model.

In the PVL-Delta model, decision-makers are assumed to estimate the expected value of decks in accordance with a subjective utility function from prospect theory (Ahn et al., 2014). In the ORL model, the expected value of decks is estimated based on the outcome of the previous choice (Haines et al., 2018). Both models update the expected value using the Delta-learning rule (Rescorla and Wagner, 1972). While the PVL-Delta includes a single learning parameter, the ORL models separate learning rates for gains and losses to capture differences in the processing of gains or losses (Haines et al., 2018).

1.3.2 Win frequency

(S) Participants often prefer decks which frequently yield a gain disregarding the expected value of said deck (Haines et al., 2018). For example, in contrary to the findings of Bechara et al. (1994), participants have been found to prefer deck B (which has a higher win frequency) over deck A (which has a lower win frequency), though they have the same net loss as seen in Figure 2 (Steingroever et al., 2013a). According to Hertwig et al. (2004) this could be explained by the propensity to underweight rare events. Thus, the extent to which win frequency affect the choices of participants is important to model.

The PVL-Delta implicitly models the win frequency through the subjective utility function (Haines et al., 2018), while the ORL explicitly accounts for win frequency.

1.3.3 Choice perseverance

(K) Participants of the IGT differ in how much they tend to explore the different decks versus exploiting what they have learned. Choice perseverance thereby seeks to represent the extent to which participants continue to choose the same deck regardless of the its outcome (Haines et al., 2018).

PVL-Delta models choice perseverance implicitly when calculating the probability of choosing a deck on the following trial. The ORL does that too, but is also includes choice perseverance explicitly by using a decay parameter which represents how much of their recent history of choices participants remember (Haines et al., 2018).

1.3.4 Reversal learning

(S) It is essential that a model of experience-based decision-making can quickly reverse its preferences (i.e., what has been learned) when a large loss is encountered (Haines et al., 2018). For example, after the first eight trials of the IGT, many prefer deck B due to its large reward. However, on the ninth trials, deck B yields a large loss which calls for a need to reverse deck preference, as seen in Figure 2 (Haines et al., 2018). The behavioral characteristics of reversal learning has been well-replicated (Roese and Summerville, 2005, as cited in Haines et al., 2018) and is well supported in cognitive modeling of experience-based tasks, like the IGT (Gläscher et al., 2008; Hampton et al., 2006, as cited in Haines et al., 2018).

Reversal-learning is not incorporated in the PVL-Delta, however it is implemented in the ORL by calculating the win frequency for the *unchosen* decks (Haines et al., 2018).

1.4 Research Question

The question of whether it is necessary to differentiate between description-based and experience-based decision-making is examined as formulated in the following research question:

- Can the ORL model sufficiently describe the IGT data and does it perform better than the PVL-Delta model?

The question is investigated by fitting, validating and comparing the two reinforcement learning models.

2 Methods

2.1 Iowa Gambling Task data

(K) The data was acquired from Steingroever et al. (2015), who collected IGT data sets across several labs. All labs used a computerized version of the IGT (Steingroever et al., 2015).

Data sets where the payoff scheme was randomized between participants and where the IGT was altered from the traditional or the modified IGT were excluded from the present data pool. For comparison reasons, data sets which had more or less than 100 trials were excluded. The final data pool consists of five studies, which resulted in a total of 232 participants. The demographics of the included studies can be found in Table 1.

Table 1

Data sets included in the current study

Study	Number of participants	IGT version	Demographics
Kjome et al. (2010)	19	modified	M = 33.9 years (SD = 11.2), 6 female
Premkumar et al. (2008)	25	modified	M = 35.4 years (SD = 11.9), 9 female
Wood et al. (2005)	153	modified	M = 45.25 years (SD = 27.21)
Worthy et al. (2013)	35	traditional	Undergraduate students, 22 female

2.2 Model specifications

2.2.1 Prospect Valence Learning with Delta rule (PVL-Delta)

(S) The Prospect Valence Learning with Delta rule (PVL-Delta) seeks to model decision-making by use of four different parameters, which are thought to represent different psychological processes of decision making (Ahn et al., 2008; Ahn et al., 2014; Steingroever et al., 2013b; Haines et al., 2018). The model was first introduced by Ahn et al. (2008) and has since been found to reliably model IGT data (Steingroever et al., 2013b).

The PVL-Delta relies on Prospect theory (Tversky and Kahneman, 1992) in that it assumes that each participant has a subjective utility of each deck, which is updated when a given deck is chosen based on the outcome of the previous choice. The subjective utility is calculated as

$$u_t^d \leftarrow \begin{cases} {X_{t-1}}^A, & \text{if } X_{t-1} \geq 0 \\ -w\,|X_{t-1}|^A, & \text{otherwise} \end{cases} \tag{2.1}$$

where u_t^d is the subjective utility of the chosen deck d on trial t, X_{t-1} is the net outcome on the previous trial. The free parameters are $A \in (0, 2)$ which is a risk aversion parameter that governs the shape of the utility function, and $w \in (0, 10)$ which represents loss aversion (Steingroever et al., 2013b).

How A and w affect the the subjective utility function has been explored in Figure 9 in Appendix A. It is seen that when the risk aversion parameter A is closer to one, the relationship between the subjective utility and the objective net outcome is more linear, while when A is closer to zero, the shape of utility function looks more like a step function (Steingroever et al., 2013b). So, when A is closer to one, there is less risk aversion, while when A is closer to zero or greater than one, there is more risk aversion.

If the net outcome on the previous trial is negative (i.e., there is a loss), then the subjective utility is modeled by use of the w parameter, which represents sensitivity to loss. When w is closer to 0, negative outcomes will to a higher extent be disregarded. When w is closer to one, the impact on the subjective utility is approaching that of the positive net outcomes but with the sign reversed. Finally, if w is greater than one, negative outcomes have a stronger impact on the subjective utility than positive outcomes, and the loss aversion is thereby higher (Steingroever et al., 2013b).

The subjective utility of the chosen deck is used to update the expected utility or value of said deck, while keeping the expected utilities of the other decks constant (Steingroever et al., 2013b). This is done by use of a simplified Delta learning rule (Rescorla and Wagner, 1972; Haines et al., 2018). The expected value (or expected utility), Ev_t^d, is calculated as

$$Ev_t^d \leftarrow Ev_{t-1}^d + a\left(u_t^d - Ev_{t-1}^d\right) \tag{2.2}$$

where the expected utility Ev_t^d of the chosen deck d on trial t is given by the expected utility on the previous trial, Ev_{t-1}, the subjective utility u_t^d on the current trial t, and the learning rate $a \in (0, 1)$. The learning rate parameter a, governs how quickly the participant will update their expected utility of the chosen deck based on the previous trial. If a is closer to one, the participant weighs the net outcome of the previous trial higher and thereby changes the expected outcome quicker, while if a is closer to zero, the net outcome of the previous trial influences the expected utility less (Haines et al., 2018).

The expected utilities of each deck d is entered into a softmax function to calculate the probability P_{t+1}^d of choosing a given deck d on the following trial $t+1$ (Haines et al., 2018). The softmax function is defined as

$$P_{t+1}^d \leftarrow \frac{\exp\left(\theta Ev_t^d\right)}{\sum_{i=1}^4 \exp\left(\theta Ev_t^i\right)}, \tag{2.3}$$

where θ represents choice consistency i.e. how consistent the participant is in choosing a given deck. θ is constrained to positive values and a smaller value of θ indicates that the participant tends to explore by changing deck preference frequently. Larger values of θ indicate that the choice of the participant is strongly related to the expected values of each deck thereby relying more on exploitation (Steingroever et al., 2013b). We used a simplified version of θ because we did not find that Haines et al. (2018) and Steingroever et al. (2013b) sufficiently justified their use of reparameterization.

Thus, the PVL-Delta has four parameters: A, which shapes the utility function and thereby the risk aversion of the participant, w, which represents loss aversion, a, which represents how quickly the participant learns, and θ, which functions as a choice consistency parameter.

2.2.2 Outcome-Representation Learning (ORL)

(K) The Outcome-Representation Learning model (ORL) was introduced in 2018 by Haines and colleagues to accommodate the IGT. Unlike the PVL-Delta model, ORL does not rely on Prospect theory, because it does not use utility functions (Haines et al., 2018). In contrast to the PVL-delta, the ORL proposes that win frequency and the expected value of each deck should be modelled separately, and not implicitly in a utility function (Haines et al., 2018). Following is a break down of the different model components and parameters of the ORL.

The ORL includes separate learning rates for positive and negative outcomes of trial t, which are used to calculate the expected utility Ev_t^d of deck d on trial t as

$$Ev_t^d \leftarrow \begin{cases} Ev_{t-1}^d + a_{\text{rew}} \cdot \left(X_{t-1} - Ev_{t-1}^d\right), & \text{if } X_{t-1} \geq 0 \\ Ev_{t-1}^d + a_{\text{pun}} \cdot \left(X_{t-1} - Ev_{t-1}^d\right), & \text{otherwise} \end{cases} \tag{2.4}$$

where $a_{rew} \in (0, 1)$ is the reward learning parameter which is used when the outcome X_{t-1} of the previous trial is more than or equal to zero, and $a_{pun} \in (0, 1)$ is the punishment learning parameter which is used when the outcome X of the previous trial $t - 1$ is less than zero. The expected utility is based on the objective outcome of the previous trial, X_{t-1} (Haines et al., 2018). Having different learning rates for positive and negative outcomes, respectively, enables the model to accommodate different sensitivities to gains and losses, much like the loss aversion parameter of the PVL-Delta (Haines et al., 2018). The expected outcome frequency, Ef_t^d (win frequency) of deck d on trial t is modeled as

$$Ef_t^d \leftarrow \begin{cases} Ef_{t-1}^d + a_{\text{rew}} \cdot \left(\text{sgn}(X_{t-1}) - Ef_{t-1}^d\right), & \text{if } X_{t-1} \geq 0 \\ Ef_{t-1}^d + a_{\text{pun}} \cdot \left(\text{sgn}(X_{t-1}) - Ef_{t-1}^d\right), & \text{otherwise} \end{cases} \tag{2.5}$$

where the learning rates, a_{pun} and a_{rew}, are the same as before, and sgn returns either -1, 0, or 1 in accordance with the outcome X_{t-1} of the previous trial, being either negative, 0, or positive, respectively (Haines et al., 2018).

Reversal learning for Ef_t^d is included in ORL by modeling the expected outcome frequency of all unchosen decks d' on trial t as

$$Ef_t^{d'} \leftarrow \begin{cases} Ef_{t-1}^{d'} + a_{\text{pun}} \cdot \left(\frac{-\,\text{sgn}(X_{t-1})}{C} - Ef_{t-1}^{d'}\right), & \text{if } X_{t-1} \geq 0 \\ Ef_{t-1}^{d'} + a_{\text{rew}} \cdot \left(\frac{-\,\text{sgn}(X_{t-1})}{C} - Ef_{t-1}^{d'}\right), & \text{otherwise} \end{cases} \tag{2.6}$$

where $Ef_t^{d'}$ is the expected outcome frequency of the unchosen decks d', and C is the number of possible alternative choices for the chosen deck, which in this study always is three. The learning rates, a_{rew} and a_{pun} are the same as before. When the outcome X_{t-1} of the previous trial is positive, it is assumed that all other options (i.e. the unchosen deck d') would have yielded a negative outcome, hence the use of the negative learning rate, a_{pun}. Vice versa, when the outcome of the previous trial $t-1$ is negative, it is assumed that d' would have yielded a positive outcome, hence the use of the positive learning rate, a_{rew} (Haines et al., 2018).

To assess the tendency of participants to stay or switch between decks independent of outcome, choice perseverance is modeled as

$$Ps_t^d \leftarrow \begin{cases} \frac{1}{1+K}, & \text{if } x_t = d \\ \frac{Ps_t^d}{1+K}, & \text{otherwise} \end{cases} \tag{2.7}$$

where Ps_t^d is the perseverance of deck d on trial t, and x_t denotes the chosen deck on trial t. K is a decay parameter, which illustrates how quickly participants forget their recent history of deck choices. K is constrained to positive values and lower values of K indicate slower forgetting, while higher values of K indicate quicker forgetting. So, for the chosen deck x_t, the perseverance is reset (i.e., set to 1), and for the remaining unchosen decks, the perseverance exponentially decays (Haines et al., 2018). We used a simplified version of K because we did not find that Haines et al. (2018) sufficiently justified their reparameterization of K.

Finally, the different components of the ORL (expected value, frequency and perseverance) are integrated into a linear function to create a value score for each deck defined as

$$V_t^d \leftarrow Ev_t^d + Ef_t^d \cdot \omega_f + Ps_t^d \cdot \omega_p \tag{2.8}$$

where V_t^d is the value score of each deck d on trial t. The expected frequency (win frequency) and the perseverance parameters are weighted by $\omega_f \in (-\infty, \infty)$ and $\omega_p \in (-\infty, \infty)$, respectively, which determines how much the expected frequency and the perseverance affect V_t^d (Haines et al., 2018). So, values of ω_f less than zero indicate that participants prefer decks with low win frequency, while values of ω_f greater than zero indicate that participants prefer decks with high win frequency. Values of ω_p lower than zero indicate that participants prefer to switch decks more often, while values of ω_p greather than zero indicate that participants prefer to 'stay' by repeatedly choose the same deck, more often (Haines et al., 2018). Given that the expected value of deck d on trial t, Ev_t^d is not weighted, it functions as form of baseline (Haines et al., 2018). The ORL uses the same softmax function to generate choice probabilities for each deck as the PVL-Delta, except it takes its own expected values, Ev_t^d, as input. The softmax function can be seen in equation 2.3.

Thus, the ORL model has six parameters to be estimated; the learning rates of a_{rew} and a_{pun}, a decay parameter K, a switch parameter θ, the weight of expected win frequency, ω_f, and the weight of perseverance, ω_p.

2.2.3 Bayesian Analysis

(S) Bayesian analysis is used to estimate the parameters of PVL-Delta and ORL at the subject- and hierarchical-level. At subject-level, the parameters of the model are thought to be independent per participant, while in the hierarchical model, the parameters of each participant are thought to be sampled from a normal distribution. Hierarchical Bayesian analysis is performed for the validation methods parameter recovery, parameter estimation, convergence diagnostics and DIC. This is chosen to enhance comparability to similar studies and because hierarchical Bayesian analysis has previously been found to provide higher accuracy in parameter recovery than subject-level analysis (Ahn et al., 2011).

