Title: The Exponential Weight Updating Model: A novel computational model for the Balloon Analogue Risk Task

Authors:

Harhim Park[1], Jaeyeong Yang[1], Jasmin Vassileva[2, 3], Woo-Young Ahn[1]

[1]Department of Psychology, Seoul National University, Seoul, Korea

[2]Department of Psychiatry, Virginia Commonwealth University, Virginia, United States of America

[3]Institute for Drug and Alcohol Studies, Virginia Commonwealth University, Virginia, United States of America

Corresponding author:

Woo-Young Ahn, Ph.D.

Department of Psychology

Seoul National University

Seoul, Korea 08826

Tel: +82-2-880-2538, Fax: +82-2-877-6428. E-mail: wahn55@snu.ac.kr

Keywords: Balloon Analogue Risk Task, risk-taking, Hierarchical Bayesian Analysis, computational modeling, substance use

Number of words in the abstract: 134

Number of words in the main text (including references): 7,144

Number of Figures/Tables in the main text: 6 / 1

Number of Figures/Tables in supplemental online material: 2 / 1

Number of References: 39

Conflict of Interest: The authors declare no competing financial interest.

Abstract

The Balloon Analogue Risk Task (BART) is a popular task used to measure risk-taking behavior. To identify cognitive processes associated with choice behavior on the BART, a few computational models have been proposed. However, the existing models are either too simplistic or fail to show good parameter recovery performance. Here, we propose a novel computational model, the exponential weight updating (EU) model that addresses the limitations of existing models. By using multiple model comparison methods including post-hoc model fits criterion and parameter recovery, we showed that the EU model outperforms the existing models. In addition, we applied the EU model to BART data from healthy controls and substance-using populations (patients with past opiate and stimulant dependence). The modeling results suggest that heroin users show increased risk preference and reduced loss aversion than other groups.

1. Introduction

Computational modeling of cognitive tasks has been widely used to address the limitations of behavioral measures, with which it is often hard to identify underlying cognitive processes (Ahn, Busemeyer, Wagenmakers, & Stout, 2008; Ahn, Dai, Vassileva, Busemeyer, & Stout, 2016; Ahn et al., 2014; Busemeyer & Stout, 2002; Chung, Christopoulos, King-Casas, Ball, & Chiu, 2015; Collins, Ciullo, Frank, & Badre, 2017; Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Guitart-Masip et al., 2012; Haines, Vassileva, & Ahn, 2018; Ratcliff, 1978; Sokol-Hessner et al., 2009; van Ravenzwaaij, Dutilh, & Wagenmakers, 2011; Wallsten, Pleskac, & Lejuez, 2005; Worthy & Maddox, 2014; Wunderlich, Smittenaar, & Dolan, 2012). For example, research examining a series of models for the Iowa Gambling Task (IGT) has accounted for the various patterns in behavioral data, provided short- and long-term prediction and good parameter recovery (Ahn et al., 2008; Ahn et al., 2014; Busemeyer & Stout, 2002; Haines et al., 2018; Worthy & Maddox, 2014), and revealed decision-making deficits in several clinical populations that were not detected by traditional performance indices (see Ahn et al., 2016 for a review).

Like the IGT, the Balloon Analogue Risk Task (BART) (Lejuez et al., 2002) was originally designed for the clinical purpose of measuring risk-taking tendencies and the identification of individuals who are prone to take risks (Lejuez et al., 2002; Lejuez, Simmons, Aklin, Daughters, & Dvir, 2004), but its scope has expanded to other areas of psychology and cognitive science (van Ravenzwaaij et al., 2011; Wallsten et al., 2005). The BART has been shown to identify high risk-taking individuals (Aklin, Lejuez, Zvolensky, Kahler, & Gwadz, 2005; Lejuez et al., 2002; Lejuez et al., 2004). Specifically, the behavioral performance of the BART is significantly correlated with self-report measures of risk-related constructs such as impulsivity and sensation seeking (Lejuez et al., 2002), and past real-world risky behaviors such as drug use and risky sex (Lejuez et al., 2004). These results presumably reflect at least two features of the BART which make the task similar to real-world situations: First, each trial of the BART includes sequential risk-taking choices and terminates when the participant does not want to take the risk any more or encounters a certain condition that makes it impossible to proceed. Second, the riskiness of the risk-taking choices within a trial increases each time the participant makes a risky choice. Specifically, as the participant makes a risky choice, the loss amount gradually increases whereas the reward amount remains constant or decreases. Many actual risky behaviors involve these features. For example, patients with substance use disorders continue to take the drug even in the face of negative consequences until they are satisfied with it or they cannot obtain the drug because of external factors such as lack of availability and insufficient money. Also, patients often start with smaller amounts of the drug, but gradually increase their dosage as their tolerance increases. As their tolerance increases, they must (and are likely to) take more risks to get the same amount of reward. The BART has the distinct advantage of effectively illustrating this kind of real-world situation in a laboratory setting.

