Computational Brain & Behavior
https://doi.org/10.1007/s42113-025-00261-9

RESEARCH

Improving Parameter Recovery in Computational Models: Employing Outlier-Insensitive Loss Functions

Mingqian Guo[1] · **Karin Roelofs**[1,2] · **Bernd Figner**[1,2]

Received: 21 January 2025 / Accepted: 23 October 2025

Abstract
The reliability of parameter estimation is crucial in using computational models for choice data in decision-making tasks, especially so since the cognitive meaningful parameters within these models often are leveraged for further analysis. Typically, model-fitting involves using the model log-likelihood as the loss function to quantify discrepancies between model predictions and observed data. However, outlier data in choice datasets can bias parameter estimation when using log-likelihood. Alternative loss functions that are less sensitive to outliers are available. In this study, we compared a total of 3 such outlier-insensitive loss functions with the log-likelihood function in terms of parameter recovery. We compared their performance in both a reinforcement learning model in a learning paradigm and a hyperbolic model in an intertemporal choice paradigm, in both systematically varying the presence of outliers (ranging from no outliers to 25% of the data being outliers). Our parameter recovery results show that even a small proportion of outlier data can substantially impair parameter identification when using the log-likelihood function, especially for the choice consistency/explore–exploit trade-off parameter. In contrast, outlier-insensitive loss functions markedly improve the recovery of computational model parameters. Moreover, our power analysis further suggests that even a small proportion of outlier trials (e.g., 5%) can potentially undermine the statistical power to detect condition differences, underscoring the importance of accounting for outliers when using cognitive models as measurement tools. Based on our results, we recommend using the outlier-insensitive loss functions for non-hierarchical model estimation as it performed well across both the learning and the intertemporal choice paradigms and under varying degrees of outlier presence.

Introduction

Computational models of decision-making have become indispensable tools in both theoretical and empirical research, offering profound insights into the cognitive processes underpinning observed behaviors (Farrell & Lewandowsky, 2018). These models not only provide a formal explanation of behavioral data but also break it down into cognitively meaningful parameters. For example, the hyperbolic discounting model, widely used in analyzing data collected with intertemporal choice tasks, disentangles raw choice data into two key parameters: the discount rate, which is assumed to reflect participants' preferences, and the inverse temperature, which captures choice consistency (Mazur et al., 1987). These parameters are valuable for probing the cognitive mechanisms underlying experimental manipulations. For instance, recent studies utilizing the hyperbolic model have demonstrated that cognitive load primarily affects choice consistency rather than altering the discount rate (Jiang & Dai, 2023; Olschewski et al., 2018).

Beyond testing conditions, computational models are also pivotal in individual differences research, particularly in contrasting patient populations with healthy groups. Collins et al. (2014), for instance, revealed that deficits in learning stimulus-response associations in individuals with schizophrenia are solely due to working memory dysfunction, while reinforcement learning abilities remain unaffected. This approach, known as theory-driven computational psychiatry, has rapidly gained traction in psychiatric research (Geng et al., 2022; Huys et al., 2016; Montague et al., 2012).

✉ Mingqian Guo
mingqian.guo@ru.nl

[1] Behavioural Science Institute, Radboud University, Thomas van Aquinostraat 4, Nijmegen, GD 6525, The Netherlands

[2] Donders Institute for Brain, Cognition and Behaviour, Radboud University, NijmegenThomas van Aquinostraat 4, GD 6525, The Netherlands

Published online: 06 November 2025

To estimate these cognitive meaningful parameters, computational models are often fitted with parameter estimation techniques like Maximum Likelihood Estimation (MLE) or Maximum a Posteriori (MAP), which identify the best-fitting parameters for the observed data. These methods often use the Log-Likelihood Function (LLF) as the loss function to quantify the difference between the model prediction and the observed data, a standard with desirable properties for model fitting. LLF also aids in computing model comparison metrics like the Akaike Information Criterion (AIC) and marginal likelihood (Akaike, 1974; Kass & Raftery, 1995; Wagenmakers & Farrell, 2004). The accuracy of these estimations is crucial for subsequent analyses, highlighting the importance of their reliability.

Outlier data is a common phenomenon in behavioral data analysis. In reaction time studies, outlier handling methods are well-studied, yet in computational models of choice data, their impact on parameter estimation is less well understood. The LLF, commonly used in computational model fitting, is relatively sensitive to outlier data, which can bias parameter estimation and compromise subsequent analyses of computational model parameters. These outlier trials might be caused by a participant's temporary inattention towards the task or possibly also by a decision strategy which is not captured by the computational model. For example, unexpected choices by participants seeking to explore unattended options can generate "outlier data," typically accounted for by using mixture models or lapse models (Nassar & Frank, 2016; Padoa-Schioppa, 2022; Vincent, 2016). Whereas these unexplained processes sometimes have minimal impact, recent studies suggest that factors like random exploration can significantly influence choice data (Pisupati et al., 2021). This consideration becomes even more critical with the increase in using online studies, which often yield lower-quality data than traditional in-lab experiments (Ratcliff & Kang, 2021). Thus, it is essential to carefully address outlier data to ensure that the estimated parameters accurately reflect the intended cognitive properties.

To solve this problem, Bayesian approaches like the Hidden Markov Model (HMM) have been integrated into computational decision-making models (Ashwood et al., 2022; Bishop, 2006; Kucharský et al., 2021; Li et al., 2024; Venditto et al., 2024). HMM offers the advantage of estimating the probability that participants are following a model-specified decision strategy versus making random choices. Nevertheless, integrating HMM into computational models can be challenging and may complicate the model, thereby potentially reducing parameter estimation reliability. Kucharský et al. (2021) used a combination of HMM and an evidence accumulation model to study the speed-accuracy trade-off. However, this integration resulted in an excessively high number of parameters for the experimental design. Consequently, they had to substantially simplified the evidence accumulation model to facilitate appropriate parameter estimation.

An alternative approach to address outlier impact is to use an outlier-insensitive loss function, which de-emphasizes incompatible data, leading to more reliable parameter estimations (Ghosh et al., 2017). This method, simpler than advanced approaches like HMM and widely tested in the machine learning literature, significantly enhances model prediction accuracy (Song et al., 2022). Yet, in computational decision-making models applied to human choice data, the effectiveness of these loss functions in parameter recovery remains untested.

To assess whether the use of alternative outlier-insensitive loss functions—commonly employed in the field of computer vision—can improve parameter estimation in the presence of outlier data, we employed (a) the reinforcement learning model in the orthogonal go-no go task (Guitart-Masip et al., 2012) and (b) the generalized hyperbolic model in an intertemporal choice task (Vincent, 2016) to simulate choice data. We systematically introduced varying proportions of random choices into the simulations and then applied different loss functions to test the reliability of the parameter estimation approaches. Specifically, we first compared parameter recovery across different loss functions when varying proportions of outlier trials were introduced. Second, we examined the impact of outlier trials on the statistical power to detect condition differences in reinforcement learning models. Finally, we compared the performance of an outlier-insensitive loss function with the Dynamic Noise approach proposed by Li et al. (2024), to illustrate that Dynamic Noise may not be well-suited for non-learning tasks. These analyses highlight the potential weakening effect of outlier data on the strength of evidence derived from computational models in disentangling preferences and choice consistency.

Methods

The reliability of the parameter estimation is often tested by the parameter recovery (Wilson & Collins, 2019). Parameter recovery includes three steps: firstly, the simulated data are generated with the model and certain parameter values. Second, the model that generated the data is fitted to the simulated data. Thirdly, the reliability of the parameter estimation is investigated by correlating the true data generation parameter values with the fitted parameter values. In the next part, we will introduce the model we used to test the parameter reliability and the loss function that we used to fit these models.

Reinforcement Learning Model

We applied a reinforcement learning model to a probabilistic monetary go/no-go bandit task in our case study. The task design orthogonally pairs valence (reward or punishment) with action (go or no-go). This task is frequently utilized to dissociate Pavlovian bias from instrumental learning (Dayan et al., 2006; Guitart-Masip et al., 2014; Guitart-Masip, Huys, Guitart-Masip et al., 2012a, b; Swart et al., 2017). It consists of four cues, each associated with different actions and outcomes. Each cue is repeated 45 times. A correct response has a 80% probability to result in a desirable outcome (i.e., receive reward/avoid punishment) and an incorrect response has a 20% probability to result in a desirable outcome (i.e., receive reward/avoid punishment).

The reinforcement learning Q-learning model is used for modeling this task (Sutton & Barto, 2018; Watkins & Dayan, 1992). For each trial, the instrumental action value is updated using the Rescorla-Wagner update rule after observing the outcome of an action:

$$Q_t(\mathrm{s},\mathrm{a}) = Q_t(\mathrm{s},\mathrm{a}) + \epsilon\ (R_t - Q_{t-1}(\mathrm{s},\mathrm{a})) \tag{1}$$

In this equation, ε is the learning rate (bounded between 0 and 1), and R_t the outcome of the action at trial t (which could be 1, 0, or −1). Given the potential for negative rewards, the instrumental action values $Q_t(S,A)$ are initialized at 0 (Ballard & McClure, 2019; Zhang et al., 2020). The difference between the outcome and the instrumental action value is the reward prediction error (RPE).

In the simplest model, it is assumed that the action weight $W_t(s_t,a_t)$ is solely determined by the instrumental action value:

$$W_t(\mathrm{s},\mathrm{a}) = Q_t(\mathrm{s},\mathrm{a}) \tag{2}$$

However, to account for the influence of Pavlovian and go biases, the two additional free parameters go bias (b) and weight for the Pavlovian biases (π) are added to the action weight:

$$W_t(\mathrm{s},\mathrm{a}) = \begin{cases} Q_t(\mathrm{s},\mathrm{a}) + \pi V_t(\mathrm{s}) + b & if\ a = Go \\ Q_t(\mathrm{s},\mathrm{a}) & else \end{cases} \tag{3}$$

Where $V_t(s_t)$ reflects the Pavlovian biases, equaling 1 for go cues and −1 for no-go cues.