Parameter recovery and posterior predictive checks are performed on the subject-level, where the latter is only possible at this level. Subject-level parameter recovery is included to further test the robustness of models. Since all hierarchical model parameters are assumed to originate from normal distributions, a mean μ and a precision $\lambda = \frac{1}{\sigma^2}$ is reported for each parameter of PVL-Delta and ORL. The prior distributions for each of the hierarchical parameters can be found in Appendix B.

Bayesian analysis is performed using Markov chain Monte Carlo simulation (MCMC). MCMC is used to approximate posterior parameter distributions given a Bayesian model description and the IGT data. MCMC works by generating Markov chains such that its equilibrium state is the posterior distribution of the model parameters in either PVL-Delta or ORL. Generally, accuracy is improved with more steps in the chain. In the present analysis a combination of Metropolis sampling, Gibbs sampling as well as other MCMC sampling methods are used. MCMC is performed in Rstudio (RStudio Team, 2020) using JAGS (Su and Yajima, 2021) with 3000-5000 steps, where the first 1000 steps are warm-up (burn-in) and three chains.

2.3 Model comparison

(K) The two models, PVL-Delta and ORL, were compared using the following validation procedures: parameter recovery, posterior predictive checks, convergence diagnostics and deviance information criterion (DIC). Prior to validating the models, they were investigated as described in the following section.

2.3.1 Preliminary model investigation

(K) Different versions of both the PVL-Delta and the ORL were inspected using parameter recovery on both subject- and hierarchical-level. All parameter recoveries were done with both 50 and 100 iterations, where 100 iterations provided the best result.

To justify the choice of only including data sets from Steingroever et al. (2015) which employed 100 trials and thereby dismissing data sets which had 95 trials, having 95 or 100 trials was explored on the subject-level. Additionally, having more or less trials and subsequently having fewer or more participants was investigated in hierarchical parameter recovery. By visually inspecting the parameter recovery plots, it was found that there was no great difference, thus increasing our confidence in only including data which had employed 100 trials. Plots can be found in Figure 10, 11, 12 and 13 in Appendix C.

(S) Having chosen to simulate 100 trials and 232 subjects with 100 iterations, the ORL was tested with scaling (dividing net outcome of each deck after 100 trials by 100) and no scaling of outcomes. It was found that scaling was necessary to reliably recover the weighting parameters of ω_f and ω_p (especially ω_p cannot be recovered without scaling). Scaling the outcomes further helped recovering θ. Plots of the recovered μ of each scaled and non-scaled parameter can be found Figure 14 in Appendix C. Consequently, the outcomes were scaled in both PVL-Delta and the ORL in the following analysis.

Haines et al. (2018) could not recover θ why it was originally fixed to one. We tested this by recovering parameters of the model both including and excluding θ and found that θ could be recovered. We therefore included θ in the model. Plots of the recovered μ of the ORL parameters with and without θ can be found Figure 15 in Appendix C.

2.3.2 Parameter recovery

(S) Parameter recovery tests the robustness of the model. In short, parameter recovery works as follows; Data (choice of deck and net outcome) is simulated with a sampled set of parameter values and the specific payoff structure of the IGT (Haines et al., 2018). The model is then fit to the simulated data, and JAGS (Su and

Yajima, 2021) is then used to estimate the sampled parameters by running a MCMC simulation. JAGS outputs a posterior density for each parameter, and the maximum point of each distribution is found and compared to the sampled (the true) parameter (Haines et al., 2018). The process is then repeated 50-100 times, where after both the 'true' and the 'inferred' (the estimated) parameters are plotted. Additionally, the R^2 of each 'true' and 'inferred' parameter was computed. For simplicity, the simulated payoff structure varied from the traditional IGT structure in that the losses of deck A and C were fixed to -250$ (net sum of trial with loss was -150$) and -50$ (net sum of trial was 0$), respectively.

2.3.3 Posterior predictive checks

(K) Posterior predictive checks is a method for assessing a model's compatibility with the data (Wilson and Collins, 2019). This is done using MCMC to predict the model posterior and compare it to the IGT data. Posterior predictive checks are performed on subject-level with the following procedure:

1. The deck choices x and the payoff structures X of each subject are provided along with a Bayesian model description to obtain posterior parameter distributions for either ORL or PVL-Delta.
2. The maximum of the posterior distribution of the parameter p is acquired. This corresponds to the model's prediction of which deck a participant would choose on each trial.
3. The number of predicted deck choices that matches the actual deck choices of the participant is calculated and divided by the total number of trials. This value is calculated for all participants and compared to chance level performance of 0.25 (25% chance of choosing each deck).

2.3.4 Parameter estimation

(S) The models are fitted to the IGT data from Steingroever et al., 2015 and the parameters of the models are thereby estimated using MCMC on the hierarchical-level. For the PVL-Delta model we estimate posterior distributions for μ_w, μ_A, μ_θ, μ_A, λ_w, λ_A, λ_θ and λ_A. For the ORL model we estimate $\mu_{a_{rew}}$, $\mu_{a_{pun}}$, μ_K, μ_θ, μ_{ω_f}, μ_{ω_p}, $\lambda_{a_{rew}}$, $\lambda_{a_{pun}}$, λ_K, λ_θ, λ_{ω_f} and λ_{ω_p}.

2.3.5 Convergence diagnostics

(S) When performing MCMC it is important to ensure converge of the Markov chains. Otherwise, the parameter estimates obtained from the posterior distributions might be unreliable (Speagle, 2019). A method that is used to investigate how well the Markov chains converged is to visually inspect the Markov chains to see how well they explore the state space and how the chains mix (Speagle, 2019). Ideally, the chains of MCMC simulation should look like hairy caterpillars by being well mixed with approximately similar variance and no obvious drifts (Roy, 2020). Furthermore, $\hat{R}$ is used to quantitatively asses the convergence by comparing the estimated between-chains and within-chains variances for each parameter (Gelman and Rubin, 1992). The implementation we use is, however, the improved rank-based $\hat{R}$ (Vehtari et al., 2020). Brooks and Gelman furthermore suggest that $\hat{R}$-values greater that 1.2 should indicate non-converge. Although, often the more stringent rule of 1.1 is applied and thus we will adhere to that convention.

2.3.6 Deviance Information Criterion (DIC)

(K) To further validate and compare the fits of PVL-Delta and ORL, the deviance information criterion (DIC) of each model is computed. DIC is an information criteria developed specifically to compare Bayesian models. The DIC is based on the principle that model fit can be viewed as a sum of the goodness of fit penalized by the models complexity (Spiegelhalter et al., 2002). The fit of the model is measured using deviance defined as

$$D(\theta) = -2 \log \mathcal{L}(y \mid \theta), \tag{2.9}$$

where y is the data, $\mathcal{L}$ is the likelihood function and θ here is the model parameters. The complexity is then defined as

$$p_D = E_{\theta|y}[D] - D\left(E_{\theta|y}[\theta]\right), \tag{2.10}$$

where $E_{\theta|y}$ is the expected posterior. Thus, the DIC can finally be defined analogously to the AIC as

$$\text{DIC} = D(\hat{\theta}) + 2p_D \tag{2.11}$$

where $\hat{\theta}$ are the bayes estimated parameters. This means that models with smaller DIC is better supported by the data. Furthermore, DIC has been found to be meaningful when comparing non-nested models such as PVL-Delta and ORL (Spiegelhalter et al., 2002). However, the rule of thumb of $\Delta\text{AIC} < 2$ for a substantial difference betweeen the models compared might not be directly applicable for DIC (Burnham and Anderson, 2004). Thus, one should be cautious in interpreting the magnitude of the difference in DIC.

2.4 System specification

(S) The analysis was performed using R (R Core Team, 2022) in RStudio (RStudio Team, 2020) employing the following packages; 'R2jags' (Su and Yajima, 2021) for JAGS sampling, 'parallel' for running chains in parallel, and 'ggpubr' (Kassambara, 2022) for plotting. Additionally, the packages 'extraDisr' (Wolodzko, 2020), 'truncnorm' (Mersmann et al., 2018), 'hesim' (Incerti and Jansen, 2021) and 'tidyverse' (Wickham et al., 2019) were used. The data handling and analysis was run on the uCloud interactive HPC system, which is administered by the eScience Center at the University of Southern Denmark. Code is available at Github. Link can be found in Appendix G.

3 Results

(S) The results from fitting and validating the PVL-Delta model and the ORL model on the IGT dataset can be found in the following sections.

3.1 PVL-Delta

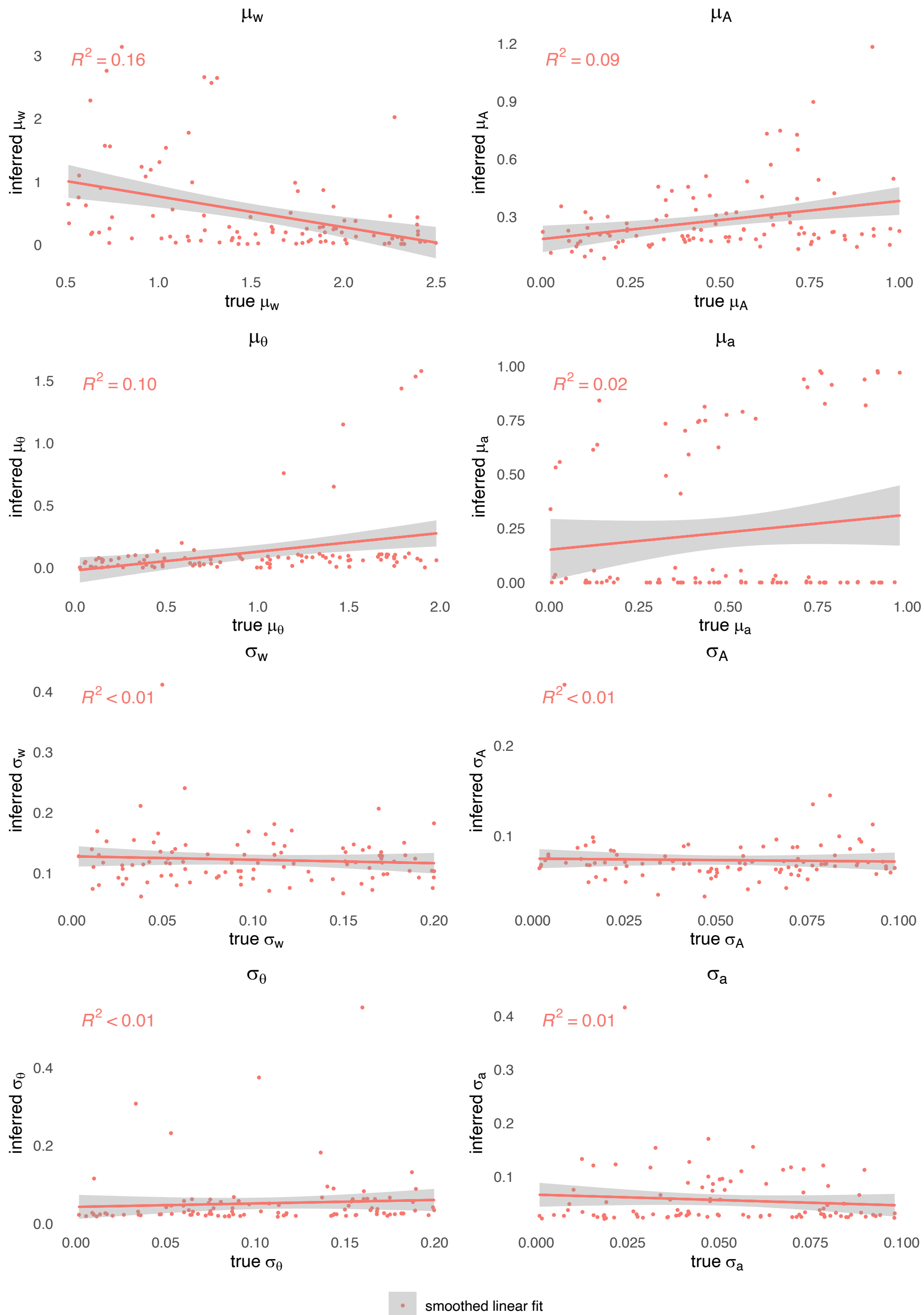
3.1.1 Parameter recovery

(S) On the subject-level, the parameters were not well-recovered. The recovered values of w were in general overestimated with great variance, while the recovered values of a are widely spread. The parameters A and θ were especially poorly recovered. The recovery of θ showed noteworthy underestimation of inferred values including a tendency to sample the lower true values around zero (Figure 16 in Appendix D).

The hierarchical parameter recovery was seemingly worse than the subject-level parameter recovery. The μ parameters of the hierarchical model were poorly recovered with a general tendency to sample the inferred values around zero, and the λ parameters (recoded as the standard deviation, σ) of the hierarchical model were not recovered (Figure 3).

Figure 3

PVL-Delta parameter recovery of μ and σ (hierarchical)



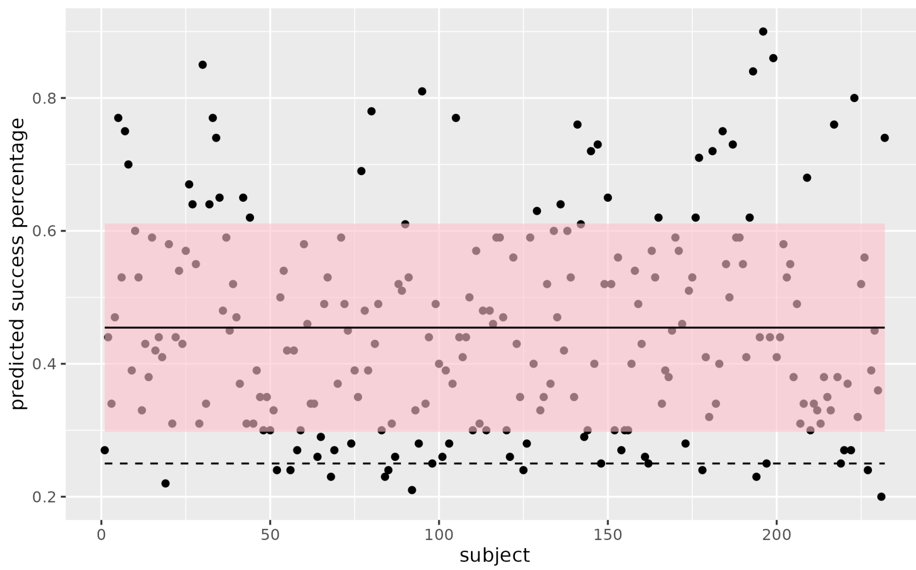
Note. Parameter recovery at hierarchical-level with linear model fit and R^2-values. Precision λ is recoded as standard deviation σ.