To quantitatively analyze the underlying cognitive processes on the BART, previous studies have proposed two computational models: a four-parameter model (Wallsten et al., 2005) and a two-parameter model (van Ravenzwaaij et al., 2011). The first model, the four-parameter model, was proposed as a winning model through comparing several computational models, and there exist correlations between the model parameters (reflecting psychological constructs) and the frequencies of past real-world risky behaviors such as gambling, substance use, and unprotected sex. However, this model failed to reproduce accurate parameter values in parameter recovery (Heathcote, Brown, & Wagenmakers, 2015; van Ravenzwaaij et al., 2011), which indicates how reliable the parameter values are. Good parameter recovery performance is crucial for interpreting results based on the model parameters (van Ravenzwaaij & Oberauer, 2009; Wagenmakers, Van Der Maas, & Grasman, 2007). Hence, poor parameter recovery performance is a critical limitation of the four-parameter model. Additionally, the learning equation of the four-parameter model is hard to interpret within a general reinforcement learning (RL) framework.

The two-parameter model was proposed to overcome this limitation of the four-parameter model (van Ravenzwaaij et al., 2011). To develop a model which shows good parameter recovery, the authors simplified the original model by removing parameters that do not exhibit good recovery. In fact, the two-parameter model is a nested model of the four-parameter model. As a result of simplification, the two-parameter model succeeded in recovering accurate parameter values (van Ravenzwaaij et al., 2011). However, the two-parameter model also has a critical limitation; it is based on a strict assumption that participants do not learn during the BART. The assumption is unrealistic unless the researcher tells the participant the actual exploding probability of virtual balloons before starting the experiment. Furthermore, because this issue is related to the task design, the two-parameter model is not applicable to the original task design. The original task design has the advantage that it similarly illustrates real-world situations, which means If we modify the task design to apply the two-parameter model, we lose the advantage. Thus, there is a need to build a new model, which shows good parameter recovery and is applicable to the original task design.

Here, we propose a novel BART model, which shows good parameter recovery, is applicable to the original task design, and is interpretable within a general RL framework. First, we introduce the existing model (the four-parameter model). Then, we reparameterize the four-parameter model to improve its parameter recovery. Finally, we develop the new model by modifying equations from the reparameterized version of the four-parameter model. Using the Leave-one-out information criterion (LOOIC) and the parameter recovery, we show that the new model outperforms the original model. To validate the new model, we compared the parameters with similar psychological constructs of competing models and apply the new model to BART data from patients with past opiate and stimulant dependence.

2. Method

2.1 Participants

593 individuals, enrolled for a study of impulsivity in opiate and stimulant users in Sofia, Bulgaria, were initially included. Then, only those who meet the following criteria were included: age between 18 and 50 years, longer than 8 years of formal education, estimated IQ of 80 or above, no history of head injury or loss of consciousness for more than 30 minutes., no history of neurological illness or psychotic disorders, HIV-seronegative status, and not currently on opioid maintenance therapy. All participants had negative breathalyzer test for alcohol and negative rapid urine toxicology screen for opiates, cannabis, amphetamines, methamphetamines, benzodiazepines, barbiturates, cocaine, MDMA, and methadone. We classified the included participants into three groups: healthy controls, heroin-dependent, and amphetamine-dependent groups. After that, group-specific criteria were applied to make each group include primarily mono-dependent ('pure') users. For the healthy control group, participants with any substance dependence or abuse symptom based on DSM-IV criteria were excluded (except for nicotine, caffeine, and past cannabis dependence). For the heroin and the amphetamine groups, mono-substance dependent participants who met DSM-IV lifetime criteria for opiate or stimulant dependence with no dependence on any other substances were included. Finally, a total of 226 subjects (135 healthy controls, 47 heroin dependent, and 44 amphetamine dependent individuals) were included in the analysis. For more details about the recruitment and screening procedures, see Ahn and Vassileva (2016). This study was approved by the Institutional Review Boards of the Virginia Commonwealth University and the Medical University in Sofia. All participants provided

informed consent. See Supplementary material for demographic and clinical characteristics of the participants (Table S1).