The decision weight is transformed into a choice probability through a softmax function:

$$p(a_t|s_t) = \frac{e^{W_t(\mathrm{s},\mathrm{a})\times\tau}}{\sum_{a'} e^{W_t(\mathrm{s},a')\times\tau}} \tag{4}$$

Where τ is the so-called inverse temperature parameter and governs the choice consistency.

To simulate choice data, we sampled 200 agents' parameters using the following distributions:

$$\in \sim TN(0.5, 0.2, 0, 1)$$

$$b \sim TN(0.3, 0.1, 0, 1)$$

$$\pi \sim TN(0.4, 0.2, 0, 1)$$

$$\tau \sim TN(4.5, 2, 1, 10)$$

Here, TN refers to the truncated normal distribution, and the last two terms in the parenthesis denote its lower and upper bounds.

Hyperbolic Discounting Model

The intertemporal choice task (ITC) involves participants deciding between an option that offers a larger reward but with a longer delivery time (later-larger option, LL) and another option that offers a smaller reward available sooner (sooner-smaller option, SS). Generally, people tend to discount the future outcome, leading to a preference for immediate outcomes. Hyperbolic discounting models are often used for modeling ITC due to the ability to account for the phenomenon of time preference reversals (Luhmann, 2013; Mazur et al., 1987). This effect describes the shift in people's preferences towards later-larger options as the delay time increases:

$$SV = \frac{OV}{1+kD} \tag{5}$$

In this equation, SV denotes the subjective value, OV represents the objective value, D stands for the delay time, and k signifies the discount rate.

The Generalized Hyperbolic Model expands upon the standard hyperbolic model by positing that discounting occurs over the subjective time. Although there are numerous variants of the generalized hyperbolic model, we opted for a modified Rachlin model (Rachlin, 2006; Vincent & Stewart, 2020)

$$SV = \frac{OV}{1+(kD)^s} \tag{6}$$

.

Here, s is the sensitivity parameter of the delay time. A prominent advantage of this modified Rachlin model is

that the two parameters have a relatively low correlation, thereby enhancing the model's identification power (Guo et al., 2025; Vincent & Stewart, 2020).

Similar to the reinforcement learning model, choices are generated through a softmax function:

$$p\left(LL\right) = \frac{e^{SV_{LL} \times \tau}}{e^{SV_{SS} \times \tau} + e^{SV_{LL} \times \tau}} \tag{7}$$

We simulated data based on the experimental design reported by Ikink et al. (2023). This study employed a binary intertemporal choice task with 60 trials. In each trial, both the reward magnitude and the delay duration were explicitly presented to participants and they did not need to actually experience the delay. The values of the SS and LL options varied across trials. SS rewards ranged from €16–70, delivered either today or in 14 days. LL rewards were 3–50% larger than the SS amount, delivered after 3, 14, or 28 days. There were 43 unique stimulus combinations in total. Among them, 26 appeared once and 17 appeared twice. Detailed information about the stimulus combinations and their frequencies is provided in the appendix. Participants had no time limit for their responses. More details about the design can be found in Ikink et al. (2023). We sampled parameters for 200 agents from the following distributions:

$k \sim TN(1, 0, 0.004, 0.06)$

$\tau \sim TN\left(0.1, 10, 0.4, 2.3\right)$

$s \sim TN\left(0.3, 2, 1, 0.3\right)$

Methods of Dealing with Outlier Data

Outlier data are commonly encountered in choice data sets. While these outliers could be insightful for understanding the cognitive process, they do not necessarily align with the specific focus of a study.

In reaction time (RT) data, where slow responses may be due to attention lapses, and fast responses may result from "fast guesses" (Lerche et al., 2017; Ratcliff, 1993), such data are detrimental for making inferences based on RT data. To address this, various methods have been developed for managing outliers in RT analysis, like the median absolute deviation (MAD) method (Leys et al., 2013) or excluding data points who are outside of the range of +/−3 standard deviations (SD) from the mean. However, these methods cannot be readily applied to the modeling of choice data.

It has been suggested that RT data could be used to exclude outliers in choice modeling (Peters & D'Esposito, 2020). While fast responses can indeed be indicative of outliers, there can be cases where participants initially lack focus at the start of a trial, but later concentrate on the stimulus and make their decisions. In such instances, while the RT data may indeed need to be excluded from the analysis, the choice data remain valuable.

In the case of a Bernoulli distribution or a categorical distribution, the cross entropy (CE) between the observed data (y) and the model prediction ($\widehat{y}$) and the LLF are mathematically identical:

$$CE = LLF = \sum\ -p\left(y\right)\log\left[p\left(\widehat{y}\right)\right] = \sum\ \ y\log\left[p\left(\widehat{y}\right)\right] \tag{8}$$

Thus, LLF can be treated as the loss function to quantify the difference between the observed data and the model prediction. Therefore, LLF is often used for determining the values of free parameters in a model with techniques such as MLE or Bayesian estimation (Daw, 2011; Gelman et al., 2013; Pawitan, 2001). These estimation techniques use gradient-based algorithms such as gradient descent, Newton's method in MLE, or Hamiltonian Monte Carlo (HMC) in Bayesian estimation. The gradient's characteristics significantly influence parameter estimation.

The logarithmic function within CE/LLF possesses a property wherein its gradient, $\Delta\ \log\left(x\right) = \frac{1}{x}$ decreases quadratically when x is less than one. This property results in heavier penalties for predictions that are incorrect but confident, compared to less confident ones. For instance, if a model confidently predicts a choice with a 0.95 probability but it is incorrect, the LLF is −2.996. In contrast, if the predicted probability is 0.5, the LLF is −0.693. This asymmetry enables the model to utilize the full informational content of the dataset, prioritizing adjustments based on incorrect predictions—a desirable trait.

However, the presence of outlier trials can disproportionately influence this process. In an outlier trial, a participant may choose the option that is opposite to the prediction that the model makes with high confidence. Because LLF assigns especially large penalties when confident predictions are contradicted, even a small number of such trials can substantially skew the overall likelihood and bias parameter estimation.

To demonstrate the significant impact of outlier trials on parameter estimation, we conducted a simulation using the reinforcement learning model in a go/no-go task. We set the go bias and Pavlovian bias parameters to 0, learning rate to $\alpha\ = 0.5$, and the inverse temperature to $\tau\ = 3$. In this case, the model is simplified to a standard two parameter Q-learning model. To generate outlier trials, we first identified trials on which the model predicted one option with high confidence (i.e., a choice probability greater than 0.75). From this subset, we randomly selected 10 trials (5% of the data). For these 10 trials, we then forced the model to output the low-probability option (i.e., the one with predicted probability < 0.25).

This procedure created outlier trials that contradicted the model's high-confidence predictions, mimicking atypical responses due to noise, attentional lapses, or erroneous responses. The results are depicted in Fig. 1, where the dark area signifies the maximum log-likelihood function (LLF) region. As proven by Ballard and McClure (2019), this region is relatively expansive in the parameter space, indicating that the reinforcement learning model is prone to an interchangeable relationship between the learning rate and temperature parameters. While this issue is not the primary focus of our current study, Ballard and McClure (2019); Shahar et al. (2019) suggest that it can be improved by jointly modeling reaction time and choice data. The true parameter, marked by a white dot, lies at the center of the bottom area in the absence of outlier data. However, with the inclusion of outliers, the lowest area shifts significantly away from the true parameter, underscoring how even a small proportion of outliers can dramatically alter the likelihood surface and, consequently, the parameter estimation results.

Besides LLF, other loss functions exhibit less sensitivity to outliers. For instance, mean squared error (MSE) is commonly used in Gaussian distribution linear regression. In the next section, we will explore several loss functions that are more robust to the presence of outlier data.

Mixture Model

One approach to handle noisy data is the mixture model or lapse model, which is a combination of a softmax function with a random choice strategy (Ashwood et al., 2022; Collins & Frank, 2012; Gelman et al., 2013; Gershman & Bhui, 2020; Nassar & Frank, 2016; Sutton & Barto, 2018). The formula can be represented as follows:

$$P(A|A,B) = (1-e)\,\frac{1}{1+\exp(-\tau\,(V(A)-V(B))} + \frac{e}{2} \tag{9}$$

The e-Softmax posits the choice is generated from a mixture distribution between a uniform distribution and a Softmax function. Parameter e is often called the lapse parameter or the "trembling hand" parameter in the literature (Harless & Camerer, 1994; Krefeld-Schwalb et al., 2022), it decides how much people use the softmax choice function versus a random choice decision rule. The output of the e-Softmax rises from a minimum value e/2 to a maximum value 1-e/2, thereby preventing the single trial LLF from being excessively small.

Empirically, the introduction of mixture can improve the model-fitting in animal and human studies. The interpretation of this parameter strongly depends on the specific model and task. In the reinforcement learning literature, an increase in the ε parameter of the softmax function is typically interpreted as reflecting random exploration.

This concept of mixture models has found wide application in RT modeling, such as the Drift-Diffusion Model (DDM). Ratcliff and Tuerlinckx (2002) demonstrated that the combination of a uniform distribution and a Wiener process improves the reliability of DDM parameters. Computational packages like HDDM (Wiecki et al., 2013) and GLAM (Molter et al., 2019) have also incorporated this structure to mitigate the impact of outlier data.