3.1.2 Posterior predictive checks

(K) Figure 4 shows that the PVL-Delta model on average predicts the behavior of the participants better than chance level. This shows that the predicted success percentage is about 0.45 compared to chance level at 0.25. The shaded red area further marks that the predictions within one standard deviation of the mean is still above chance level.

Figure 4

Posterior predictive checks for subject-level PVL-Delta



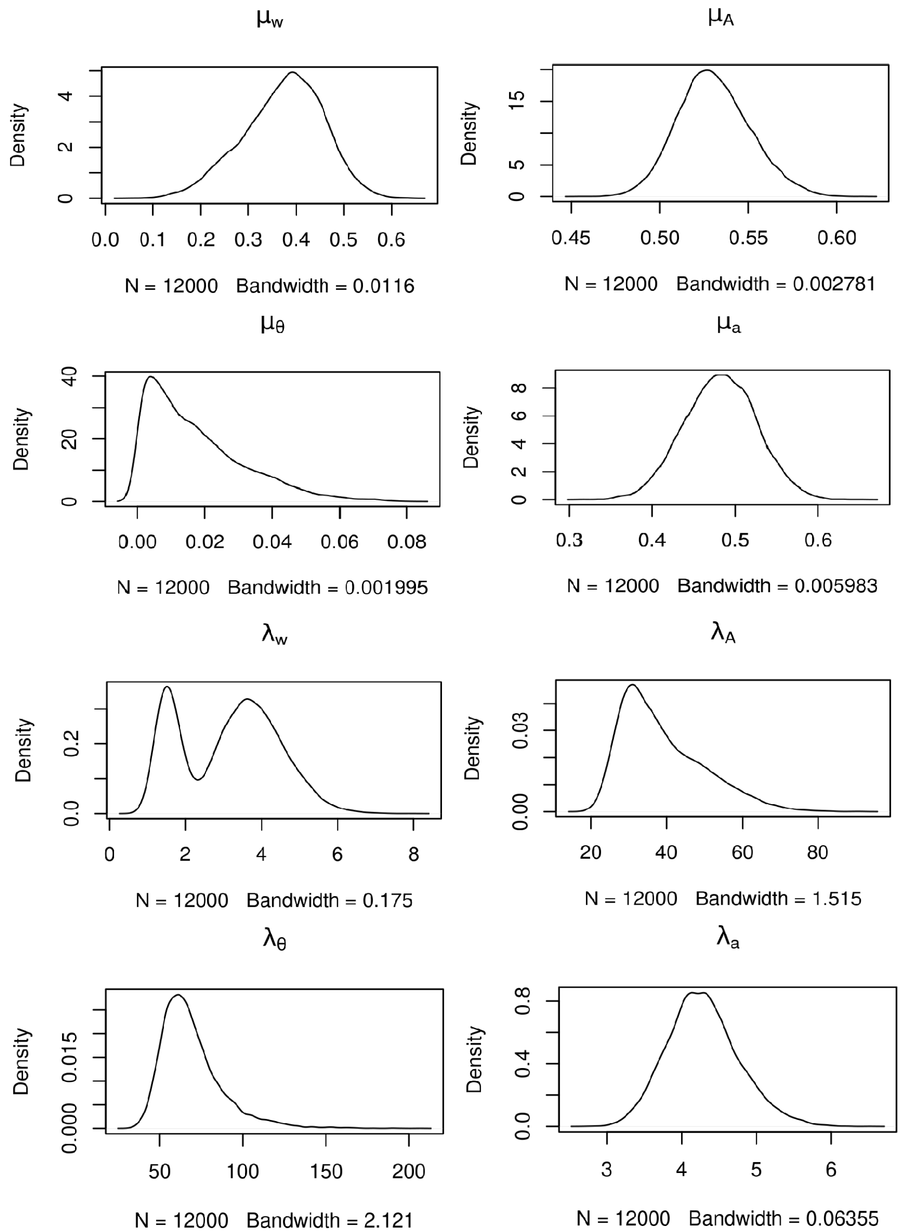
Note. Each point in the plot marks a predicted success percentage per subject (i.e., the number of times the PVL-Delta model correctly predicts the behaviour of a participant). The solid black line is the average over all participants and the shaded red area covers the prediction within one standard deviation of the mean. The dashed line marks chance level.

3.1.3 Parameter estimation

(S) The posterior distributions of the parameters of the hierarchical PVL-Delta can be viewed in Figure 5a. Additionally, the maximum of the posterior distributions as well as their 95% credible interval are reported in Figure 5b.

Figure 5

a) Parameter estimations for hierarchical PVL-Delta parameters.



b) Maximum a posteriori

Parameter	λ_A	λ_a	λ_θ	λ_w	μ_A	μ_a	μ_θ	μ_w
95% CI upper	64.5770	5.3254	117.2940	5.4479	0.5735	0.5661	$5.3854 \cdot 10^{-2}$	0.5216
MAP	30.9747	4.1369	61.1268	1.5198	0.5268	0.4804	$3.9325 \cdot 10^{-3}$	0.3925
95% CI lower	24.1226	3.4130	44.3441	1.1617	0.4932	0.3970	$5.5698 \cdot 10^{-4}$	0.1944

Note. Parameter estimations of PVL-Delta. Posterior distributions of parameters (a). Maximum a posteriori values and 95% credible intervals (CI) (b).

3.1.4 Convergence diagnostics and DIC

(K) Plots of the three Markov chains from the MCMC simulation can be seen in Figure 18 in Appendix E. Here it is visually inspected that only the chains for estimating μ_A and λ_A appear like hairy caterpillars. $\hat{R}$ of each hierarchical PVL-Delta parameter can be found in Table 2 in Appendix F. It is evident that only the Deviance and μ_θ obtain $\hat{R}$-values below 1.1. Additionally, PVL-Delta yielded a DIC of 59222.85.

3.2 ORL

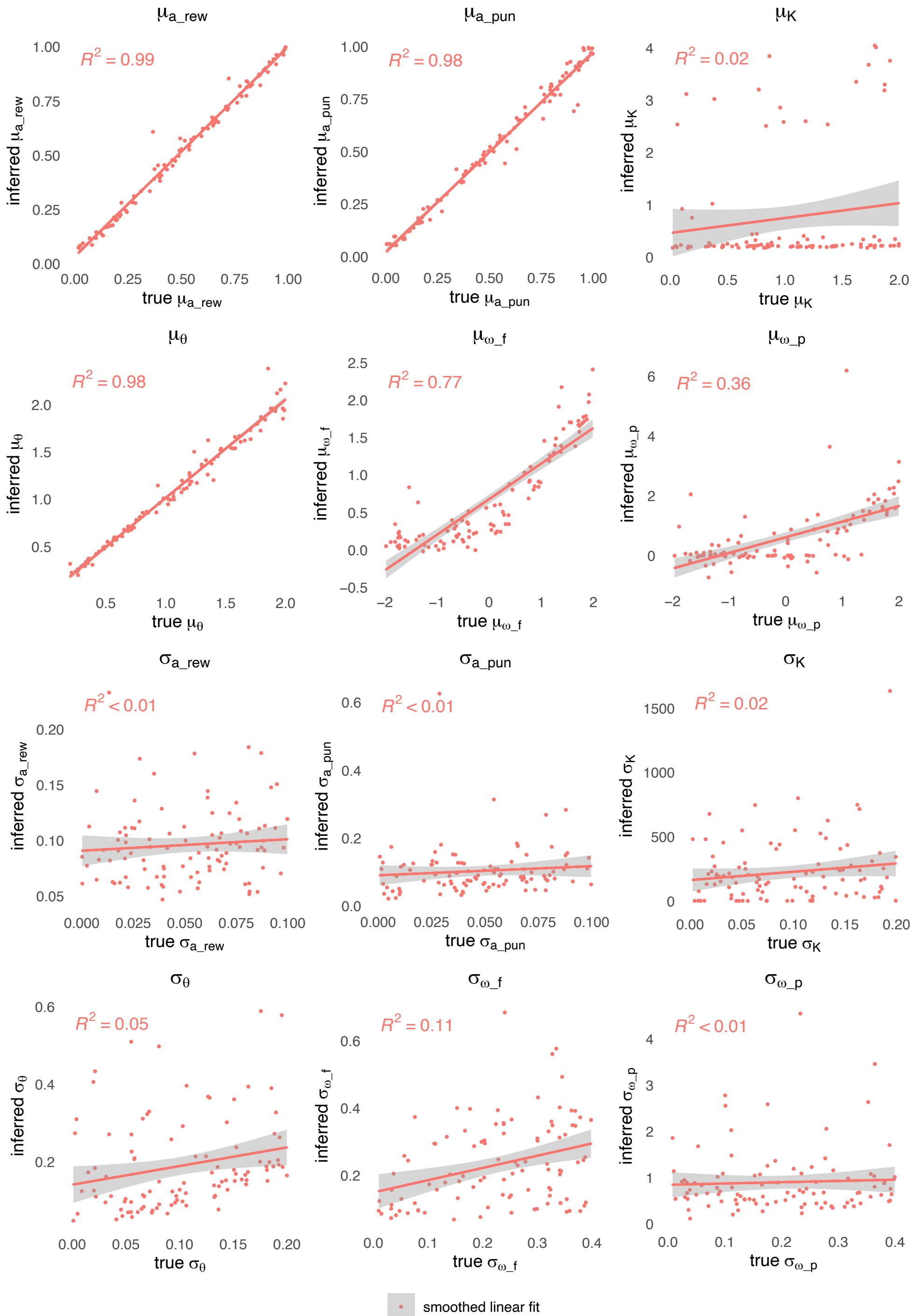
3.2.1 Parameter recovery

(K) On the subject-level, all but one of the parameters were recovered relatively well, despite a tendency to underestimate the inferred values of θ and ω_p, especially. The one parameter which could not be recovered was K, which showed no linear relationship between the true and the inferred values (Figure 17 in Appendix D).

The hierarchical parameter recovery of μ was in general found to be better for most parameters ($\mu_{a_{rew}}$, $\mu_{a_{pun}}$, and μ_θ, especially), while other of the μ parameters showed a tendency to sample the inferred values around zero (μ_K, μ_{ω_p}). The parameter of μ_{ω_f} seemed to be relatively well recovered, but values less than zero were, however, not recovered. The λ parameters (recoded as the standard deviation, σ) were not recovered (Figure 6).

Figure 6

ORL parameter recovery of μ and σ (hierarchical)



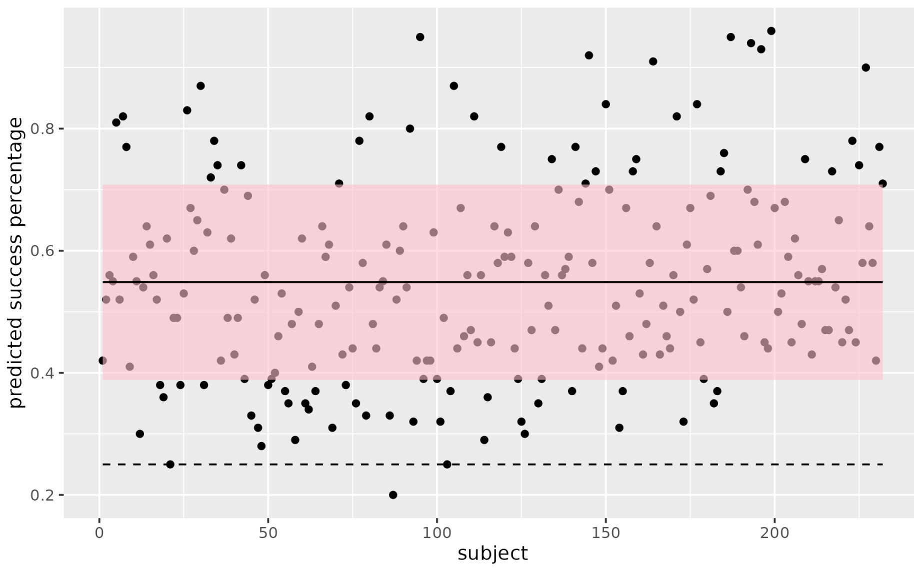
Note. Parameter recovery for each of the parameters on the hierarchical-level with a linear model fit and R^2-values. Precision λ is recoded as standard deviation σ.

3.2.2 Posterior predictive checks

(S) The posterior predictive checks of the ORL model as seen in Figure 7 shows that the ORL model on average predicts the behavior of the participants at 0.55, which is better than chance level. This is also the case for the predictions within one standard deviation of the mean.