2.2 Task

In the BART, a virtual balloon is presented to the participant on each trial. Participants need to decide whether to pump the balloon to accumulate some predefined amount of reward (i.e., pump), or cash-out and receive the reward that has been accumulated so far (i.e., transfer). Each trial ends when the participant chooses to cash-out the accumulated reward or the balloon explodes. If the balloon explodes, the participant loses all the accumulated reward on that trial. Participants are not informed about the probability of the balloon exploding. Typically, the degree of risk-taking on the BART is measured by the adjusted BART score, which is the average number of pumps for unexploded balloons (Lejuez et al., 2002). The adjusted BART score is preferable because it is not directly affected by the explosion probability and reflects pure intention. To examine group differences of the adjusted BART score between the three groups, we conducted Bayesian t-test by using the R package *BEST* (Kruschke, 2013; Meredith & Kruschke, 2018).

2.3 Models

2.3.1 The four-parameter model

The four-parameter model (Wallsten et al., 2005) is based on two assumptions. First, the participants update the belief about the probability of the balloon exploding after each trial. Second, the participants decide the optimal number of pumps before each trial.

From a computational modeling perspective, the first assumption means that $p_{\mathrm{k}}^{\mathrm{burst}}$, the participant's perceived probability that pumping the balloon on trial k will make the balloon explode, is constant during the trial k. The participant initially has a prior belief about the probability of the balloon exploding and updates the prior belief based on observation on each trial. The updating process is described as follows:

$$p_k^{burst} = 1 - \frac{\alpha + \sum_{i=0}^{k-1} n_i^{success}}{\mu + \sum_{i=0}^{k-1} n_i^{pumps}} \; with \; 0 < \alpha < \mu. \qquad (1)$$

In Equation (1), the initial value of p_k^{burst} is $1 - \alpha/\mu$, which reflects the participant's initial belief that pumping will make the balloon explode. The magnitudes of α and μ indicate the degree of learning from observations; high values indicate that the prior belief is strong and the perceived probability is hard to be affected by the observed data. $\sum_{i=0}^{k-1} n_i^{success}$ is the sum of the number of successful pumps up to trial $k-1$, and $\sum_{i=0}^{k-1} n_i^{pumps}$ is the sum of the total number of pumps up to trial $k-1$.

The second assumption that the participant evaluates the optimal number of pumps before each trial is reflected in the equations for calculating the probability that the participant will pump the balloon. Adopting the prospect theory (Kahneman & Tversky, 2013), the expected utility after l pumps on trial k, U_{kl}, is given by:

$$U_{kl} = \left(1 - p_k^{burst}\right)^l (lr)^\gamma. \qquad (2)$$

In Equation (2), r is the amount of reward per successful pump, and γ is risk-taking propensity. We can calculate the optimal number of pumps by setting the first derivative of Equation (2) for l equals zero. Then, we can easily derive the optimal number of pumps on trial k, ω_k, as follows:

$$\omega_k = \frac{-\gamma}{\ln\left(1 - p_k^{burst}\right)} \ with\, \gamma \geq 0. \qquad (3)$$

Based on ω_k, we can calculate the probability that the participant will pump the balloon on trial k for pump l, p_{kl}^{pump}:

$$p_{kl}^{pump} = \frac{1}{1 + e^{\tau(l - \omega_k)}} \ with\, \tau \geq 0. \qquad (4)$$

In this logistic equation, τ is the inverse temperature parameter. The inverse temperature of the choice rule determines how deterministic or random the choice is; The higher τ, the more deterministic. If $l < \omega_k$, p_{kl}^{pump} becomes greater than 0.5. Similarly, if $l > \omega_k$, p_{kl}^{pump} becomes less than 0.5. In sum, the four-parameter model has four parameters to be estimated: $\alpha, \mu, \gamma,$ and τ.