Outlier Insensitive Loss Function

In contrast to asymmetric objective functions like LLF, symmetric objective functions treat all incorrect predictions of the model equally, regardless of the model prediction's strength. For instance, Ghosh et al. (2017) demonstrated Mean Absolute Error (MAE) is insensitive to outlier data:

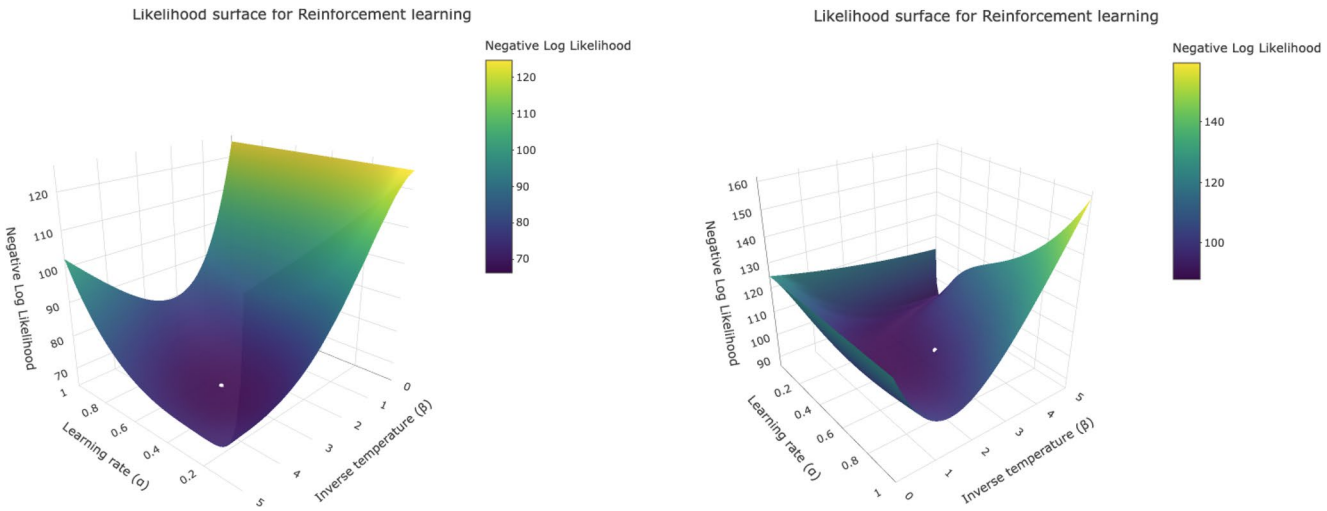

Fig. 1 Log-likelihood surface of the reinforcement learning model. Plot A shows the likelihood surface without the outlier data, whereas the Plot B shows likelihood surface with 10 outlier trials. The white represents the true parameter value used in the simulation ($\alpha = 0.5,\ \tau = 3$)

$$MAE = \sum |y - \widehat{y}| / N \tag{10}$$

Where N is the data point number, y is the observed data and $\widehat{y}$ is the model prediction. The gradient of MAE is:

$$\frac{dMAE}{d\widehat{y}} = \begin{cases} +1, & if\ model\ \ prediction\ is\ correct \\ -1, & if\ model\ prediciton\ is\ incorrect \end{cases} \tag{11}$$

While MAE reduces the influence of outliers, it poses a crucial challenge—the convergence of model fitting requires more iterations and data points than other approaches like LLF. Studies focusing on imaging classifications reveal that to achieve similar performance as LLF in the absence of or with very few outlier data, MAE requires several times more data points and model-fitting time (Song et al., 2022).

Zhang and Sabuncu (2018) suggest that, instead of using a log function to transform the likelihood, a Box-Cox transformation might be preferable as it can be considered a combination of MAE and LLF, termed generalized cross entropy (GCE):

$$GCE = y\frac{1-\widehat{y}^q}{q} \tag{12}$$

In this formula, q is a hyperparameter defined before model fitting. When $q = 1$, GCE is equivalent to MAE since their gradient is identical. When $q = 0$, GCE has an identical gradient as LLF. Therefore, a q in the range $[0,1]$ can be viewed as a combination of MAE and LLF. As q approaches 0, GCE performs like MAE, while as q approaches 1, GCE closely aligns with LLF. One point to note is that, beyond the Box–Cox–based GCE, many other transformation functions exist. We focused on the Box–Cox–based GCE because it is well established and has demonstrated good performance in previous studies. The Yeo–Johnson transformation, which has the advantage of handling negative values, was not considered since our data are non-negative, in which case it reduces to Box–Cox. Given the wide range of possible transformation and loss functions, a systematic exploration was not feasible.

The third loss function introduced is the so-called t-log function by Ding and Vishwanathan (2010).

The t-log function is the extension of the log function derived from statistical physics (Naudts, 2002).

$$\log_t(y) = \begin{cases} \log(y), t = 1 \\ \frac{y^{1-t}-1}{1-t}, \ otherwise \end{cases} \tag{13}$$

Where t is a hyperparameter like q in GCE. When systematically verifying the temperature hyperparameter, we observe that the curvature of the t-log function diminishes with decreasing temperature. Additionally, the t-log function is bounded above by $-\frac{1}{1-t}$, hence it is also referred to as a bounded-log function. In Fig. 2, we illustrate the relationship between the input and output of outlier-insensitive loss functions, such as the GCE and t-log function, along with the mixture model. As the temperature parameter in the t-log function decreases and the q parameter in GCE increases, the slope of these functions diminishes. This indicates that the impact of incorrect but confident predictions made by the model is reduced.

In addition to these loss functions, there are several prevalent options like symmetric cross entropy (SCE) and so on (Wang et al., 2019). However, most of them rely on two hyperparameters, which can complicate their usage. Furthermore, many studies show their performance is close to GCE (Song et al., 2022). Therefore, in our subsequent simulation and parameter recovery studies, we will only test these three outlier-insensitive loss functions and their mixture model performance described above in handling outlier data.

Simulation and Model-Fitting

We randomly sampled 200 parameter combinations from the beforementioned distributions, with each combination representing a hypothetical agent. For each agent, synthetic datasets were simulated 100 times to reduce the influence of stochastic variability. In each simulation, the indices of

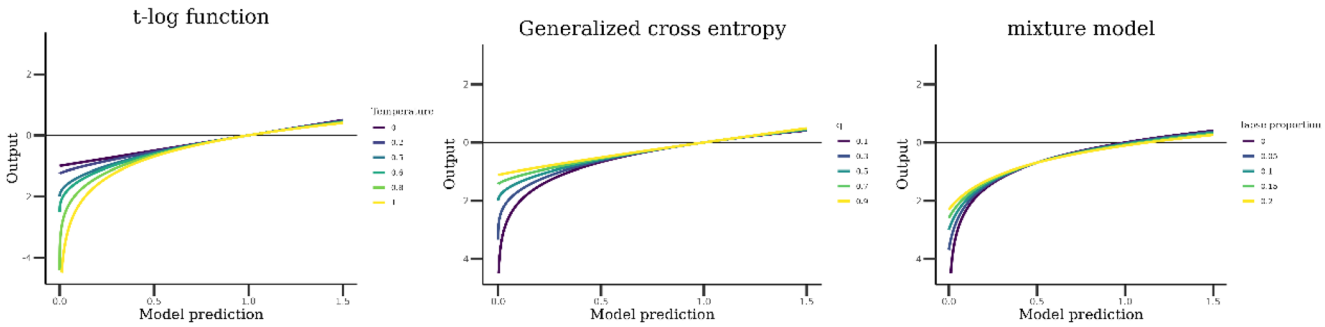

Fig. 2 t-log function, generalized cross entropy, and mixture model with different hyperparameters. The x axis represents the input of the function, whereas the y axis denotes the output of the function

outlier trials were sampled independently, such that the specific trials classified as outliers could differ across repetitions (e.g., trial 1, 5, and 20 might be outliers in one simulation, while trial 7 and 22 might be outliers in another simulation). For every agent and every simulation, the respective model was fitted (reinforcement learning or discounting model), resulting in a total of 20,000 model-fitting runs for each outlier level.

Outlier data were generated at five levels (0%, 5%, 10%, 15%, and 25%). At each level, outliers were introduced by replacing the model-generated choice with a random choice sampled from a Bernoulli distribution with rate of 0.5. Importantly, outlier trials at lower levels were nested within higher levels (e.g., the 5% outliers are included in the 10% condition).

Within the reinforcement learning paradigm, four trial types were included: go-to-reward, go-to-avoid-punishment, no-go-to-reward, and no-go-to-avoid-punishment. In the intertemporal choice task, each trial featured distinct values for the SS and LL options in terms of both reward magnitude and delay duration.

Following data generation, we fitted the true model to the synthetic data using several loss functions: the mixture model, MAE, GCE, t-log function, and LLF as the loss functions to fit the true model to the synthetic data to check the parameter recovery. For loss functions like GCE and t-log function, they required a pre-defined hyperparameter. Exploiting the hyperparameter space requires extensive computational resources which might not be realistic. Therefore, we adopted the suggestion in Zhang and Sabuncu (2018) and Amid et al. (2019) to set q in GCE to 0.2, whereas the temperature parameter in the t-log function was set to 0.8. To check the robustness of our analysis, we also include a condition where $q = 0.7$ and temperature = 0.7. The model fitting process utilized the optimization package DEoptim (Mullen et al., 2011) in the R programming language (Core Team, 2022).

Parameter Recovery Results

We examined the parameter identification of different loss functions through correlating the fitted parameters with the data-generating parameters. The simulation was conducted 100 times, yielding 100 correlation coefficients. The mean and standard deviation of these coefficients are detailed in the appendix, while their distribution is depicted in the violin plots of Figs. 3 and 4.

Previous work investigating statistical power of computational models has shown that parameter recovery is influenced by the sampled parameter distribution (Beeckmans et al., 2024). Broader distributions lead to a better recovery. Also, there is no consensus on an acceptable correlation value for parameter recovery. One study which performed a power analysis for computational models suggests a correlation of 0.7 as "significant" for parameter recovery (Beeckmans et al., 2023). Adopting this standard, most parameters in our reinforcement learning model fall short of this level of acceptability, even without outliers. Consequently, we relaxed this criterion to 0.5.

Generally, the hyperbolic discounting model outperforms the reinforcement learning model in parameter recovery, regardless of the outlier proportion. In the reinforcement learning model, all parameters' identification substantially decreases with increasing outlier proportions, especially so for the inverse temperature parameter governing choice consistency.

Without outlier trials, most loss functions perform comparably for all parameters, except for the go bias parameter, potentially due to its limited sample distribution standard deviation. Interestingly, for MAE, no parameter correlation exceeds 0.5, and the correlation for the inverse temperature parameter is nearly zero, indicating that MAE's theoretical outlier insensitivity does not translate into practical utility.

When the outlier proportion is low (5%), GCE with a hyperparameter of 0.2 performs best, with all parameter correlations above 0.5. The t-log function also performs well, though slightly less so. Traditional methods like MLE and the mixture model lag behind GCE, but their correlations are near 0.5. MLE's sensitivity to outliers explains its poorer performance, while the mixture model's underperformance for the inverse temperature parameter may be due to its high collinearity with the lapse parameter e.