Figure 7

Posterior predictive checks ORL



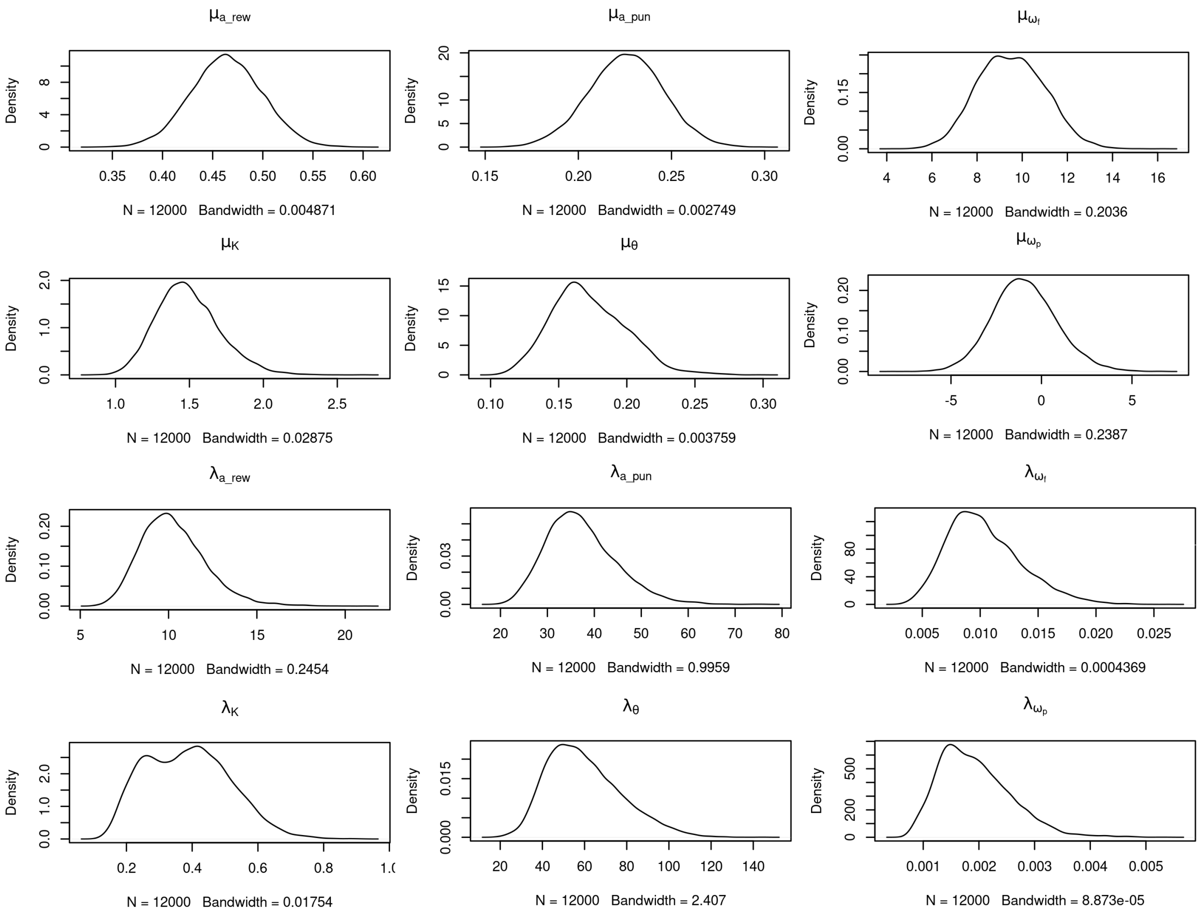
Note. Each point marks a predicted success percentage per subject (i.e. the number of times the PVL-Delta model correctly predicts the behavior of a participant). The solid black line is the average over all participants and the dashed line is chance level. The shaded red area covers the predictions within one standard deviation of the mean.

3.2.3 Parameter estimation

(K) The posterior distributions of parameter estimates from hierarchical Bayesian analysis of ORL is seen in Figure 8a. The maximum values of the posterior distributions as well as the 95% credible intervals are seen in Figure 8b.

Figure 8

a) Parameter estimation of hierarchical ORL



b) Maximum a posteriori

Parameter	95% CI lower bound	MAP	95% CI upper bound
λ_K	0.1783	0.4162	0.6455
$\lambda_{a_{pun}}$	24.9772	34.8273	53.9292
$\lambda_{a_{rew}}$	7.2712	9.8622	14.4697
λ_{ω_f}	$5.2356 \cdot 10^{-3}$	$8.6532 \cdot 10^{-3}$	$1.7692 \cdot 10^{-3}$
λ_{ω_p}	9.4898 ·10−4	$14.926 \cdot 10^{-4}$	$33.724 \cdot 10^{-4}$
λ_θ	32.6651	49.5338	99.5956
μ_K	1.1214	1.4523	1.9567
$\mu_{a_{pun}}$	0.1857	0.2249	0.26445
$\mu_{a_{rew}}$	0.3946	0.4627	0.5334
μ_{ω_f}	6.7697	8.9207	12.4206
μ_{ω_p}	-4.2575	-1.2545	2.7574
μ_θ	0.1248	0.1614	0.2296

Note. The MAP column contains the maximum of the posterior distribution. Additionally, the 95% credible intervals (CI) are presented.

3.2.4 Convergence diagnostics and DIC

(S) The markov chains from MCMC are seen in Figure 19 in Appendix E. By visually inspecting the chains, it seems that only the chains for estimating μ_{ω_p} appear as hairy caterpillars. The quantitative measure of convergence $\hat{R}$ can be found in Table 3 in Appendix F. Here, only the $\hat{R}$-value of $\lambda_{a_{rew}}$ is below 1.1. Lastly, DIC yielded a value of DIC = 51414.1 for ORL.

4 Discussion

(S) The results provide the overall indication that the ORL model describes the behavior in our IGT data set better. This is evident by the ORL outperforming the PVL-Delta in all the investigated model validation metrics. The ORL model obtains a superior parameter recovery at the hierarchical level, higher predicted success percentage in posterior predictive checks and lower DIC. However, it should be noted that major convergence problems are present for both PVL-Delta and ORL. Hence, the results and the comparison to Haines et al. (2018) and Steingroever et al. (2013b) in the following sections should be interpreted with care.

4.1 Comparing with the literature

4.1.1 PVL-Delta

(K) Both Haines et al. (2018) and Steingroever et al. (2013b) validate the PVL-Delta model. However, it is only Haines et al. who perform parameter recovery using a similar procedure to ours. Haines et al. report good parameter recovery across all model parameters for the PVL-Delta. In comparison, we do not recover the PVL-Delta parameters as well. Especially the precision values in the hierarchical PVL-Delta were challenging to recover, which could indicate that more data is needed to recover precision parameters. However, running parameter recovery on 232 subjects on 100 trials took up to 12 days and so, it seems infeasible to increase the number of simulated subjects with the current implementation. One drawback of our parameter recovery compared to the parameter recovery from Haines et al. is that our parameter spaces for A and w are constrained to narrower intervals, namely $A \in (0, 1)$ and $w \in (0.5, 2.5)$. This means that the full parameter space of $A \in (0, 2)$ and $w \in (0, 10)$, respectively, are not investigated.

Steingroever et al. and Haines et al. did not perform posterior predictive checks. Haines et al. did, however, compute mean squared deviations (MSD) which is a simulation method to compare long-term prediction accuracy across models similarly to posterior predictive checks. What separates MSD from posterior predictive checks is that it does not condition on the observed IGT data in the same manner as posterior predictive checks. Using MSD, Haines et al. found that PVL-Delta obtained the lowest MSD on 5 out of the 9 investigated data sets. In comparison our posterior predictive checks yield a predicted success percentage above chance level. Thus, it seems that PVL-Delta achieves good posterior predictive checks both in Haines et al. and in our study.

Steingroever et al. report hierarchical mean parameter estimations for the PVL-Delta parameters. When comparing the posterior distributions of subjects performing 150 trials IGT to our posterior distributions, it seems that comparable distributions are achieved for some parameters (A and w). Although, this is hard to determine because of the scaling of results reported by Steingroever et al. One possible explanation for our parameter estimation results differing from those of Steingroever et al. and Haines et al. is the fact that our Markov chains have not converged. Steingroever et al. do outline some initial convergence problems with PVL-Delta, which they solved by re-initializing all chains with values close to the mean of their initial unconverged Markov chains. This mostly solved their convergence problems, however, they did exclude some participants to further enhance convergence. Steingroever et al. also used longer Markov chains of 9000 steps (1000 being burn-in). Thus, potentially more reliable parameter estimates could have been obtained if convergence of the Markov chains was ensured by providing better initialization and more steps.

4.1.2 ORL

(S) The ORL model was introduced by Haines et al. (2018) and to our knowledge it has yet to be validated by others. In their parameter recovery of the ORL, Haines et al. find that all model parameters are well-recovered. In our parameter recovery, subject-level parameters are relatively well-recovered with the exception of K. In hierarchical ORL, some parameters ($\mu_{a_{rew}}$, $\mu_{a_{pun}}$ and μ_θ) appear very well-recovered, whilst the recovery of μ_K and μ_{ω_p} appear less successful. The parameter recoveries for the precision values were again challenging to recover. Thus, we were not able to create a parameter recovery comparable to Haines et al. It should be noted that we do recover μ_θ which Haines et al. chose to fixate at 1 because they could not recover the parameter. This fixation of θ could explain some of the differences between the parameter recovery in Haines et al. and the one we present.

Haines et al. also compute the MSD for the ORL. They found that the ORL model obtained the lowest MSD score in 4 out of the 9 data sets, which is approximate to MSD performance of PVL-Delta. In our validation of posterior predictive checks, the predicted success percentage of ORL was higher than PVL-Delta. This suggests that the ORL model fits the behavior in our data set relatively better than the PVL-Delta.

In the ORL parameter estimations reported by Haines et al., parameters of different groups of controls and various subjects with SUDs are included. Thus, we compare our posterior distributions of the ORL parameters to those of the controls in Haines et al. It appears that our parameter estimates are generally higher than those of Haines et al.

A possible explanation for our parameter estimates differing from those of Haines et al. is the problems with the Markov chain convergence. In contrast, Haines et al. report successful convergence of all parameters, despite them using Markov chains of similar length to ours with 4000 steps (1500 being warm-up).

4.2 Limitations of the models

(K) Cognitive models are inherently limited as they seek to quantify latent cognitive processes which are theoretical and abstract. A cognitive process must therefore be simplified in attempts to model it. Consequently, cognitive models such as the PVL-Delta and the ORL may never be able to reveal the full extent of a complex cognitive process such as decision-making. Using extensive model validation techniques we can, however, come close.

Steingroever et al. (2013b) investigated the limitations of the PVL-Delta model using use two validation techniques; parameter space partitioning and test of selective influence. They concluded that PVL-Delta rarely generates choice patterns with preferences for 'bad' over 'good' decks. This limitation might not be a problem when modeling subjects who rarely exhibit this behavioral pattern. However, when modeling data from participants with decision-making deficits, it could be an issue as these individuals often show this preference (Bechara et al., 1994; Bechara et al., 1997). In the test of selective influence, Steingroever et al. manipulate the IGT such that it should result in a change in the different model parameters. They successfully influenced the loss aversion parameter w and the choice consistency parameter θ but they were unable to affect the learning parameter a during these manipulations. This indicates that a might not model the learning process as assumed. Steingroever et al., however, note that this should be further investigated.

(S) Naturally, ORL has its own limitations. It seemingly captures many of the cognitive strategies observed in behavioral IGT data. However, a drawback of ORL is that it is sensitive to scaling of the of net outcomes as found during preliminary parameter recovery (see section 2.3.1). Furthermore, we encountered major problems in recovering the decay parameter K. Haines et al. did not have the same issue, but they reparameterized it which we did not. Given that we could not recover K, we should have performed additional parameter recoveries with the reparameterized K. However, this was not possible due to the aforementioned run time of the models.

Unlike, Haines et al., we could recover θ. Interestingly, in the parameter recovery without θ, ω_p was notably better recovered than in the recovery of the model including θ. This might suggest that some of the same variance is assigned to those two parameters. As aforementioned, the θ which Haines et al. could not recover was reparameterized. They might have decided to reparameterize θ in ORL, because it was originally reparameterized in PVL-Delta. Thus, we might have succeeded in recovering θ because we did not reparameterize it. It should have been tested, but again due to the running time of the models, this was not done. It thereby seems that the ORL parameters of K and θ might be slightly dubious.

It should be noted that any comparison of different models such as the ORL and PVL-Delta should be done with great care. That is, it is much simpler to compare nested models which share many of the same parameters. When comparing non-nested models, one should investigate many different model validation metrics besides looking at the performance of said models on the same data. A drawback of comparing non-nested models is that it might be difficult to disentangle the performance of the model from the contributions of specific parameters. That is, the finding that the ORL outperforms the PVL-Delta needs to be followed by a thorough investigation of ORL, much like Steingroever et al.'s rigorous testing of the PVL-Delta. Thereby a better understanding of the relative success of ORL to model decision-making in the IGT might be achieved.

4.3 Further perspectives

(S) The presented results add an interesting perspective to the debate about decision from experience vs. decision from description. According to Hertwig et al. (2004)'s arguments, the results of our study indicate that there is a need for two different models of decision-making. Also, the sampling error which Fox and Hadar (2006) claimed Hertwig et al. accidentally based their findings on is not present in the IGT, because the 'rare event', which in the IGT is a great infrequent loss, is present somewhere within every block of 10 draws from deck B and D. Thus, every participant drawing from deck B would within ten draws meet this infrequent loss, and thereby potentially learn to avoid said deck. To test this, the models should be tested on a descriptive decision-making task. Then, if there truly is a need for two different models, PVL-Delta should outperform ORL.

However, the fact that ORL performed relatively better on the IGT data than PVL-Delta, does not necessarily support the argument that prospect theory is unnecessary. It could be that ORL would improve if it was based on prospect theory. In the same vein, it could be that prospect theory is a good way of modeling both types of decision-making data, but it might be overshadowed by other parameters of the PVL-Delta which might be less relevant.

No matter the verdict, one should be cautious in interpreting the models solely by reference to whether they include prospect theory or not. That is, the parameters of the respective models differ greatly besides whether they calculate the expected value of a deck based on the objective outcome or on a subjective utility function. For example, ORL includes different learning rates for wins and losses and reversal learning, while PVL-Delta models risk aversion through the shape parameter. Thus, the models incorporate different parameters and can thereby be used to cast light on different aspects of decision-making.

4.4 Future research

(K) Immediate next steps include running parameter recoveries for the full parameter space, and explore ways to improve convergence of Markov chains (e.g., increase length of chains). Additionally, the different parameters of the ORL should be thoroughly investigated by, e.g., inspecting the simulated choice patterns.

To investigate decision from description and the extent to which it differs from decision from experience, the IGT could be altered to facilitate decision from description instead of experience. One could imagine that instead of turning cards to learn which decks are 'good' versus 'bad', the probabilities of the given gains and losses of each deck are explicitly stated. If a separate model of decision from experience is needed, the PVL-Delta would outperform the ORL when modeling the descriptive IGT. Introducing descriptive IGT might be an interesting way to investigate the question of whether two different theories, or models, are needed to explain decision-making.