We can calculate the likelihood of the data given the parameters by multiplying the probability that the participant will pump on trial k for pump l, p_{kl}^{pump}. The probability of the data given the parameters, $p(D|\alpha, \mu, \gamma, \tau)$, is given by:

$$p(D|\alpha, \mu, \gamma, \tau) = \prod_{k=1}^{k^{last}} \prod_{l=1}^{l_k^{last}} p_{kl}^{pump} \left(1 - p_{k, l_k^{last}+1}^{pump}\right)^{d_k}, \qquad (5)$$

where k^{last} is the last number of trials, l_k^{last} is the last number of pumping opportunities on trial k. if the participant transfers the accumulated reward on trial k to the virtual bank account, $d_k = 1$ and if the balloon explodes on trial k, $d_k = 0$.

Although previous researchers have primarily used the four-parameter model, it is known that α and μ of the four-parameter are not well recovered (van Ravenzwaaij et al., 2011). We suspected that the strong association between α and μ (the ratio reflects the participant's initial belief and the magnitudes indicate the degree of learning from observations) may be problematic and tested if reparametrizing α and μ would improve the parameter recovery performance of the model.

2.3.2 Reparametrized version (Par4 model)

The parameters α and μ are associated with two processes. The ratio of α to μ, α/μ, refers to the participant's initial belief that pumping will make the balloon explode, and the magnitudes of both α and μ determine the degree of learning from observations. Thus, we wanted to reparametrize them so that each parameter is uniquely associated with just one process. Also, we wanted to remove the constraint that α is less than μ because the constraint might lead to inefficient sampling for Bayesian parameter estimation (see Section 2.4).

For the goal, we reparametrized α and μ into ϕ and η: $\phi = \alpha/\mu$ and $\eta = 1/\mu$. Substituting these parameters into Equation (1) yields:

$$p_k^{burst} = 1 - \frac{\phi + \eta \sum_{i=0}^{k-1} n_i^{success}}{1 + \eta \sum_{i=0}^{k-1} n_i^{pumps}} \; with \; 0 < \phi < 1, \eta > 0. \quad (6)$$

After the reparameterization, the initial value of p_k^{burst} equals $1 - \phi$. Thus, ϕ indicates the participant's initial belief that pumping will not make the balloon explode. Also, η is an updating coefficient of the participant's belief by the observed data. If $\eta = 0$, p_k^{burst} is not affected by the observed data. If η is very large, p_k^{burst} rapidly comes close to the observed probability of burst. The Equations (3) and (4) remain the same. Note that we included the reparameterized four-parameter version in the hBayesDM package as a function named *bart_par4* (Ahn, Haines, & Zhang, 2017).

2.3.3 The Exponential Weight Updating model (EU model)

The four-parameter model has two critical limitations even after the reparameterization. The reparametrized version of the four-parameter model consists of two parts. First, we calculate p_k^{burst}. Subsequently, we calculate p_{kl}^{pump} from p_k^{burst}. First, although the participant updates p_k^{burst} from the prior belief with the observed data, the form of the Equation (6) does not provide intuitive interpretation about the learning process. We modified the updating equation in an attempt to more clearly show the learning process. Second, the assumption that the participant determines the optimal number of pumps before each trial may be unjustified. Instead, the participant may decide whether to pump the balloon or not just before each pump.

To address the first issue, we defined a parameter, $\psi = 1 - \phi$, which is the initial value of p_k^{burst}. Substituting this parameter into Equation (6) yields:

$$p_k^{burst} = \omega_{k-1}\psi + (1 - \omega_{k-1})P_{k-1} \; with \; 0 < \psi < 1, \eta > 0, \quad (7)$$