At higher outlier proportions (10%, 15%, 25%), all loss functions see a marked decline in parameter recovery. The t-log with a hyperparameter of 0.7 performs best, but most correlations hover around 0.4, suggesting even outlier-insensitive functions struggle at higher outlier levels.

The results of the hyperbolic model in intertemporal choice tasks mirrors those of the reinforcement learning model, suggesting that the gained insights might be stable across different types of modelling scenarios. With a 10% outlier proportion, only the t-log and GCE with a hyperparameter of 0.2 maintain correlations above 0.5. At 15% and 25%, the t-log is slightly better for the inverse temperature, maintaining a 0.4 correlation, while others, including GCE, drop to 0.2. However, the discount rate k and the time perception parameter s remain acceptably identified (around a 0.6 correlation) even at a 25% outlier proportion.

The weight parameter e in the mixture model can be regarded as the estimation of the outlier proportion. We also checked the estimation of this parameter in the hyperbolic model and the reinforcement learning model. The result is presented in Table 1. In the reinforcement learning model,

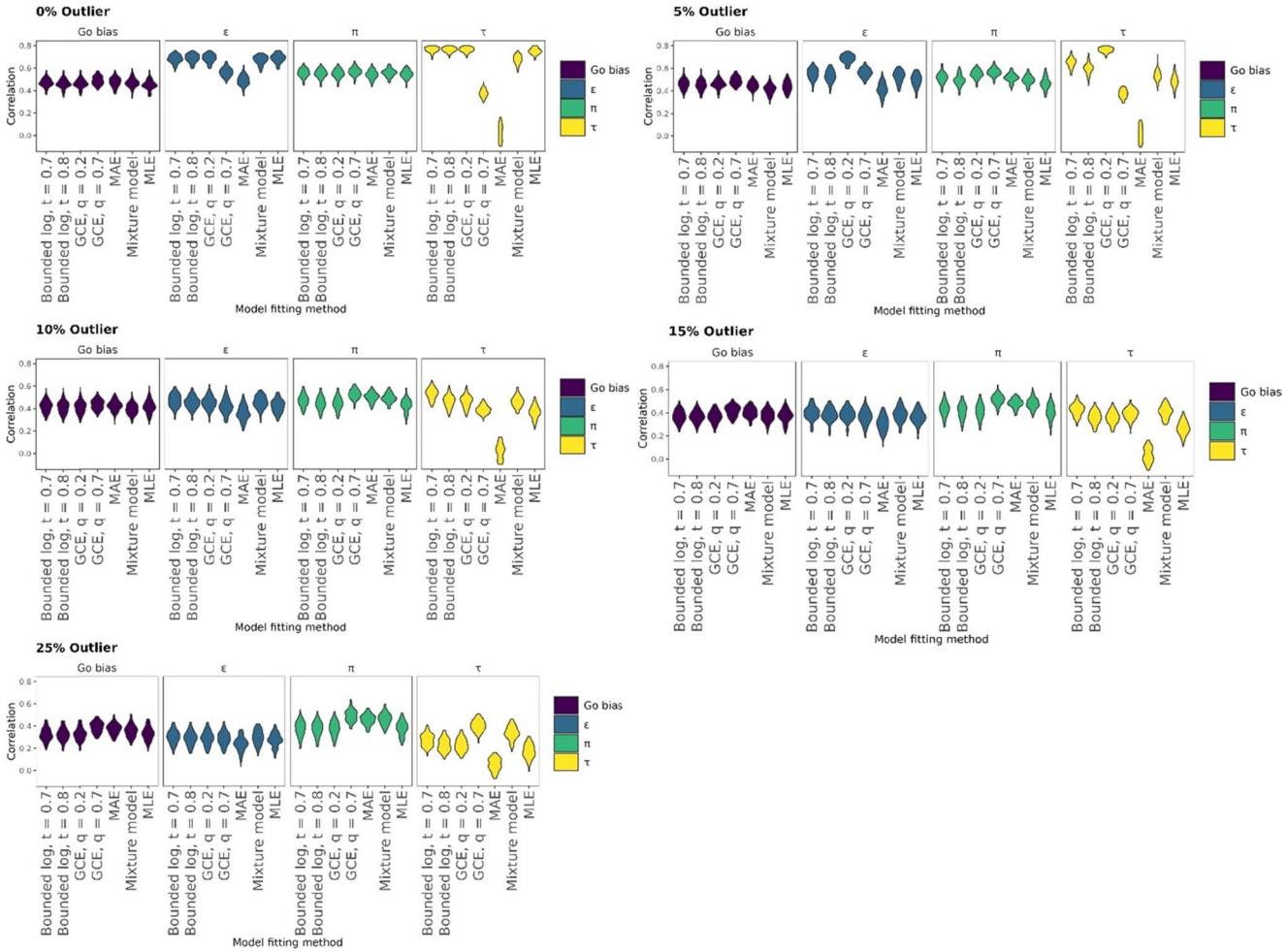

Fig. 3 Parameter recovery of the reinforcement learning model. The x-axis represents the various loss functions, while the y-axis depicts the correlation in parameter recovery

the outlier proportion is accurately estimated, while in the hyperbolic model it tends to be overestimated. This discrepancy might arise from the reinforcement learning model's consideration of outliers in future decisions, unlike the independent decisions of the hyperbolic model.

In conclusion, outlier-insensitive functions like t-log and GCE outperform MLE and the mixture model in our parameter recovery, demonstrating comparable or superior performance both in the absence and in the presence of outliers. At a 5% outlier proportion, GCE and t-log yield acceptable results for both the reinforcement and hyperbolic models.

Power Analysis Results

Cognitive modeling is increasingly used as a measurement tool, either to quantify experimental manipulations or to capture individual differences. In this role, the reliability and sensitivity of parameter estimates are critical, much like in traditional psychometric approaches. Recent studies in computational modeling of decision-making emphasize the importance of power analysis (Beeckmans et al., 2024; Gluth & Jarecki, 2019; van Ravenzwaaij et al., 2017). For example, van Ravenzwaaij et al. (2017) demonstrated that the EZ diffusion model has greater statistical power to detect the effect of experimental conditions on model parameters than the Ratcliff full diffusion model. In line with this, we applied a similar power analysis protocol as van Ravenzwaaij et al. (2017) to reinforcement learning models, using simulated data without outlier trials and with 5% outlier trials, to demonstrate how outlier trials can undermine the utility of cognitive models as measurement tools.

We simulated two groups of datasets to mimic a two-condition between-subjects experiment. Consistent with the parameter recovery analysis described in the previous section, we sampled 200 agents and ran 100 simulations to estimate statistical power. The first group served as the "baseline" group, with parameters drawn from the same distribution as in the parameter recovery analysis. For the second group, we allowed the learning rate to differ from

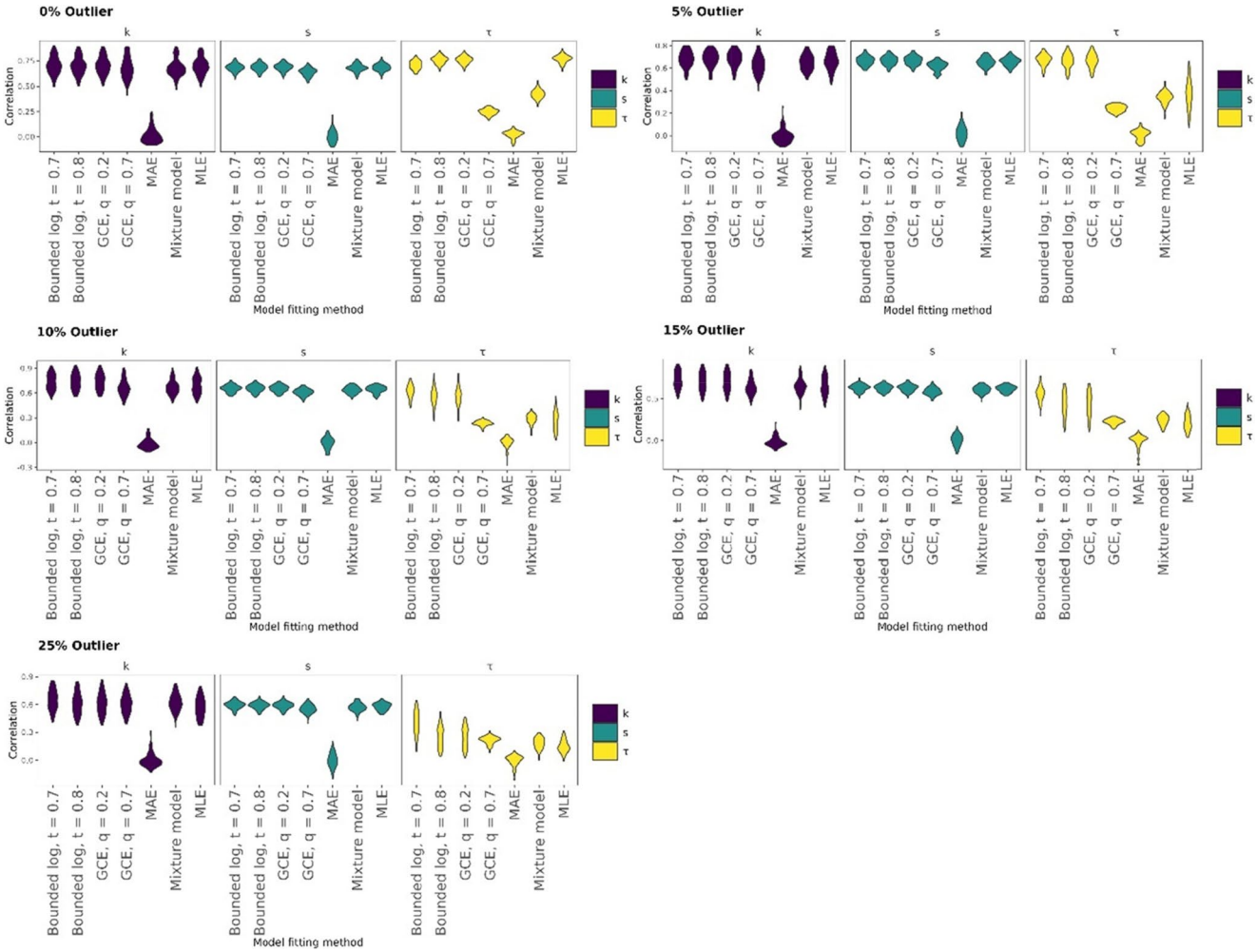

Fig. 4 Parameter recovery of the hyperbolic model. The x-axis represents the various loss functions, while the y-axis depicts the correlation in parameter recovery