Another way to test if a division of decision-making is necessary is to alter the ORL so it includes prospect theory. The performance of the altered ORL could then be tested on both the original and the hypothetical descriptive IGT, which could be compared to the performance of the original ORL. Vice versa, one could alter the PVL-Delta so it is not based on prospect theory. Investigating the difference of calculating the expected value of decks based on a subjective utility function or on the objective monetary outcome would remedy the discussion of whether prospect theory is necessary to model decision-making. Understanding how prospect theory affects model performance, would help the interpretation of the remaining parameters.

5 Conclusion

We have investigated the relationship between decisions from description and decisions from experience using two different cognitive models. We examined the PVL-Delta model, which includes prospect theory, and the ORL model, which does not include prospect theory, and their respective performance of modeling Iowa Gambling Task data. The results indicate that ORL seems promising in modeling experience-based decision-making data despite convergence problems. It is, however, premature to conclude whether or not prospect theory is needed for modeling experience-based decision-making, and future studies should, among other things, test the models on a descriptive based decision-making task.

References

Ahn, Woo-Young, Busemeyer, Jerome R., Wagenmakers, Eric-Jan, and Stout, Julie C. (2008). "Comparison of Decision Learning Models Using the Generalization Criterion Method". In: *Cognitive Science* 32.8, pp. 1376–1402. DOI: https://doi.org/10.1080/03640210802352992. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1080/03640210802352992. URL: https://onlinelibrary.wiley.com/doi/abs/10.1080/03640210802352992.

Ahn, Woo-Young, Krawitz, Adam, Kim, Woojae, Busemeyer, Jerome, and Brown, Joshua (May 2011). "A Model-Based fMRI Analysis With Hierarchical Bayesian Parameter Estimation". In: *Journal of Neuroscience, Psychology, and Economics* 4, pp. 95–110. DOI: 10.1037/a0020684.

Ahn, Woo-Young, Vasilev, Georgi, Lee, Sung-Ha, Busemeyer, Jerome R., Kruschke, John K., Bechara, Antoine, and Vassileva, Jasmin (2014). "Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users". In: *Frontiers in psychology* 5, p. 849.

Barberis, Nicholas C (2013). "Thirty years of prospect theory in economics: A review and assessment". In: *Journal of Economic Perspectives* 27.1, pp. 173–96.

Barron, Greg and Erev, Ido (July 2003). "Small Feedback-based Decisions and Their Limited Correspondence to Description-based Decisions". In: *Journal of Behavioral Decision Making - J BEHAV DECIS MAKING* 16, pp. 215–233. DOI: 10.1002/bdm.443.

Bechara, Antoine, Damasio, Antonio R, Damasio, Hanna, and Anderson, Steven W (1994). "Insensitivity to future consequences following damage to human prefrontal cortex". In: *Cognition* 50.1-3, pp. 7–15.

Bechara, Antoine, Damasio, Hanna, Tranel, Daniel, and Damasio, Antonio R. (1997). "Deciding Advantageously Before Knowing the Advantageous Strategy". In: *Science* 275.5304, pp. 1293–1295. DOI: 10.1126/science.275.5304.1293. eprint: https://www.science.org/doi/pdf/10.1126/science.275.5304.1293. URL: https://www.science.org/doi/abs/10.1126/science.275.5304.1293.

Bechara, Antoine, Dolan, Sara, Denburg, Natalie, Hindes, Andrea, Anderson, Steven W, and Nathan, Peter E (2001). "Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers". In: *Neuropsychologia* 39.4, pp. 376–389. ISSN: 0028-3932. DOI: https://doi.org/10.1016/S0028-3932(00)00136-6. URL: https://www.sciencedirect.com/science/article/pii/S0028393200001366.

Beitz, Kevin M, Salthouse, Timothy, and Davis, Hasker P. (2014). "Performance on the Iowa Gambling Task: From 5 to 89 years of age." In: *Journal of experimental psychology. General* 143 4, pp. 1677–1689.

Burnham, Kenneth P. and Anderson, David R. (2004). "Multimodel Inference: Understanding AIC and BIC in Model Selection". In: *Sociological Methods & Research* 33.2, pp. 261–304. DOI: 10.1177/0049124104268644. eprint: https://doi.org/10.1177/0049124104268644. URL: https://doi.org/10.1177/0049124104268644.

Erev, Ido (May 1998). "Signal Detection by Human Observers: A Cutoff Reinforcement Learning Model of Categorization Decisions under Uncertainty". In: *Psychological review* 105, pp. 280–98. DOI: 10.1037//0033-295X.105.2.280.

Fox, Craig and Hadar, Liat (Nov. 2006). ""Decisions from experience" = sampling error + prospect theory". In: *Judgment and Decision Making* 1.

Fridberg, Daniel, Queller, Sarah, Ahn, Woo-Young, Kim, Woojae, Bishara, Anthony, Busemeyer, Jerome, Porrino, Linda, and Stout, Julie (Feb. 2010). "Cognitive Mechanisms Underlying Risky Decision-Making in Chronic Cannabis Users". In: *Journal of mathematical psychology* 54, pp. 28–38. DOI: 10.1016/j.jmp.2009.10.002.

Gelman, Andrew and Rubin, Donald B. (1992). "Inference from Iterative Simulation Using Multiple Sequences". In: *Statistical Science* 7.4, pp. 457–472. ISSN: 08834237. URL: http://www.jstor.org/stable/2246093 (visited on 01/04/2023).

Haines, Nathaniel, Vassileva, Jasmin, and Ahn, Woo-Young (2018). "The outcome-representation learning model: A novel reinforcement learning model of the iowa gambling task". In: *Cognitive Science* 42.8, pp. 2534–2561.

Hertwig, Ralph, Barron, Greg, Weber, Elke, and Erev, Ido (Sept. 2004). "Decisions From Experience and the Effect of Rare Events in Risky Choice". In: *Psychological science* 15, pp. 534–9. DOI: 10.1111/j.0956-7976.2004.00715.x.

Incerti, Devin and Jansen, Jeroen P. (2021). *hesim: Health Economic Simulation Modeling and Decision Analysis*. arXiv: 2102.09437 [stat.AP]. URL: https://arxiv.org/abs/2102.09437.

Kahneman, Daniel and Tversky, Amos (1979). "Prospect Theory: An Analysis of Decision under Risk". In: *Econometrica* 47.2, pp. 263–291.

Kassambara, Alboukadel (2022). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.5.0. URL: https://rpkgs.datanovia.com/ggpubr/.

Kjome, Kimberly L, Lane, Scott D, Schmitz, Joy M, Green, Charles, Ma, Liangsuo, Prasla, Irshad, Swann, Alan C, and Moeller, F Gerard (2010). "Relationship between impulsivity and decision making in cocaine dependence". In: *Psychiatry research* 178.2, pp. 299–304.

Mersmann, Olaf, Trautmann, Heike, Steuer, Detlef, and Bornkamp, Björn (2018). *truncnorm: Truncated Normal Distribution*. R package version 1.0-8. URL: https://github.com/olafmersmann/truncnorm.

Premkumar, Preethi, Fannon, Dominic, Kuipers, Elizabeth, Simmons, Andrew, Frangou, Sophia, and Kumari, Veena (2008). "Emotional decision-making and its dissociable components in schizophrenia and schizoaffective disorder: a behavioural and MRI investigation". In: *Neuropsychologia* 46.7, pp. 2002–2012.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Rescorla, R. and Wagner, Allan (Jan. 1972). "A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement". In: *Classical Conditioning: Current Research and Theory*.

Roy, Vivekananda (2020). "Convergence Diagnostics for Markov Chain Monte Carlo". In: *Annual Review of Statistics and Its Application* 7.1, pp. 387–412. DOI: 10.1146/annurev-statistics-031219-041300. eprint: https://doi.org/10.1146/annurev-statistics-031219-041300. URL: https://doi.org/10.1146/annurev-statistics-031219-041300.

RStudio Team (2020). *RStudio: Integrated Development Environment for R*. Version 4.2.0. RStudio, PBC. Boston, MA. URL: http://www.rstudio.com/.

Salmond, C.H., Menon, D.K., Chatfield, D.A., Pickard, J.D., and Sahakian, B.J. (2005). "Deficits in Decision-Making in Head Injury Survivors". In: *Journal of Neurotrauma* 22.6. PMID: 15941371, pp. 613–622. DOI: 10.1089/neu.2005.22.613. eprint: https://doi.org/10.1089/neu.2005.22.613. URL: https://doi.org/10.1089/neu.2005.22.613.

Speagle, Joshua S. (2019). *A Conceptual Introduction to Markov Chain Monte Carlo Methods*. DOI: 10.48550/ARXIV.1909.12313. URL: https://arxiv.org/abs/1909.12313.

Spiegelhalter, D.J., Best, Nicky, Carlin, B.P., and Linde, A. (Jan. 2002). "Bayesian measures of model complexity and fit (with discussion)". In: *Journal of the Royal Statistical Society, Series B* 64, pp. 1–34. DOI: 10.1111/1467-9868.02022.

Steingroever, Helen, Fridberg, Daniel J, Horstmann, Annette, Kjome, Kimberly L, Kumari, Veena, Lane, Scott D, Maia, Tiago V, McClelland, James L, Pachur, Thorsten, Premkumar, Preethi, et al. (2015). "Data from 617 healthy participants performing the Iowa gambling task: A" many labs" collaboration". In: *Journal of Open Psychology Data* 3.1, pp. 340–353.

Steingroever, Helen, Wetzels, Ruud, Horstmann, Annette, Neumann, Jane, and Wagenmakers, Eric-Jan (2013a). "Performance of healthy participants on the Iowa Gambling Task." In: *Psychological assessment* 25.1, p. 180.

Steingroever, Helen, Wetzels, Ruud, and Wagenmakers, Eric-Jan (2013b). "Validating the PVL-Delta model for the Iowa gambling task". In: *Frontiers in psychology* 4, p. 898.

Su, Yu-Sung and Yajima, Masanao (2021). *R2jags: Using R to Run 'JAGS'*. R package version 0.7-1.

Tversky, Amos and Kahneman, Daniel (1992). "Advances in prospect theory: Cumulative representation of uncertainty". In: *Journal of Risk and uncertainty* 5.4, pp. 297–323.

Vehtari, Aki, Gelman, Andrew, Simpson, Daniel, Carpenter, Bob, and Bürkner, Paul-Christian (July 2020). "Rank-Normalization, Folding, and Localization: An Improved $\widehat{R}$ for Assessing Convergence of MCMC". In: *Bayesian Analysis* -1. DOI: 10.1214/20-BA1221.

Weber, Elke, Shafir, Sharoni, and Blais, Ann-Renee (May 2004). "Predicting Risk Sensitivity in Humans and Lower Animals: Risk as Variance or Coefficient of Variation". In: *Psychological review* 111, pp. 430–45. DOI: 10.1037/0033-295X.111.2.430.

Wickham, Hadley, Averick, Mara, Bryan, Jennifer, Chang, Winston, McGowan, Lucy D'Agostino, François, Romain, Grolemund, Garrett, Hayes, Alex, Henry, Lionel, Hester, Jim, Kuhn, Max, Pedersen, Thomas Lin, Miller, Evan, Bache, Stephan Milton, Müller, Kirill, Ooms, Jeroen, Robinson, David, Seidel, Dana Paige, Spinu, Vitalie, Takahashi, Kohske, Vaughan, Davis, Wilke, Claus, Woo, Kara, and Yutani, Hiroaki (2019). "Welcome to the tidyverse". In: *Journal of Open Source Software* 4.43, p. 1686. DOI: 10.21105/joss.01686.

Wilson, Robert C and Collins, Anne GE (Nov. 2019). "Ten simple rules for the computational modeling of behavioral data". In: *eLife* 8. Ed. by Timothy E Behrens, e49547. ISSN: 2050-084X. DOI: 10.7554/eLife.49547. URL: https://doi.org/10.7554/eLife.49547.

Wolodzko, Tymoteusz (2020). *extraDistr: Additional Univariate and Multivariate Distributions*. R package version 1.9.1. URL: https://github.com/twolodzko/extraDistr.

Wood, S, Busemeyer, J, Koling, A, Cox, C R, and Davis, H (2005). "Older adults as adaptive decision makers: Evidence from the Iowa Gambling Task". In: *Psychology and Ageing* 20.2, pp. 220–225. DOI: DOI:http://dx.doi.org/10.1037/0882-7974.20.2.220.

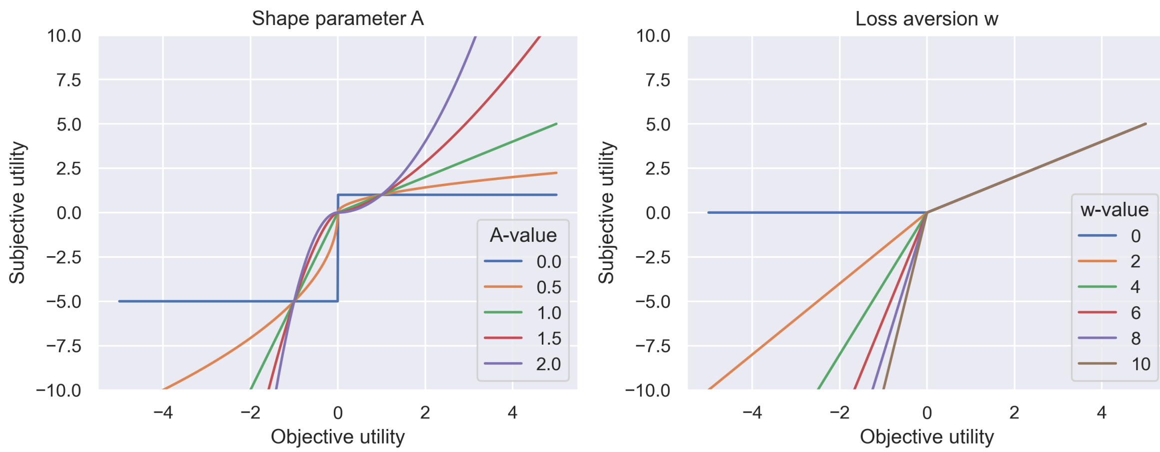
Worthy, Darrell A, Pang, Bo, and Byrne, Kaileigh A (2013). "Decomposing the roles of perseveration and expected value representation in models of the Iowa gambling task". In: *Frontiers in psychology* 4, p. 640.