where $P_{k-1} = \frac{\sum_{i=0}^{k-1}(n_i^{pumps} - n_i^{success})}{\sum_{i=0}^{k-1} n_i^{pumps}}$, which is the observed probability that pumping has made the balloon explode up to trial $k-1$, and $\omega_{k-1} = \frac{1}{1+\eta\sum_{i=0}^{k-1} n_i^{pumps}}$, which is the weight indicating how much weight is given to the prior belief on trial k when estimating the probability of the balloon exploding. Each component of the Equation (7) has a clear role, which is interpretable in a general RL framework. Specifically, the current value (p_k^{burst}) is estimated as a weighted average of the initial value (ψ) and the observed value (P_{k-1}). As data accumulates, the participant updates the weight (ω_{k-1}) and the observed value (P_{k-1}). The weight (ω_{k-1}) and the observed value (P_{k-1}) are determined by the total number of the data ($\sum_{i=0}^{k-1} n_i^{pumps}$) and the number of the data that meet a certain condition (explosion, $\sum_{i=0}^{k-1}(n_i^{pumps} - n_i^{success})$). In this framework, η indicates how rapidly the participant depends on experience. If $\eta \to \infty$, learning entirely depends on the present outcome. If $\eta = 0$, no further learning occurs.

To improve the model performance within this framework, we modified the functional form of the weight, ω_{k-1}. If we define x as $\eta\sum_{i=0}^{k-1} n_i^{pumps}$, in Equation (7), $\omega_{k-1} = \frac{1}{1+x}$, which means the weight is hyperbolic. Instead of the hyperbolic function, other functional forms are possible if they meet two conditions: the participant's learning starts with the prior belief (if $x = 0$, $\omega_{k-1} = 1$) and primarily depends on the observed value after observing enough data (if $x \to \infty$, $\omega_{k-1} \to 0$). The exponential decay, e^{-x}, is a reasonable alternative because it meets the two conditions and is commonly used to describe natural phenomena such as the voltage of the resistor-capacitor circuit, the number of remain radioactive atoms, and the concentration of the first-order chemical reaction. Replacing ω_{k-1} with e^{-x} in Equation (7) yields:

$$p_k^{burst} = e^{-\xi\sum_{i=0}^{k-1} n_i^{pumps}}\psi + (1 - e^{-\xi\sum_{i=0}^{k-1} n_i^{pumps}})P_{k-1} \; with \; 0 < \psi < 1, \xi > 0. \quad (8)$$

To avoid confusion, we replaced η with ξ and named ξ an updating exponent.

For the second issue (the assumption that the participant determines the optimal number of pumps before each trial), we tested a new model which assumed that the participant decides whether to pump the balloon or not before each pump instead of each trial. Using the prospect theory

(Kahneman & Tversky, 2013), we calculated the subjective utilities for pumping and not-pumping a balloon on trial k for pump l as follows:

$$U_{kl}^{pump} = \left(1 - p_k^{burst}\right) r^{\rho} - p_k^{burst} \lambda \{(l-1)r\}^{\rho} \quad with\, 0 < \rho < 2, \lambda > 0, \qquad (9)$$

$$U_{kl}^{transfer} = 0, \qquad (10)$$

where r is the amount of reward for each successful pump, ρ is risk preference, and λ is loss aversion. Then, we can calculate the probability that the participant will pump the balloon on trial k for pump l, p_{kl}^{pump}, by using these subjective utilities.

$$p_{kl}^{pump} = \frac{1}{1 + e^{\tau\left(U_{kl}^{transfer} - U_{kl}^{pump}\right)}} \; with\, \tau \geq 0, \qquad (11)$$

where τ is inverse temperature. We noticed that this model (Equations (9), (10), and (11)) is similar to a model reported in Wallsten et al. (2005), which was not the best-fitting model. However, the model fit of the unchosen model was close to that of the winning model (the four-parameter model). Unlike the model reported in Wallsten et al. (2005), our EU model has a single parameter for risk preference instead of having separate risk preference for gain and loss. In summary, the EU model has five free parameters to be estimated: ψ (prior belief of burst), ξ (updating exponent), ρ (risk preference), τ (inverse temperature), and λ (loss aversion).

2.4 Hierarchical Bayesian Analysis (HBA)

We used hierarchical Bayesian Analysis (HBA) for parameter estimation (Berger, 2013; Gelman, Carlin, Stern, & Rubin, 2004; Lee, 2011). HBA offers several benefits over conventional non-hierarchical approaches such as individual-level ordinary least squares and maximum likelihood estimation (MLE) methods. First, HBA estimates parameters as posterior distributions instead of point estimates. Posterior distributions provide us with more information about the parameters than point estimates because distributions show the uncertainty of the estimated values. Second, with HBA, we can systematically characterize similarities and differences across subjects within a Bayesian framework based on the amount of information from each individual. Previous studies suggest that HBA allow us to estimate model parameters more accurately than individual- or group-level MLE methods (Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011).