Table 1 Mean Estimation of the e parameter from different outlier proportions. The standard deviation of the Estimation is shown between parentheses

Outlier proportion	RL Model	Hyperbolic Model
0%	0.012(0.026)	0.077(0.008)
5%	0.053(0.038)	0.118(0.009)
10%	0.097(0.054)	0.159(0.010)
15%	0.141(0.063)	0.200(0.010)
25%	0.227(0.083)	0.280(0.011)

those of the first group. The learning rate parameter is typically the primary focus in reinforcement learning models. Previous research has shown that the learning rate adapts to environmental uncertainty, and individual differences in this adaptive ability have been linked to psychiatric disorders such as anxiety and depression (Behrens et al., 2007; Browning et al., 2015; Piray et al., 2019). Specifically, we simulated three effect sizes—small, medium, and large. Depending on the effect size, the mean of the parameter of interest in Group 2 exceeded the Group 1 mean by 0.5, 0.8, or 1.0 within-group standard deviations, respectively. Individual participant values of learning rate for Group 2 were sampled from the following truncated normal distributions:

$$\in_{small} \sim \mathrm{TN}\,(0.6, 0.2, 1)$$

$$\in_{medium} \sim TN\,(0.66, 0.2, 0, 1)$$

$$\in_{large} \sim TN\,(0.8, 0.2, 0, 1)$$

We then fitted the RL model using the outlier-insensitive loss functions introduced in the previous sections and performed independent-samples *t*-tests between the estimated parameters of the two groups. Statistical power was defined as the proportion of simulations in which the *t*-test *p*-value was smaller than 0.001 rather than the conventional 0.05) in

Table 2 Power analysis results for the learning rate parameter (ϵ) under 0% outlier trials

Loss function	Low effect size	Medium effect size	High effect size
MLE	0.26	0.93	0.99
Mixture model	0.23	0.90	0.98
MAE	0.32	0.98	0.96
Bounded log t=0.7	0.42	0.99	1.00
Bounded log t=0.8	0.35	0.96	1.00
GCE q=0.2	0.34	0.96	1.00
GCE q=0.7	0.34	0.96	0.99

Table 3 Power analysis results for the learning rate parameter (ε) under 5% outlier trials

Loss function	Low effect size	Medium effect size	High effect size
MLE	0.00	0.09	0.04
Mixture model	0.23	0.48	0.60
MAE	0.17	0.71	0.69
Bounded log t=0.7	0.07	0.29	0.32
Bounded log t=0.8	0.15	0.59	0.67
GCE q=0.2	0.04	0.29	0.31
GCE q=0.7	0.07	0.37	0.53

order to reduce the probability of Type I errors and ensure that parameter differences were identified only under strong statistical evidence.

Table 2 shows the power analysis results for the learning rate parameter when no outlier trials were included. Across all loss functions, statistical power was high for medium and large effect sizes ($\geq$0.90), with most values approaching 1.00. For small effect sizes, however, notable differences emerged: MLE (0.26) and the Mixture model (0.23) showed relatively low power, whereas MAE (0.32), GCE (q=0.2/0.7; 0.34), and Bounded log (t=0.7; 0.42) achieved higher values, with Bounded log (t=0.7) performing best.

When 5% of the trials were replaced with outliers (Table 3), statistical power dropped substantially, especially for small and medium effect sizes. MLE almost completely failed (power$\approx$0), confirming its vulnerability to outliers. Among the robust loss functions, MAE showed comparatively higher power for the learning rate parameter (0.17, 0.71, and 0.69 for small, medium, and large effect sizes, respectively), followed by the Mixture model (0.23, 0.48, 0.60) and Bounded log (t=0.8; 0.15, 0.59, 0.67). In contrast, GCE (q=0.2/0.7) and Bounded log (t=0.7) yielded lower power under outlier conditions ($\leq$0.37 for small and medium effect sizes).

Taken together, these results indicate that while most loss functions perform adequately in the absence of outliers, even a small proportion of outlier trials can dramatically reduce statistical power. Although MAE appeared to preserve relatively higher sensitivity for the learning rate parameter under outlier conditions, it does not provide reliable fits for other key parameters such as the inverse temperature. This suggests that the choice of loss function should balance robustness to outliers with accurate estimation across all model parameters.

Comparison between Outlier-Insensitive Loss Function and Dynamic Noise Estimation

The purpose of the following section is to provide a simple, conceptual comparison between Dynamic Noise and outlier-insensitive loss functions. Specifically, we aim to illustrate that while Dynamic Noise can be an effective approach in reinforcement learning tasks, it is less suitable for non-learning tasks such as intertemporal choice. In the following section, we first introduce the Dynamic Noise framework and explain why it is not appropriate for non-learning tasks. We then provide a conceptual comparison by applying Dynamic Noise to the intertemporal choice task under the 5% outlier condition, to illustrate how its performance differs from that of outlier-insensitive loss functions.

Recent studies, such as Ashwood et al. (2022) and Li et al. (2024), have employed Bayesian HMM to estimate the hidden state of a participant (i.e., whether they are engaged in the task). Combing HMM and computational models can reduce the impact of outliers on parameter estimation in computational models. However, their approach differs in the usage of observation sequences: Ashwood et al.'s (2022) GLM-HMM approach utilizes the entire sequence of observations to deduce hidden states, in contrast to Li et al.'s (2024) Dynamic noise approach which relies exclusively on previous observations to forecast future hidden states. The usage of the entire sequence of observations of GLM-HMM requires an iterative expectation-maximization (EM) algorithm for parameter estimation, complicating the model-fitting process. Conversely, Dynamic noise is more straightforward, requiring only two additional parameters for model fitting.

One limitation of Dynamic noise is its potential unsuitability for non-learning tasks. For example, in a case where several outliers are present among the most recent 10 trials, Dynamic noise, relying heavily on prior observations, might erroneously forecast the next trial as an outlier, unlike GLM-HMM, which might predict otherwise by utilizing information about future events.

In the learning tasks, participants may disengage from the task and lose track of the candidate options' reward probability for a few trials, resulting in a high temporal correlation among outlier trials. In such a case, since the most recent history is full of outlier trials, it is plausible for an algorithm to predict that future trials will also very likely be outlier trials. However, in non-learning tasks, such as

intertemporal choice tasks, participants might initially follow a hyperbolic model, randomly select an option in subsequent trials, and then revert to the hyperbolic model. Hence, the temporal correlation of outlier trials in non-learning tasks might not be as pronounced as in learning tasks.

To evaluate Dynamic noise against outlier-insensitive loss functions, we applied the same model used for parameter recovery to the dataset of Ikink et al. (2023) employing three different methodologies: Dynamic noise, LLF, and the t-log function as the loss function. In the original paper that proposes the Dynamic noise, the author suggested Dynamic noise can have better or at least equivalent performance as MLE. Therefore, we used the Akaike Information Criterion (AIC) as the model comparison index (Wagenmakers & Farrell, 2004). Here, we briefly introduce the Dynamic Noise protocol proposed by Li et al. (2024), which we use in this section.

In a HMM, the hidden state at the current time step ($\boldsymbol{h_t}$) generates the current observation ($\boldsymbol{o_t}$), such that $\boldsymbol{p}\left(\boldsymbol{o_t}|\,\boldsymbol{h_t}\right)$. The hidden state at time t is in turn influenced only by the hidden state at the previous time step: $\boldsymbol{p}\left(\boldsymbol{h_t}|\,\boldsymbol{h_{t-1}}\right)$. The Dynamic Noise approach assumes two possible hidden states: (i) an engaged state (E), in which choices follow the decision policy of the underlying cognitive model (e.g., reinforcement learning or discounting model), and (ii) a random-choice state (R), in which choices are random. Transitions between these two states are governed by two transition probabilities: $T_E^R = p\left(R|\,E\right)$, the probability of moving from the engaged state to the random-choice state, and $T_R^E = p\left(E|\,R\right)$, the probability of moving from the random-choice state back to the engaged state.

At each trial, the probability of being in the engaged state is updated as:

$$p\left(E_t\right) = \sum\nolimits_i p\left(E_t|h_i\right) \times\ p\left(h_i\right) = p\left(E_{t-1}\right) \times\ \left(1 - T_E^R\right) + p\left(R_{t-1}\right) \times\ T_R^E$$

with $p\left(R_t\right) = 1 - p\left(E_t\right)$. Here, the probability that a participant is in the random-choice state, $p\left(R_t\right)$, is conceptually similar to theϵ parameter in the mixture model. The key difference is that $p\left(R_t\right)$ dynamically varies from trial to trial, whereas the ϵ parameter in the mixture model is fixed across trials.

The likelihood of the observed choice is then given by:

$$\begin{aligned} p\left(o_t\right) &= \sum\nolimits_i p\left(o_t|h_i\right) = p\left(o_t|E_t\right) \times\ p\left(E_t\right) + p\left(o_t|R_t\right) \times\ p\left(R_t\right) \\ &= p\left(o_t|E_{t-1}\right) \times\ p\left(E_t\right) + p\left(o_t|R_t\right) \times\ \left(1 - p\left(E_t\right)\right) \end{aligned}$$

In binay choice tasks such as the reinforcement learning and intertemporal choice tasks used in this study, $p\left(o_t|\,E_t\right)$ is given by the decision policy of the respective model, while $p\left(o_t|\,R_t\right)$ is the emission probability for the random state, in the binary choice task, it is 0.5.