Appendix A

Subjective utility parameter investigation

Figure 9

Subjective utility function with varying A and w.



Note. Plot of the subjective utility function in PLV-Delta. In the left plot, the loss aversion parameter w is kept constant at 5 while the value of A is varied. In the right plot the shape parameter A is kept constant at 1 and w is varied. This shows how the two parameters affect the subjective utility function.

Appendix B

Prior distributions

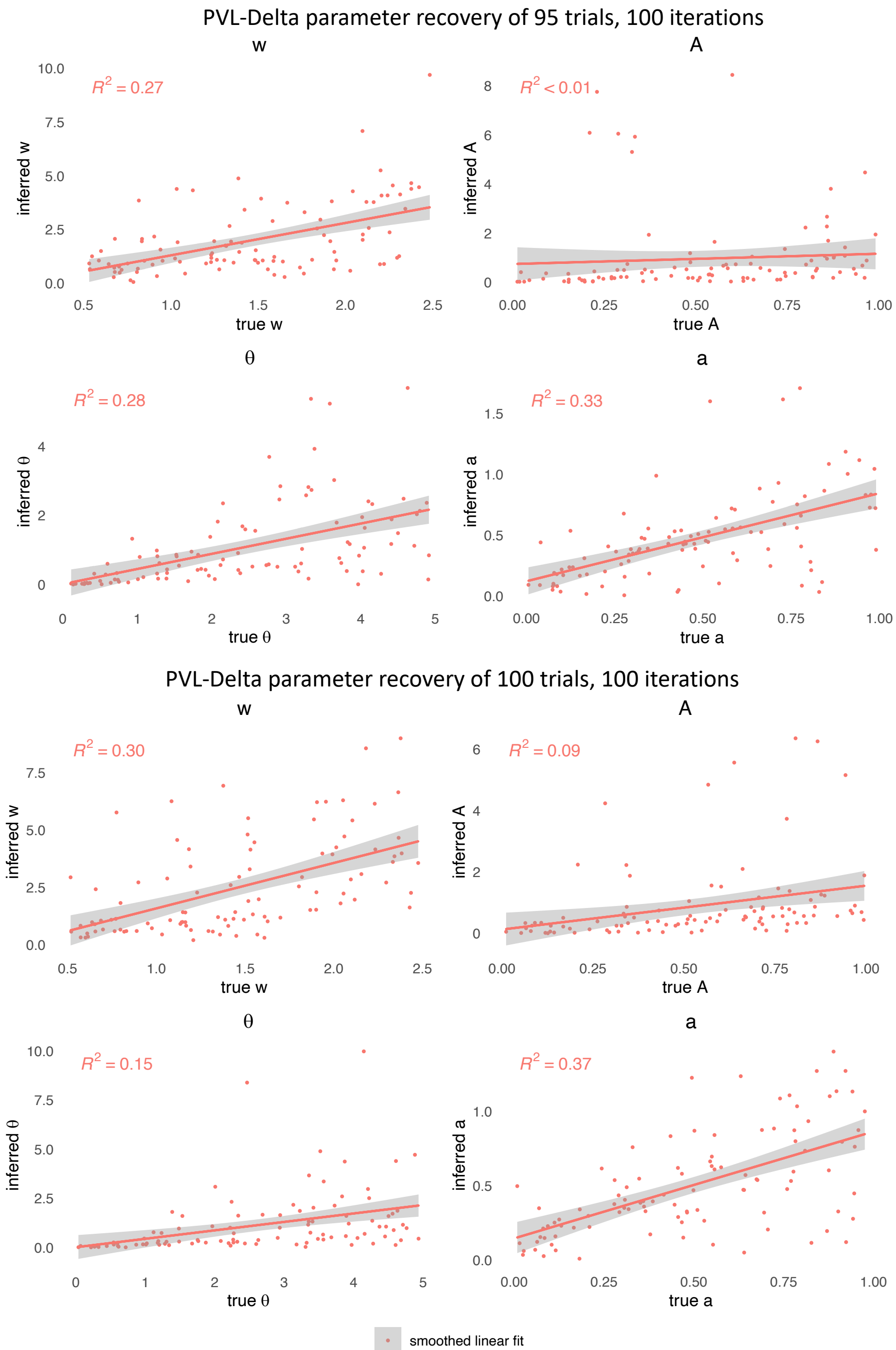
Level	Model	Parameter	Prior distribution
Subject	PVL-Delta	w	$w \sim N\left(0, 0.01^2\right), w > 0$
Subject	PVL-Delta	A	$A \sim N\left(0, 0.01^2\right), A > 0$
Subject	PVL-Delta	θ	$\theta \sim N\left(0, 0.01^2\right), \theta > 0$
Subject	PVL-Delta	A	$A \sim N\left(0, 0.01^2\right)$
Hierarchical	PVL-Delta	μ_w	$\mu_w \sim N\left(0, 1^2\right), \mu_w > 0$
Hierarchical	PVL-Delta	μ_A	$\mu_A \sim N\left(0, 1^2\right)$
Hierarchical	PVL-Delta	μ_θ	$\mu_\theta \sim N\left(0, 1^2\right), \mu_\theta > 0$
Hierarchical	PVL-Delta	μ_a	$\mu_a \sim N\left(0, 1^2\right), 1 > \mu_a > 0$
Hierarchical	PVL-Delta	λ_w	$\lambda_w \sim \Gamma(0.01, 0.01)$
Hierarchical	PVL-Delta	λ_A	$\lambda_A \sim \Gamma(0.01, 0.01)$
Hierarchical	PVL-Delta	λ_θ	$\lambda_A \sim \Gamma(0.01, 0.01)$
Hierarchical	PVL-Delta	λ_a	$\lambda_a \sim \Gamma(0.01, 0.01)$
Subject	ORL	a_{rew}	$a_{rew} \sim U\left(0, 0.01^2\right)$
Subject	ORL	a_{pun}	$a_{pun} \sim U\left(0, 0.01^2\right)$
Subject	ORL	K	$K \sim N\left(0, 0.1^2\right), K > 0$
Subject	ORL	θ	$\theta \sim N\left(0, 0.1^2\right), \theta > 0$
Subject	ORL	ω_f	$\omega_f \sim N\left(0, 0.1^2\right)$
Subject	ORL	ω_p	$\omega_p \sim N\left(0, 0.1^2\right)$
Hierarchical	ORL	$\mu_{a_{rew}}$	$\mu_{a_{rew}} \sim N\left(0, 1^2\right), 1 > \mu_{a_{rew}} > 0$
Hierarchical	ORL	$\mu_{a_{pun}}$	$\mu_{a_{pun}} \sim N\left(0, 1^2\right), 1 > \mu_{a_{pun}} > 0$
Hierarchical	ORL	μ_K	$\mu_K \sim N\left(0, 1^2\right), \mu_K > 0$
Hierarchical	ORL	μ_θ	$\mu_\theta \sim N\left(0, 1^2\right), \mu_\theta > 0$
Hierarchical	ORL	μ_{ω_f}	$\mu_{\omega_f} \sim N\left(0, 0.1^2\right)$
Hierarchical	ORL	μ_{ω_p}	$\mu_{\omega_p} \sim N\left(0, 0.1^2\right)$
Hierarchical	ORL	$\lambda_{a_{rew}}$	$\lambda_{a_{rew}} \sim \Gamma(0.01, 0.01)$
Hierarchical	ORL	$\lambda_{a_{pun}}$	$\lambda_{a_{pun}} \sim \Gamma(0.01, 0.01)$
Hierarchical	ORL	λ_K	$\lambda_K \sim \Gamma(0.01, 0.01)$
Hierarchical	ORL	λ_θ	$\lambda_\theta \sim \Gamma(0.01, 0.01)$
Hierarchical	ORL	λ_{ω_f}	$\lambda_{\omega_f} \sim \Gamma(0.01, 0.01)$
Hierarchical	ORL	λ_{ω_p}	$\lambda_{\omega_p} \sim \Gamma(0.01, 0.01)$

Appendix C

Preliminary parameter recovery

Figure 10

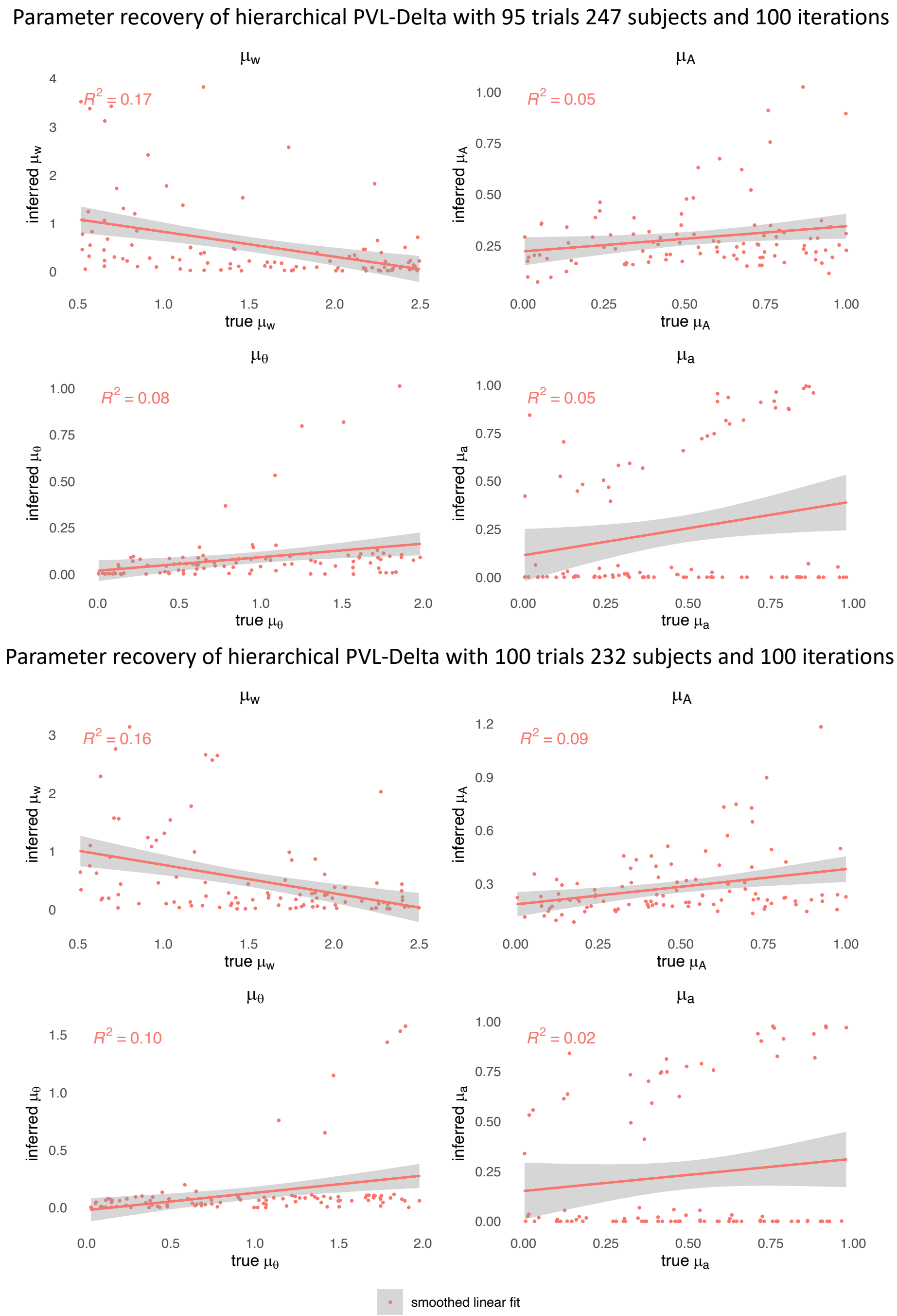
Parameter recovery of PVL-Delta (subject-level)



Note. PVL-Delta with 95 trials (top) and with 100 trials (bottom). The difference is not noteworthy.

Figure 11

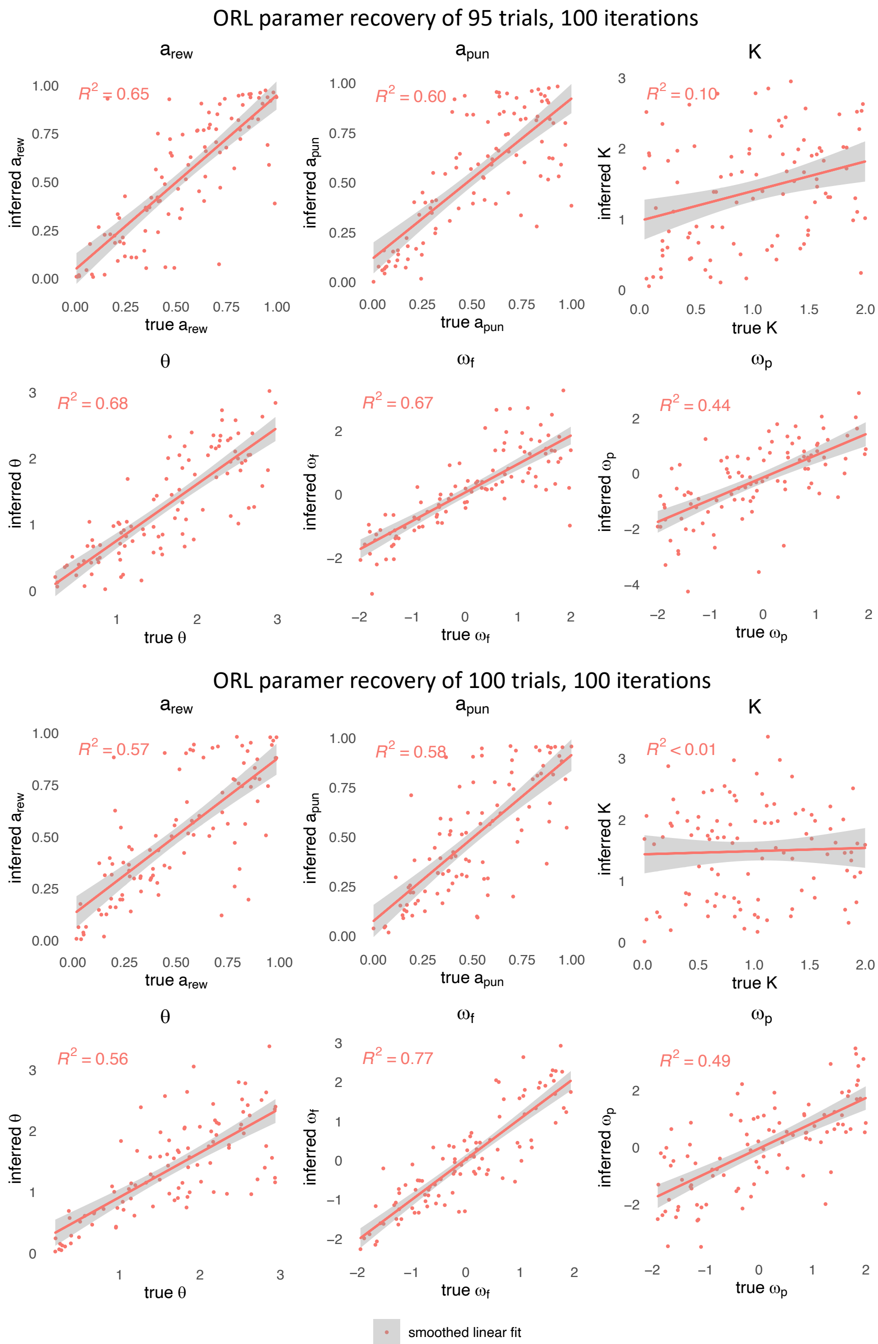
Parameter recovery of hierarchical PVL-Delta



Note. Hierarchical PVL-Delta with 95 trials and 247 subjects (top) and with 100 trials and 232 subjects (bottom). The extra five trials and 15 subjects less do not make a noteworthy difference in the parameter recovery of 100 trials and 232 subjects.