We conducted HBA by using Stan (version 2.15.1), a probabilistic programming language for specifying statistical models (Carpenter et al., 2017). Stan uses Hamiltonian Monte Carlo (HMC) for sampling from high-dimensional parameter space. Specifically, we implemented the models in the hBayesDM (Ahn et al., 2017) environment, which uses Stan. We used 4000 samples including 2000 burn-in samples and 4 independent chains to make sure that the estimated parameter values are convergent. The trace plots indicated that chains were well mixed and the $\hat{R}$ values (Gelman & Rubin, 1992) for all model parameters were lower than 1.1, which indicates that the estimated parameter values converged to their target posterior distributions.

2.5 Model comparison

2.5.1 Leave-one-out information criterion (LOOIC)

LOOIC is an information criterion calculated from the Leave-one-out cross-validation. Leave-one-out cross-validation is a method to estimate out-of-sample prediction accuracy from a fitted Bayesian model based on the log-likelihood evaluated from the posterior distributions (Vehtari, Gelman, & Gabry, 2017). It is well-known that LOOIC has various benefits over simpler estimates such as Akaike Information Criterion (AIC, Akaike, 1998) and Bayesian Information Criterion (BIC, Schwarz, 1978). We used the R package *loo* (Vehtari et al., 2017) to estimate LOOIC for each model. Because LOOIC is calculated from the log-likelihood, the lower LOOIC is, the better its model fit is.

2.5.2 Parameter recovery

We also used parameter recovery to evaluate how accurate a model estimates true parameter values from the simulation data generated from the true parameter values (e.g., Ahn et al., 2011; Ahn et al., 2014; Haines et al., 2018; Wagenmakers et al., 2007). For the comparison, we did parameter recovery analysis for the reparametrized version of the four-parameter model and the EU model. In each model, we used estimated posterior means from healthy controls' data as the true parameter values because we wanted to use plausible combinations of parameter values for parameter recovery analysis. Then, we generated simulation data by using the true parameter values (for 135 subjects and 30 trials per subject). Lastly, we estimated parameter values from the simulation data. Correlations between the true parameter values and the predicted parameter values were used to evaluate the model performance.

3. Results

3.1 Model comparison

3.1.1 Leave-one-out information criterion (LOOIC)

Table 1 shows the LOOIC for the EU and the Par4 models. The EU model outperforms the Par4 model in all three groups. LOOIC weights, the calculated likelihoods of the models based on LOOIC, strongly favored the EU model as indicated by its LOOIC weights of 1.

Group	Model	LOOIC	LOOIC Weights
HC	EU	20454.267	1.000
	Par4	20666.613	0.000
Her	EU	7080.670	1.000
	Par4	7182.598	0.000
Amp	EU	6793.605	1.000
	Par4	6877.514	0.000

Table 1. Leave-one-out information criterion for each model. EU*:* the new BART model, Par4: the reparametrized version of the four-parameter model, HC: healthy control group, Her: heroin group, Amp: amphetamine group. Lower LOOIC indicates better model fit. The LOOIC weight is the likelihood of the model calculated based on its LOOIC.

3.1.2 Parameter recovery

Figure 1 shows the results of parameter recovery for the Par4 model. We evaluated the quality of parameter recovery as the correlation coefficient between the true and estimated values. As shown in Figure 1, overall all parameters were relatively well recovered in the Par4 model including the prior belief of success (ϕ) and the updating coefficient (η), which were not well recovered and systematically overestimated in the previous studies (Heathcote et al., 2015; van Ravenzwaaij et al., 2011). This suggests that our reparameterization may have improved the parameter recovery performance by separating the roles of the two parameters. To directly compare the parameter recovery of the four-parameter model and the Par4 model, we attempted to recover the parameters of the four-parameter model, but the parameters of the original four-parameter model failed to converge even after many (e.g., 4000) burn-in samples. We can think about two possible reasons. First, given that the magnitudes of α and μ commonly indicate the degree of learning from observations, the high correlation between the two parameters might make the sampling process fail to work well even with HMC. Second, the constraint that μ is always larger than α may cause issues in the sampling process.