Finally, the trial-by-trial posterior probability of being in the engaged state is updated via Bayes' rule:

$$p\left(E_t\right) = \frac{p\left(o_t|E_t\right) \times\ p\left(E_t\right)}{p\left(o_t\right)}$$

The results indicated that Dynamic noise underperforms relative to the other approaches (median AIC value for Dynamic noise is 55.45, median AIC value for MLE is 49.13, median AIC value for outlier-insensitive loss function is 51.51). Although Dynamic noise shows slightly better absolute model-fitting than other approaches, it introduces two additional parameters. These parameters cause significant penalties when using AIC as the model comparison criterion. Further, the t-log function also shows a worse AIC model weight than the LLF loss function. But it does not necessarily mean that an outlier-tolerant loss function like t-log function is worse than the LLF loss function, since the property to reduce the impact of outlier trials can be regarded as another form of penalty of the loss function. Outlier-insensitive loss functions reduce over-fitting, potentially resulting in worse absolute model-fitting but better out-of-sample prediction performance.

Additionally, when applying Dynamic noise to an intertemporal choice simulation with 5% outlier trials, it has an inferior parameter recovery compared to the LLF or t-log function, as shown in Table 4. To be specific, for the discounting rate parameter k and the time sensitivity parameter s, Dynamic noise's parameter recovery is slightly worse than for the LLF and t-log functions. Moreover, concerning the inverse temperature parameter τ, Dynamic noise is significantly worse than the other two approaches. These results reveal the unsuitability of Dynamic noise for non-learning tasks.

Discussion

Outlier data is a common phenomenon in behavioral data analysis. In reaction time studies, outlier handling methods are well-studied, yet in computational models of choice data, their impact on parameter estimation is less well understood and addressed. The current study introduces

Table 4 Dynamic noise parameter recovery for intertemporal choice task simulation with 5% outlier trials. The standard deviation of the Estimation is shown between parentheses

Parameter	Correlation
k	0.642 (0.086)
τ	0.341 (0.045)
s	0.649 (0.040)

several outlier-insensitive loss functions—commonly used in the computer vision field, as alternatives to the LLF function commonly used in computational model fitting. We simulated data using (a) the reinforcement learning model for go/no-go bandit tasks and (b) the hyperbolic model for intertemporal choice tasks, systematically examining the impact of outlier data, and comparing these loss functions in terms of their parameter recovery performance. Our results reveal that while the default loss function, LLF, is heavily influenced by outlier trials, outlier-insensitive functions like the t-log function and GCE (with appropriate hyperparameters) have superior parameter recovery. Our power analysis further suggests that even a small proportion of outlier trials (e.g., 5%) can potentially undermine the statistical power to detect condition differences, underscoring the importance of accounting for outliers when using cognitive models as measurement tools. These functions are effective even in the absence of outliers, matching the performance of LLF and traditional methods like the mixture model.

The outlier-insensitive loss functions perform relatively well when the outlier proportion is not that high (5% or 10%). Among all parameters, outlier-insensitive loss functions especially improve the parameter recovery of the inverse temperature parameter which governs the choice consistency. The inverse temperature parameter is always considered to be difficult to disentangle from the parameters that are responsible for computing the decision values, such as the risk aversion parameter in prospect theory (Krefeld-Schwalb et al., 2022; Stewart et al., 2018) or the learning rate parameter in reinforcement learning models (Ballard & McClure, 2019; Gershman, 2016). The presence of outliers worsens this problem whereas using outlier-insensitive loss functions can improve the parameter identification. This implies that achieving accurate interpretations regarding experimental manipulations or individual differences in choice consistency necessitates relatively high data quality.

When there are many outlier trials, for instance, 15% or 25% of the trials, these outlier-insensitive loss functions perform slightly better than LLF, but the overall performance is still not ideal. This implies that outlier-insensitive loss functions are useful in reducing the impact of outliers primarily when the proportion of outlier trials is relatively low.

Besides the outlier-insensitive loss functions, HMM can also estimate the latent state of a trial, indicating whether participants are engaged in the task or randomly selecting an option. Ashwood et al. (2022) utilized HMMs to infer the latent state using the full observation sequence. However, compared to this approach, outlier-insensitive loss functions are simpler to use and do not overly complicate the model. Another study, by Li et al. (2024), proposed a simplified HMM known as Dynamic noise. While Dynamic noise is straightforward to implement, it is less suitable for static tasks where computational models assume no learning occurs, where outlier-insensitive loss functions are more suitable in this case.

A critical point, which we highlight explicitly here, is how the hyperparameters of outlier-insensitive loss functions were chosen. When implementing these loss functions, hyperparameters must be pre-defined. Due to computational constraints, we limited our study to two hyperparameters per loss function. Hyperparameters not only affect the tolerance to outliers but also impact parameter estimation when outlier proportions are low. For example, GCE with a hyperparameter close to 1 would perform similar to MAE, which performs the worst among all loss functions tested in the current study. One potential way is to follow the strategy adopted in the current study, where we observed that (i) a hyperparameter equal to 0.2 for GCE provided a better parameter recovery than MLE and the mixture model when there are few outlier trials and (ii) equivalent performance as MLE and the mixture model when there are no outlier trials. Another approach to decide on the hyperparameter is to use the estimation of the proportion of outlier trials obtained from the mixture model. Despite the mixture model showing relatively worse performance in parameter recovery, it still provides a fairly accurate estimation of the proportion of outlier trials in the data. The process involves fitting the mixture model to the behavioral data firstly to determine the proportion of outlier trials. Once the proportion of outlier trials is established, parameter recovery is performed using different hyperparameters of the outlier-insensitive loss function. The hyperparameter that yields the best parameter recovery performance is then selected.

One limitation of the current study is that we only tested these outlier insensitive loss functions in non-hierarchical model-fitting approaches, i.e., without exploring hierarchical model fitting performance. This is relevant because hierarchical modelling approaches have been proved to be more stable and to have better performance in parameter recovery (Ahn et al., 2017; Huys et al., 2011). Furthermore, compared to the non-hierarchical model-fitting, hierarchical model fitting has a higher tolerance of outlier trials due to the so called shrinkage effect (Gelman et al., 2013). It is also possible to replace the LLF with the outlier insensitive loss functions in the hierarchical model-fitting procedure. Future work could investigate performance of different loss functions in a hierarchical modelling context.

While our study focused on outlier-insensitive loss functions as a strategy to mitigate the impact of noisy trials, other approaches may also enhance robustness in cognitive modeling. For instance, model averaging explicitly addresses model uncertainty by combining evidence across multiple candidate models rather than relying on a single specification. Boehm et al. (2023) demonstrated this approach in

hierarchical diffusion models using inclusion Bayes factors, showing that model averaging can yield more stable parameter inferences.

In conclusion, the current study highlights the substantial impact of outlier data on parameter identification. The utilization of outlier-insensitive loss functions such as GCE and the t-log function, instead of the standard LLF, significantly improves parameter identification, especially when dealing with a relatively small proportion of outlier data. Unlike other approaches that directly estimate hidden states for each trial, outlier-insensitive loss functions are user-friendly and adaptable to various computational models focused on choice modeling. Given the increasing use of online and mobile phone tasks for individual differences studies where parameter identification poses a significant challenge (Jangraw et al., 2023; Steyvers et al., 2019), future research should pay more attention to the handling of outlier data in computational modeling.

Appendix

We present the detailed parameter correlations for the reinforcement learning model and the hyperbolic model in the following series of tables. Note that because the simulation and model-fitting were performed 100 times for each agent, we not only present the mean of the correlations between the true and fitted parameter but also show the standard deviations of these correlations in parantheses.

Reinforcement Learning Model

Table 5 0% outlier

	ε	π	τ	$Go\ bias$
MLE	0.685 (0.038)	0.542(0.038)	0.754(0.031)	0.450(0.045)
Mixture model	0.669 (0.039)	0.556(0.035)	0.674(0.040)	0.467(0.045)
t-log, temp = 0.7	0.678(0.036)	0.555(0.038)	0.763(0.026)	0.470(0.041)
t-log, temp = 0.8	0.685(0.037)	0.549(0.038)	0.771(0.027)	0.462(0.043)
GCE, $q = 0.2$	0.685(0.037)	0.549(0.038)	0.771(0.027)	0.462(0.042)
GCE, $q = 0.7$	0.678(0.040)	0.555(0.035)	0.766(0.038)	0.470(0.039)
MAE	0.490(0.047)	0.540(0.039)	-0.043(0.110)	0.473(0.043)

Table 6 5% outlier

	ε	π	τ	$Go\ bias$
MLE	0.493 (0.056)	0.469(0.053)	0.489(0.058)	0.441(0.049)
Mixture model	0.526(0.052)	0.497(0.040)	0.534(0.051)	0.415(0.046)
t-log, temp = 0.7	0.551(0.051)	0.518(0.047)	0.649(0.043)	0.460(0.050)
t-log, temp = 0.8	0.532(0.053)	0.494(0.050)	0.598(0.051)	0.453(0.051)
GCE, $q = 0.2$	0.685(0.037)	0.549(0.039)	0.771(0.027)	0.462(0.042)
GCE, $q = 0.7$	0.562(0.040)	0.568(0.036)	0.374(0.037)	0.489(0.038)
MAE	0.416(0.061)	0.516(0.039)	0.006(0.077)	0.447(0.043)

Table 7 10% outlier

	ε	π	τ	$Go\ bias$
MLE	0.420 (0.057)	0.427(0.060)	0.366(0.057)	0.413(0.059)
Mixture model	0.453 (0.056)	0.492(0.041)	0.460(0.050)	0.394(0.050)
t-log, temp = 0.7	0.468(0.059)	0.469(0.054)	0.530(0.050)	0.419(0.053)
t-log, temp = 0.8	0.450(0.055)	0.444(0.054)	0.462(0.054)	0.411 (0.056)
GCE, $q = 0.2$	0.450(0.058)	0.443(0.058)	0.461(0.055)	0.411(0.056)
GCE, $q = 0.7$	0.416 (0.067)	0.525(0.042)	0.389(0.039)	0.433(0.047)
MAE	0.359(0.067)	0.500(0.041)	0.013(0.072)	0.423(0.048)