Figure 12

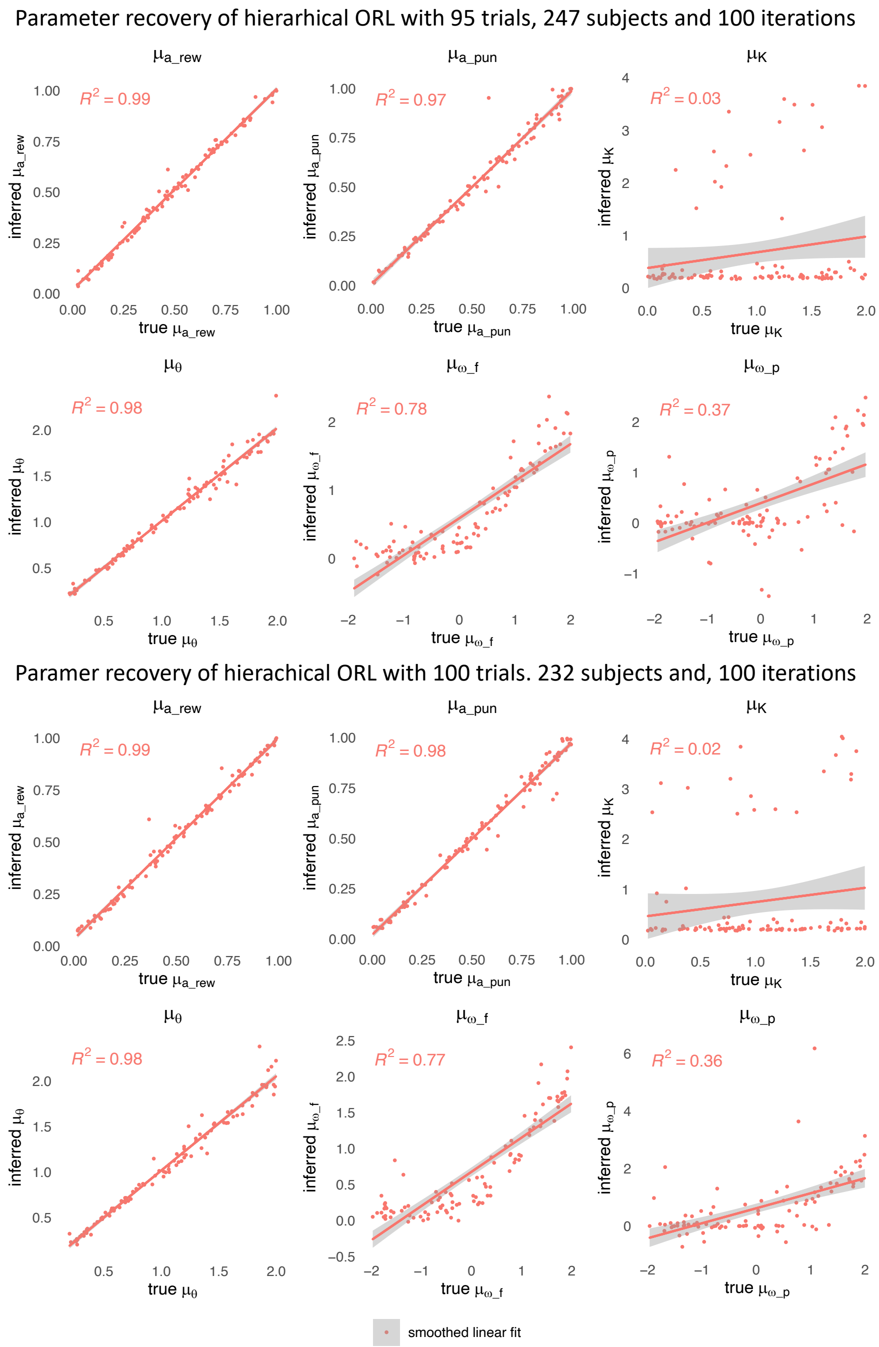
Parameter recovery of ORL (subject-level)



Note. ORL with 95 trials (top) and with 100 trials (bottom). The extra five trials do not make a noteworthy difference in the parameter recovery of 100 trials.

Figure 13

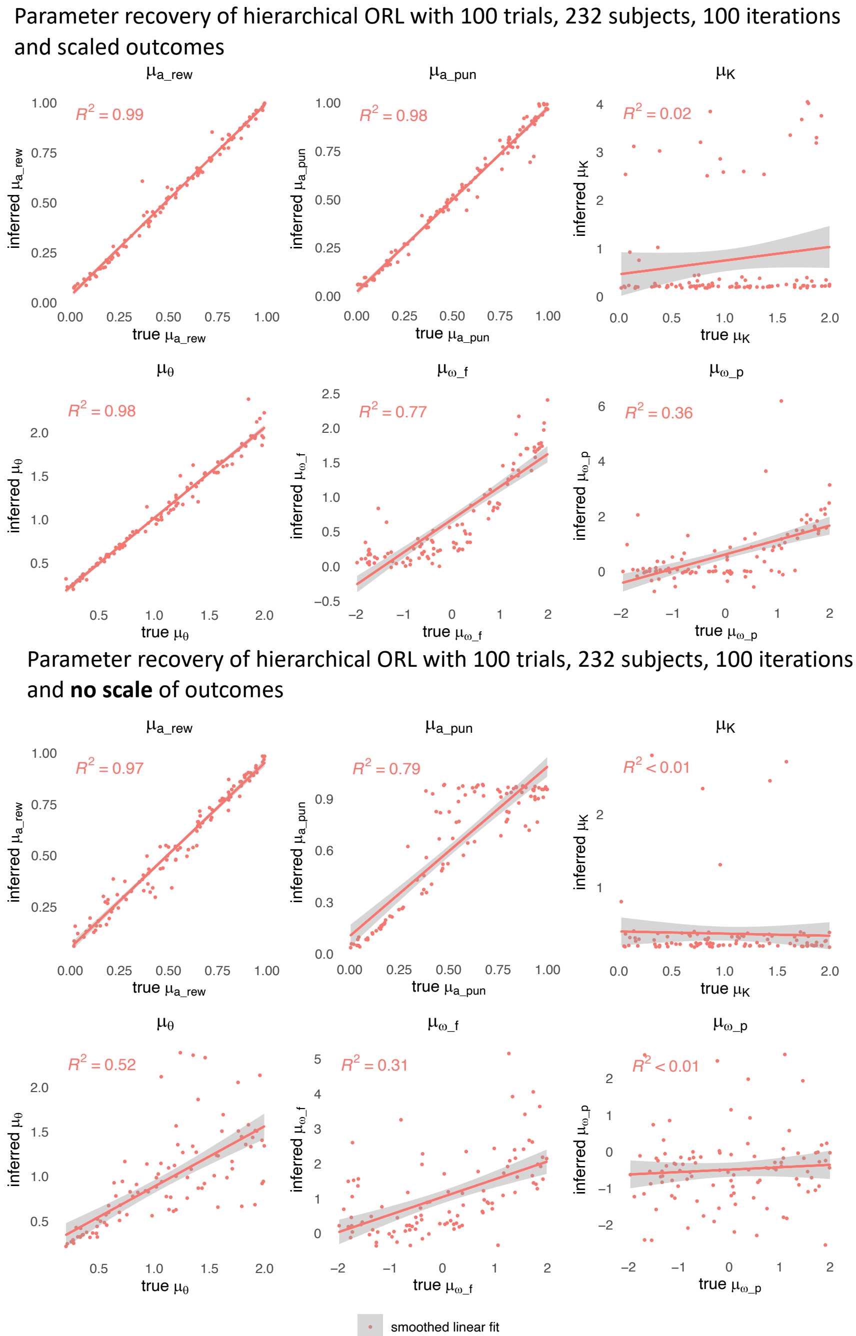
Parameter recovery of hierarchical ORL



Note. Hierarchical ORL with 95 trials and 247 subjects (top) and with 100 trials and 232 subjects (bottom). The extra five trials and 15 subjects less do not make a noteworthy difference in the parameter recovery of 100 trials and 232 subjects.

Figure 14

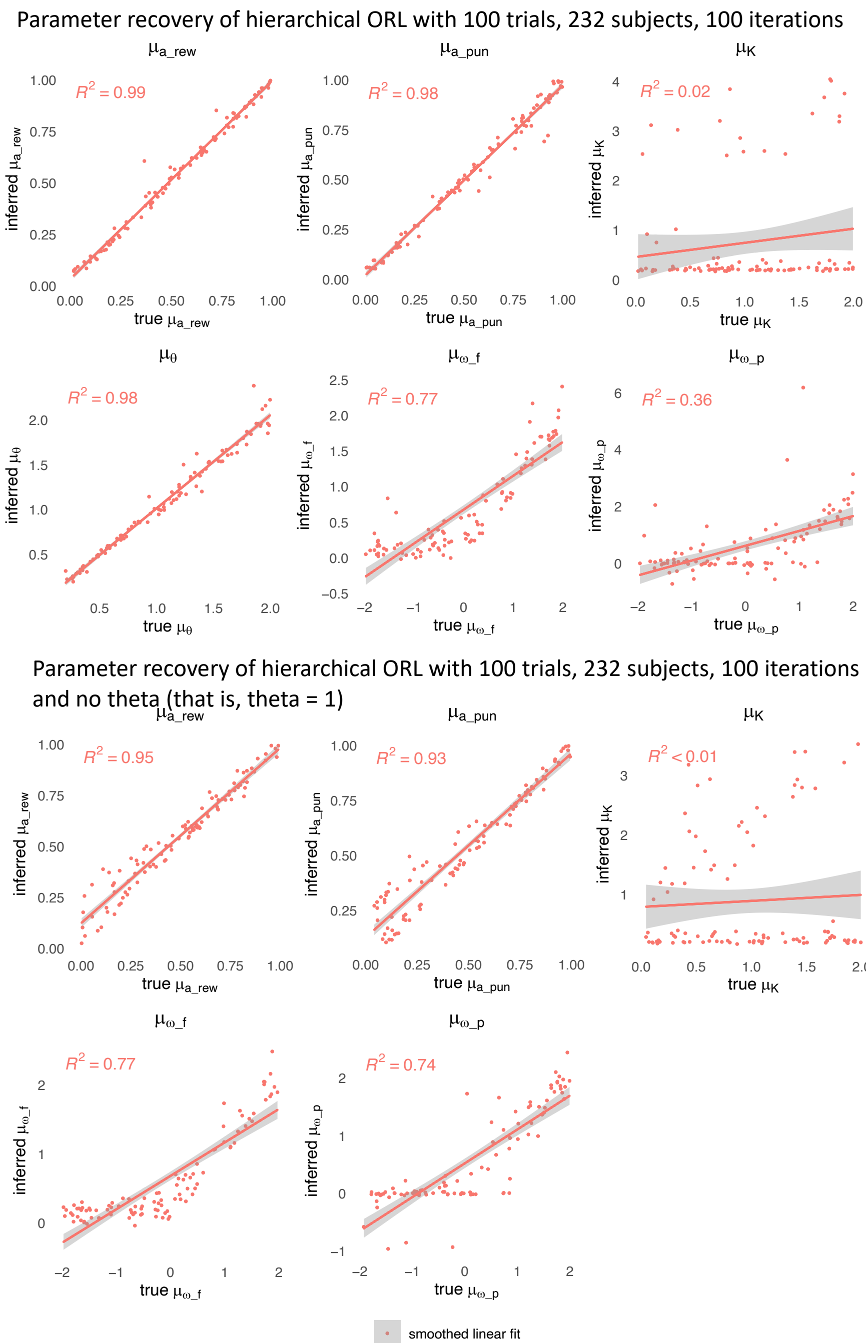
Parameter recovery of hierarchical ORL with and without scaling of outcomes



Note. ORL model with scaling (top) and without scaling (bottom). Scaling the outcomes helps to recover ω_f, ω_p and θ.

Figure 15

Parameter recovery of hierarchical ORL with and without θ



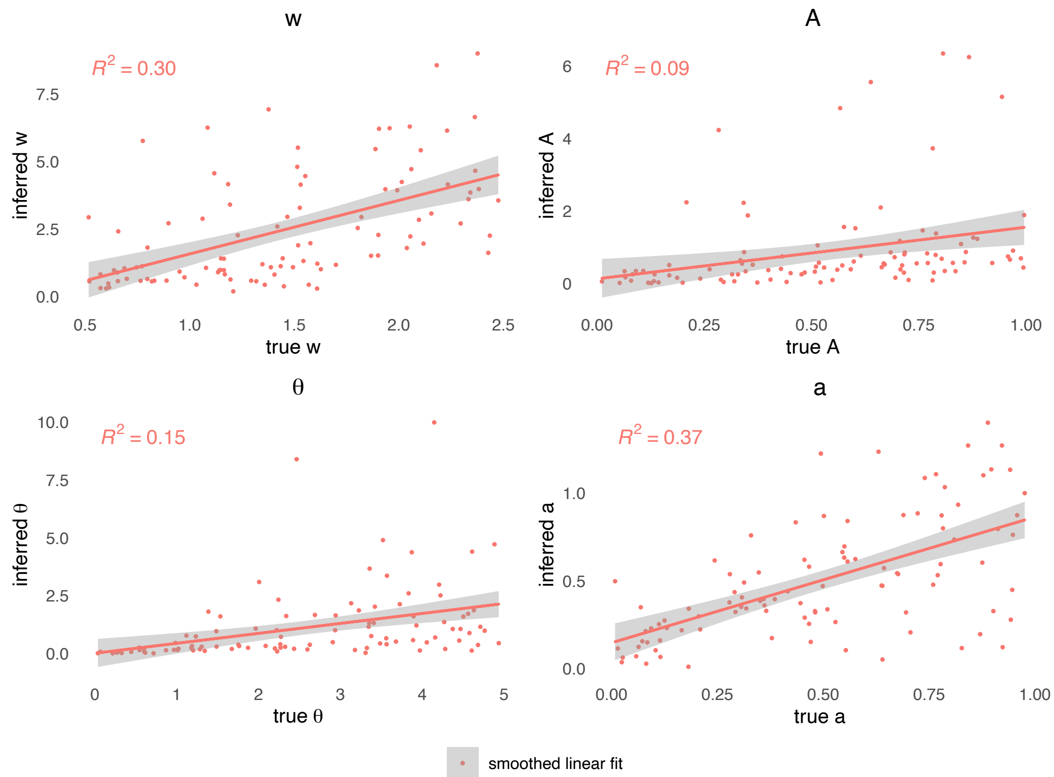
Note. ORL model including θ (top) and excluding θ (bottom). Excluding θ improves the recovery of ω_p, though it still will not sample values below zero.

Appendix D

Subject-level parameter recovery

Figure 16

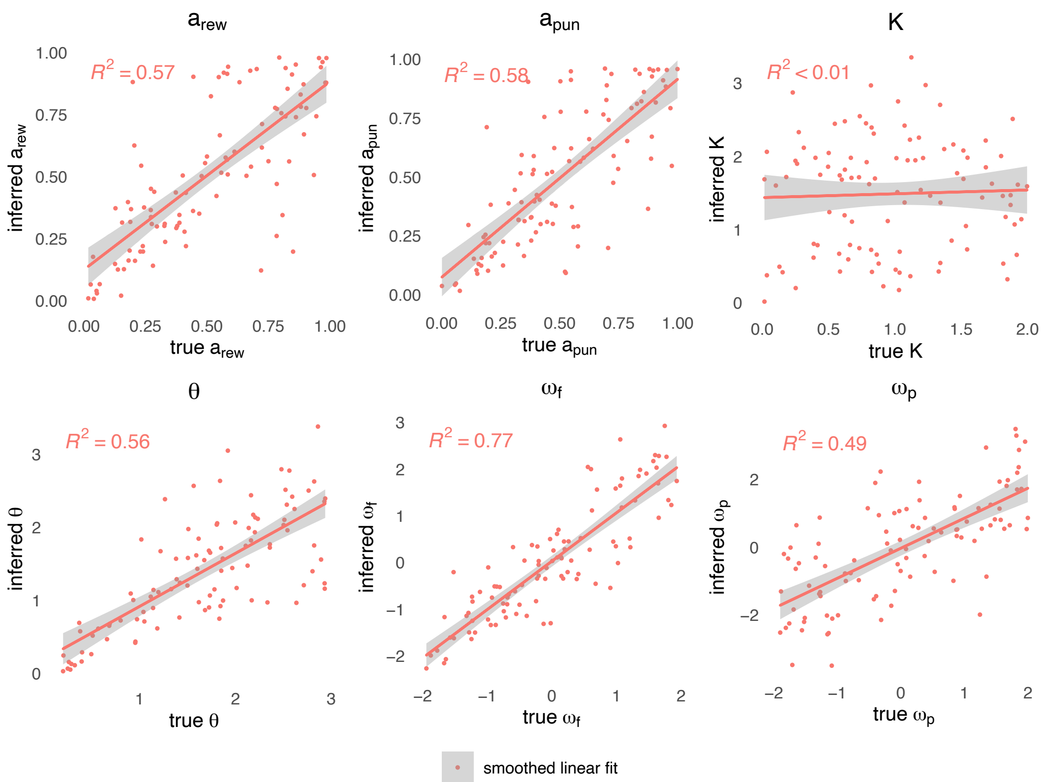
PVL-Delta parameter recovery (subject-level)



Note. Parameter recoveries for each subject-level, a linear model fit and R^2-values.

Figure 17

ORL parameter recovery (subject-level)



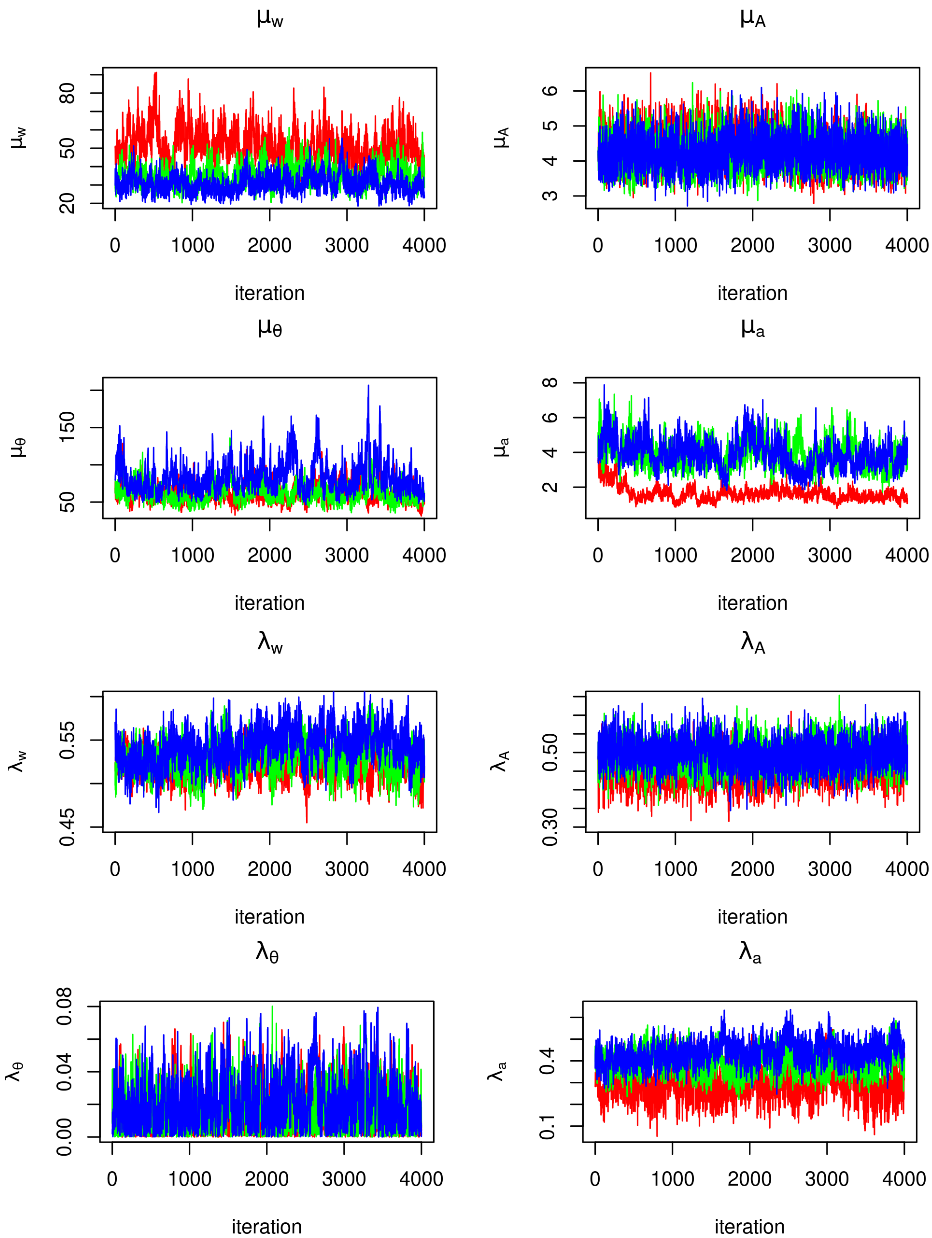
Note. Parameter recoveries for each subject-level parameter as well as a linear fit and R^2-values.

Appendix E

Markov chains

Figure 18

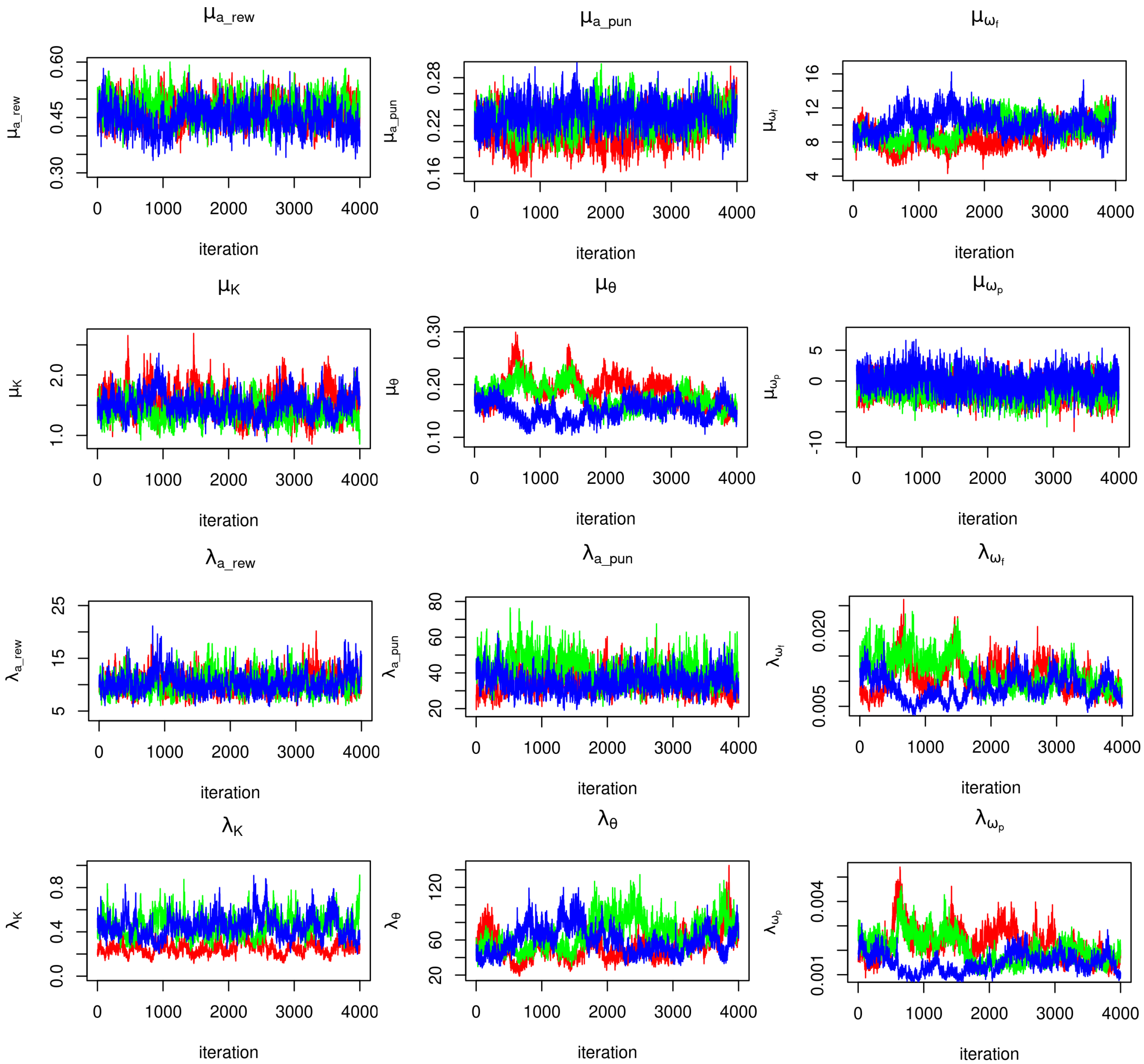
Markov chain convergence plots for hierarchical PVL-Delta parameters.



Note. Markov chains from MCMC simulation for parameter estimation of the hierarchical PVL-Delta.

Figure 19

Markov chain convergence plots for hierarchical ORL parameters.



Note. Markov chains from the MCMC simulation for parameter estimation of the hierarchical ORL.

Appendix F

$\hat{R}$-values

Table 2

$\hat{R}$-values for PVL-Delta

Parameter	Deviance	λ_A	λ_a	λ_θ	λ_w	μ_A	μ_a	μ_θ	μ_w
$\hat{R}$	1.0431	2.4888	1.0154	1.5091	3.4988	1.3095	1.3138	1.0400	2.0094

Table 3

$\hat{R}$-values for ORL

Parameter	$\hat{R}$
Deviance	1.926522
λ_K	2.393207
$\lambda_{a_{pun}}$	1.283013
$\lambda_{a_{rew}}$	1.001669
λ_{ω_f}	1.242726
λ_{ω_p}	1.477710
λ_θ	1.170795
μ_K	1.128628
$\mu_{a_{pun}}$	1.278411
$\mu_{a_{rew}}$	1.152920
μ_{ω_f}	1.362650
μ_{ω_p}	1.231957
μ_θ	1.557384

Appendix G

Source code

The code used for data pre-processing and analysis is freely available at our github repository:

[redacted]