Table 8 15% outlier

	ε	π	τ	$Go\ bias$
MLE	0.365(0.060)	0.412(0.066)	0.271(0.059)	0.374(0.059)
Mixture model	0.387(0.056)	0.482(0.041)	0.411(0.050)	0.378(0.050)
t-log, temp = 0.7	0.400(0.059)	0.431(0.062)	0.427(0.052)	0.370(0.053)
t-log, temp = 0.8	0.385(0.058)	0.416(0.065)	0.362(0.055)	0.367(0.052)
GCE, $q = 0.2$	0.387(0.058)	0.417(0.058)	0.361(0.055)	0.367(0.056)
GCE, $q = 0.7$	0.367(0.065)	0.516(0.045)	0.391(0.046)	0.418(0.046)
MAE	0.309(0.067)	0.483(0.040)	0.033(0.068)	0.405(0.046)

Table 9 25% outlier

	ε	π	τ	$Go\ bias$
MLE	0.272(0.063)	0.376(0.063)	0.170(0.064)	0.328(0.061)
Mixture model	0.295(0.066)	0.460(0.041)	0.336(0.061)	0.355(0.057)
t-log, temp = 0.7	0.300(0.061)	0.374(0.067)	0.271(0.056)	0.324(0.058)
t-log, temp = 0.8	0.286(0.061)	0.372(0.066)	0.228(0.059)	0.324(0.059)
GCE, $q = 0.2$	0.288(0.062)	0.417(0.058)	0.373(0.067)	0.324(0.060)
GCE, $q = 0.7$	0.285(0.061)	0.489(0.053)	0.387(0.049)	0.387(0.049)
MAE	0.232(0.066)	0.454(0.047)	0.045(0.047)	0.372(0.049)

Hyperbolic Model

Table 10 0% outlier

	k	τ	s
MLE	0.704(0.091)	0.769(0.039)	0.685(0.036)
Mixture model	0.680(0.089)	0.423(0.044)	0.679(0.039)
t-log, temp = 0.7	0.710(0.093)	0.708(0.042)	0.683(0.038)
t-log, temp = 0.8	0.707(0.093)	0.757(0.039)	0.685(0.038)
GCE, $q = 0.2$	0.709(0.092)	0.757(0.038)	0.686(0.038)
GCE, $q = 0.7$	0.682(0.107)	0.248(0.030)	0.637(0.041)
MAE	0.008(0.077)	-0.008(0.078)	-0.009(0.072)

Table 11 5% outlier

	k	τ	s
MLE	0.706(0.109)	0.372(0.126)	0.659(0.126)
Mixture model	0.666(0.083)	0.340(0.053)	0.655(0.042)
t-log, temp = 0.7	0.658 (0.106)	0.644(0.030)	0.619(0.042)
t-log, temp = 0.8	0.745(0.104)	0.672(0.071)	0.669(0.039)
GCE, $q = 0.2$	0.745(0.106)	0.672(0.068)	0.669(0.039)
GCE, $q = 0.7$	0.658(0.106)	0.243(0.030)	0.619(0.042)
MAE	-0.001(0.083)	0.005(0.053)	0.005(0.042)

Table 12 10% outlier

	k	τ	s
MLE	0.688(0.102)	0.266(0.116)	0.642(0.042)
Mixture model	0.667(0.088)	0.287(0.059)	0.636(0.042)
t-log, temp = 0.7	0.734(0.096)	0.622(0.079)	0.653(0.043)
t-log, temp = 0.8	0.739(0.097)	0.555(0.114)	0.653(0.044)
GCE, $q = 0.2$	0.738(0.096)	0.554(0.114)	0.652(0.041)
GCE, $q = 0.7$	0.663(0.100)	0.235(0.031)	0.609(0.042)
MAE	-0.014(0.063)	0.002(0.064)	-0.003(0.069)

Table 13 15% outlier

	k	τ	s
MLE	0.653 (0.115)	0.216(0.093)	0.623(0.038)
Mixture model	0.652(0.095)	0.247(0.058)	0.614(0.044)
t-log, temp = 0.7	0.722(0.103)	0.565(0.089)	0.635(0.043)
t-log, temp = 0.8	0.700(0.111)	0.438(0.142)	0.632(0.041)
GCE, $q = 0.2$	0.700(0.111)	0.437(0.143)	0.632(0.042)
GCE, $q = 0.7$	0.638(0.093)	0.229(0.034)	0.591(0.045)
MAE	-0.013(0.058)	-0.005(0.081)	0.005(0.070)

Table 14 25% outlier

	k	τ	s
MLE	0.573(0.107)	0.150(0.066)	0.573(0.040)
Mixture model	0.622(0.092)	0.185(0.059)	0.571(0.045)
t-log, temp = 0.7	0.650(0.110)	0.402(0.128)	0.593(0.041)
t-log, temp = 0.8	0.610(0.116)	0.252(0.114)	0.589(0.038)
GCE, $q = 0.2$	0.612(0.115)	0.250(0.115)	0.589(0.039)
GCE, $q = 0.7$	0.605(0.100)	0.218(0.041)	0.553(0.046)
MAE	0.003(0.092)	-0.001(0.059)	-0.001(0.045)

Table 15 We provide detailed information about the stimuli used in Ikink et al. (2023) in the following table

LLamount	SSamount	LLtime	SStime	count
19	16	42	14	2
36	33	3	0	2
67	56	17	14	2
21	16	17	14	2
73	56	42	14	2
43	42	17	14	2
43	33	28	14	2
42	41	14	0	2
41	34	3	0	2
57	44	14	0	2
55	42	3	0	2
33	22	28	14	2
32	29	28	14	2
32	29	17	14	2
29	28	28	0	2
21	20	28	14	2
23	19	14	0	2
83	69	28	14	1
68	45	42	14	1
57	52	42	14	1
58	52	42	14	1
67	45	42	14	1
74	62	28	0	1
84	70	28	14	1
81	54	3	0	1
84	56	14	0	1
80	54	3	0	1
75	62	28	0	1
46	45	3	0	1
53	36	17	14	1
53	35	17	14	1
51	46	14	0	1
50	46	14	0	1
47	46	3	0	1
45	44	42	14	1
44	43	42	14	1
41	32	28	0	1
40	31	28	0	1
37	25	28	0	1
36	24	28	0	1
27	25	28	0	1
26	24	28	0	1
85	56	14	0	1

Author Contributions M.G. designed and analyzed the data. All authors reviewed the manuscript.

Data Availability No datasets were generated or analysed during the current study.

Code Availability The code used in this study is available upon request.

Declarations

Ethics Approval This study only involves the simulation and therefore does not require ethics approval.

Financial Support M.G. thanks the Guangzhou Elite Project for a fellowship.

Consent for Publication Not applicable.

Consent to participate Not applicable.

Competing interests The authors declare no competing interests.

References

Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational Psychiatry, 1*, 24–57. https://doi.org/10.1162/CPSY_a_00002

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Amid, E., Warmuth, M. K., Anil, R., & Koren, T. (2019). Robust bi-tempered logistic loss based on bregman divergences. *Advances in neural information processing systems*, *32*. https://proceedings.neurips.cc/paper_files/paper/2019/file/8cd7775f9129da8b5bf787a063d8426e-Paper.pdf

Ashwood, Z. C., Roy, N. A., Stone, I. R., Brain, I., Urai, L., Churchland, A. E., Pouget, A. K., A., & Pillow, J. W. (2022). Mice alternate between discrete strategies during perceptual decision-making. *Nature Neuroscience*, *25*(2), 201–212. https://doi.org/10.1038/s41593-021-01007-z

Ballard, I. C., & McClure, S. M. (2019). Joint modeling of reaction times and choice improves parameter identifiability in reinforcement learning models. *Journal of Neuroscience Methods, 317*, 37–44. https://doi.org/10.1016/j.jneumeth.2019.01.006

Beeckmans, M., Huycke, P., Verguts, T., & Verbeke, P. (2023). *How much data do we need to estimate computational models of decision making?* The Compass toolbox.

Beeckmans, M., Huycke, P., Verguts, T., & Verbeke, P. (2024). How much data do we need to estimate computational models of decision-making? The COMPASS toolbox. *Behavior Research Methods*, *56*(3), 2537–2548. https://doi.org/10.3758/s13428-023-02165-7

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

Boehm, U., Evans, N. J., Gronau, Q. F., Matzke, D., Wagenmakers, E. J., & Heathcote, A. J. (2023). Inclusion Bayes factors for mixed hierarchical diffusion decision models. *Psychological Methods*, *Advance online publication*. https://doi.org/10.1037/met0000582

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590–596. https://doi.org/10.1038/nn.3961

Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A

behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience, 35*(7), 1024–1035. https://doi.org/10.1111/j.1460-9568.2011.07980.x

Collins, A. G. E., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. *The Journal of Neuroscience, 34*(41), 13747–13756. https://doi.org/10.1523/jneurosci.0989-14.2014

Core Team, R. (2022). R. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. https://www.R-project.org/

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In *Decision making, affect, learning: Attention performance XXIII* (Vol. 23).

Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural Networks, 19*(8), 1153–1160. https://doi.org/10.1016/j.neunet.2006.03.002

Ding, N., & Vishwanathan, S. (2010). t-Logistic regression. *Advances in neural information processing systems*, *23*. https://proceedings.neurips.cc/paper_files/paper/2010/file/96da2f590cd7246bbde0051047b0d6f7-Paper.pdf

Farrell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*. Cambridge University Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis. 3rd edition*. Chapman and Hall/CRC.

Geng, H., Chen, J., Chuan-Peng, H., Jin, J., Chan, R. C. K., Li, Y., Hu, X., Zhang, R. Y., & Zhang, L. (2022). Promoting computational psychiatry in China. *Nature Human Behaviour, 6*(5), 615–617. https://doi.org/10.1038/s41562-022-01328-4

Gershman, S. J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology, 71*, 1–6. https://doi.org/10.1016/j.jmp.2016.01.006

Gershman, S. J., & Bhui, R. (2020). Rationally inattentive intertemporal choice. *Nature Communications, 11*(1), Article 3365. https://doi.org/10.1038/s41467-020-16852-y

Ghosh, A., Kumar, H., & Sastry, P. S. (2017). Robust loss functions under label noise for deep neural networks. Proceedings of the AAAI conference on artificial intelligence.

Gluth, S., & Jarecki, J. B. (2019). On the importance of power analyses for cognitive modeling. *Computational Brain & Behavior*, *2*(3–4), 266–270. https://doi.org/10.1007/s42113-019-00039-w

Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., & Dolan, R. J. (2012a). Action controls dopaminergic enhancement of reward representations. *Proceedings of the National Academy of Sciences of the United States of America, 109*(19), 7511–7516. https://doi.org/10.1073/pnas.1202229109

Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012b). Go and no-go learning in reward and punishment: Interactions between affect and effect. *Neuroimage, 62*(1), 154–166. https://doi.org/10.1016/j.neuroimage.2012.04.024

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus Valence in decision making. *Trends in Cognitive Sciences*, *18*(4), 194–202. https://doi.org/10.1016/j.tics.2014.01.003

Guo, M., Ikink, I., Roelofs, K., & Figner, B. (2025). Ambiguity preferences in intertemporal and risky choice: A large–scale study using drift–diffusion modelling. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-025-02709-2

Harless, D. W., & Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. *Econometrica, 62*(6), 1251–1289. https://doi.org/10.2307/2951749

Huys, Q. J., Cools, R., Golzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the roles of approach, activation and valence in instrumental and Pavlovian responding. *PLoS Computational Biology, 7*(4), Article e1002028. https://doi.org/10.1371/journal.pcbi.1002028

Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience, 19*(3), 404–413. https://doi.org/10.1038/nn.4238

Ikink, I., van Duijvenvoorde, A. C. K., Huizenga, H., Roelofs, K., & Figner, B. (2023). Age differences in intertemporal choice among children, adolescents, and adults. *Journal of Experimental Child Psychology, 233*, Article 105691. https://doi.org/10.1016/j.jecp.2023.105691

Jangraw, D. C., Keren, H., Sun, H., Bedder, R. L., Rutledge, R. B., Pereira, F., Thomas, A. G., Pine, D. S., Zheng, C., Nielson, D. M., & Stringaris, A. (2023). A highly replicable decline in mood during rest and simple tasks. *Nature Human Behaviour, 7*(4), 596–610. https://doi.org/10.1038/s41562-023-01519-7

Jiang, T., & Dai, J. (2023). Cognitive load enhances patience rather than impulsivity. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02403-1

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal Of The American Statistical Association, 90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Krefeld-Schwalb, A., Pachur, T., & Scheibehenne, B. (2022). Structural parameter interdependencies in computational models of cognition. *Psychological Review, 129*(2), 313–339. https://doi.org/10.1037/rev0000285

Kucharský, Š., Tran, N. H., Veldkamp, K., Raijmakers, M., & Visser, I. (2021). Hidden Markov models of evidence accumulation in speeded decision tasks. *Computational Brain & Behavior*, *4*(4), 416–441. https://doi.org/10.1007/s42113-021-00115-0

Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods, 49*(2), 513–537. https://doi.org/10.3758/s13428-016-0740-2

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal Of Experimental Social Psychology, 49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Li, J. J., Shi, C., Li, L., & A. Collins (2024). Dynamic noise estimation: A generalized method for modeling noise fluctuations in decision-making. *Journal of Mathematical Psychology*, *2023.2006*(2019.545524). https://doi.org/10.1016/j.jmp.2024.102842

Luhmann, C. C. (2013). Discounting of delayed rewards is not hyperbolic. *Journal of Experimental Psychology: Learning Memory & Cognition*, *39*(4), 1274–1279. https://doi.org/10.1037/a0031170

Mazur, J. E., Mazur, J., & Nevin, J. (1987). An adjusting procedure for studying delayed reinforcement. In *Quantitative Analysis of Behavior* (pp. 55–73). Hillsdale, NJ (1987).

Molter, F., Thomas, A. W., Heekeren, H. R., & Mohr, P. N. C. (2019). GLAMbox: A Python toolbox for investigating the association between gaze allocation and decision behaviour. *PLoS One, 14*(12), Article e0226428. https://doi.org/10.1371/journal.pone.0226428

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences, 16*(1), 72–80. https://doi.org/10.1016/j.tics.2011.11.018

Mullen, K., Ardia, D., Gil, D. L., Windover, D., & Cline, J. (2011). DEoptim: An R package for global optimization by differential evolution. *Journal of Statistical Software, 40*(6), 1–26.

Nassar, M. R., & Frank, M. J. (2016). Taming the beast: Extracting generalizable knowledge from computational models of cognition. *Current Opinion in Behavioral Sciences, 11*, 49–54. https://doi.org/10.1016/j.cobeha.2016.04.003

Naudts, J. (2002). Deformed exponentials and logarithms in generalized thermostatistics. *Physica A: Statistical Mechanics and Its Applications, 316*(1–4), 323–334.

Olschewski, S., Rieskamp, J., & Scheibehenne, B. (2018). Taxing cognitive capacities reduces choice consistency rather than preference: A model-based test. *Journal Of Experimental Psychology: General, 147*(4), 462–484. https://doi.org/10.1037/xge0000403

Padoa-Schioppa, C. (2022). Logistic analysis of choice data: A primer. *Neuron, 110*(10), 1615–1630. https://doi.org/10.1016/j.neuron.2022.03.002

Pawitan, Y. (2001). *In all likelihood: Statistical modelling and inference using likelihood*. Oxford University Press.

Peters, J., & D'Esposito, M. (2020). The drift diffusion model as the choice rule in inter-temporal and risky choice: A case study in medial orbitofrontal cortex lesion patients and controls. *PLoS Computational Biology, 16*(4), Article e1007615. https://doi.org/10.1371/journal.pcbi.1007615

Piray, P., Ly, V., Roelofs, K., Cools, R., & Toni, I. (2019). Emotionally aversive cues suppress neural systems underlying optimal learning in socially anxious individuals. *The Journal of Neuroscience, 39*(8), 1445–1456. https://doi.org/10.1523/jneurosci.1394-18.2018

Pisupati, S., Chartarifsky-Lynn, L., Khanal, A., & Churchland, A. K. (2021). Lapses in perceptual decisions reflect exploration. *eLife*. https://doi.org/10.7554/eLife.55490

Rachlin, H. (2006). Notes on discounting. *Journal of the Experimental Analysis of Behavior, 85*(3), 425–435. https://doi.org/10.1901/jeab.2006.85-05

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510. https://doi.org/10.1037/0033-2909.114.3.510

Ratcliff, R., & Kang, I. (2021). Qualitative speed-accuracy tradeoff effects can be explained by a diffusion/fast-guess mixture model. *Scientific Reports, 11*(1), Article 15169. https://doi.org/10.1038/s41598-021-94451-7

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. https://doi.org/10.3758/BF03196302

Shahar, N., Hauser, T. U., Moutoussis, M., Moran, R., Keramati, M., consortium, N., & Dolan, R. J. (2019). Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Computational Biology, 15*(2), Article e1006803. https://doi.org/10.1371/journal.pcbi.1006803

Song, H., Kim, M., Park, D., Shin, Y., & Lee, J. G. (2022). Learning from noisy labels with deep neural networks: A survey. *IEEE Transactions on Neural Networks and Learning Systems*.

Stewart, N., Scheibehenne, B., & Pachur, T. (2018). Psychological parameters have units: A bug fix for stochastic prospect theory and other decision models.

Steyvers, M., Hawkins, G. E., Karayanidis, F., & Brown, S. D. (2019). A large-scale analysis of task switching practice effects across the lifespan. *Proceedings of the National Academy of Sciences of the United States of America, 116*(36), 17735–17740. https://doi.org/10.1073/pnas.1906788116

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Swart, J. C., Frobose, M. I., Cook, J. L., Geurts, D. E., Frank, M. J., Cools, R., & den Ouden, H. E. (2017). Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *Elife*, *6*. https://doi.org/10.7554/eLife.22169

van Ravenzwaaij, D., Donkin, C., & Vandekerckhove, J. (2017). The EZ diffusion model provides a powerful test of simple empirical effects. *Psychonomic Bulletin & Review*, *24*(2), 547–556. https://doi.org/10.3758/s13423-016-1081-y

Venditto, S. J. C., Miller, K. J., Brody, C. D., & Daw, N. D. (2024). Dynamic reinforcement learning reveals time-dependent shifts in strategy during reward learning. *Elife*. https://doi.org/10.7554/elife.97612.2

Vincent, B. T. (2016). Hierarchical bayesian estimation and hypothesis testing for delay discounting tasks. *Behavior Research Methods, 48*(4), 1608–1620. https://doi.org/10.3758/s13428-015-0672-2

Vincent, B. T., & Stewart, N. (2020). The case of muddled units in temporal discounting. *Cognition*. https://doi.org/10.1016/j.cognition.2020.104203

Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, *11*(1), 192–196. https://doi.org/10.3758/BF03206482

Wang, Y., Ma, X., Chen, Z., Luo, Y., Yi, J., & Bailey, J. (2019). Symmetric cross entropy for robust learning with noisy labels. Proceedings of the IEEE/CVF international conference on computer vision.

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine Learning, 8*(3–4), 279–292. https://doi.org/10.1007/BF00992698

Wiecki, T., Sofer, I., & Frank, M. (2013). HDDM: Hierarchical bayesian Estimation of the Drift-Diffusion model in python. *Frontiers in Human Neuroscience*, *7*(14). https://doi.org/10.3389/fninf.2013.00014

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*. https://doi.org/10.7554/eLife.49547

Zhang, Z., & Sabuncu, M. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. *Advances in neural information processing systems*, *31*. https://proceedings.neurips.cc/paper_files/paper/2018/file/f2925f97bc13ad2852a7a551802feea0-Paper.pdf

Zhang, L., Lengersdorff, L., Mikus, N., Glascher, J., & Lamm, C. (2020). Using reinforcement learning models in social neuroscience: Frameworks, pitfalls and suggestions of best practices. *Social Cognitive and Affective Neuroscience, 15*(6), 695–707. https://doi.org/10.1093/scan/nsaa089