Figure 2 shows the results of parameter recovery for the EU model. The EU model shows similar parameter recovery performance to the Par4 model (the average of the correlation coefficients for the Par4 model and the EU model are 0.888 and 0.895, respectively). Although the parameter recovery performance of the two models is similar and it is hard to test whether the difference is significant, we chose the EU model as a winning model in parameter recovery because the EU model includes one more parameter than the Par4 model. Parameter recovery performance tends to get worse as the number of parameters included in the model increases.

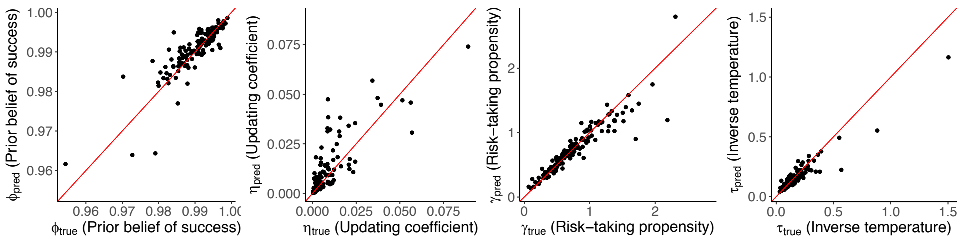

Figure 1. Parameter recovery results for the reparametrized version of the four-parameter model (Par4 model). The red line denotes $y = x$. The correlation coefficient of each scatter plot is as follows. ϕ (prior belief of success): 0.863, η (updating coefficient): 0.809, γ (risk-taking propensity): 0.934, τ (inverse temperature): 0.948. The average of the correlation coefficients is 0.888.

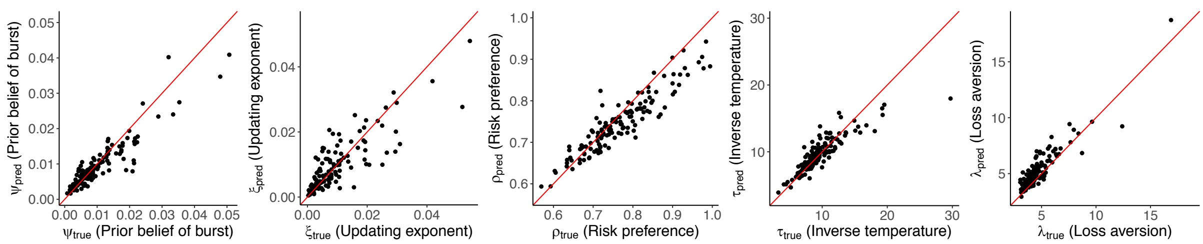

Figure 2. Parameter recovery results for the new BART model (EU model). The red line denotes $y = x$. The correlation coefficient of each scatter plot is as follows. ψ (prior belief of burst): 0.924, ξ (updating exponent): 0.842, ρ (risk preference): 0.923, τ (inverse temperature): 0.866, λ (loss aversion): 0.918. The average of the correlation coefficients is 0.895.

3.2 Correlation analysis

Model parameters of the Par4 and the EU models are seemingly corresponding to each other: (γ, ρ), (ϕ, ψ), (η, ξ), and (τ, τ), the former one is the parameter of the Par4 model and the latter one is the parameter of the EU model. The risk-taking propensity (γ) and the risk preference (ρ) relate to the risk-taking tendency. The prior belief of success (ϕ) and the prior belief of burst (ψ) correspond to the participant's prior belief about the balloon. The updating coefficient (η) and the updating exponent (ξ) mean updating rate of observation. The two inverse temperatures (τ) reflect how much the participant is deterministic.

We examined the correlations between parameters that have similar constructs in each of the two models. Figure 3 shows the correlations between the corresponding parameter pairs. All of the pairs have strong correlations. Although the correlation between the two inverse temperatures (τ) is relatively weak, it is acceptable given that they are related to different quantities; one is related to the number of pumps, and the other is related to the subjective utility. The prior belief of success (ϕ) and the prior belief of burst (ψ) are negatively correlated because the sum of the two probabilities should be 1 in an ideal case. The updating coefficient (η) and the updating exponent (ξ) are positively correlated because both of them represent how rapidly the participant updates the belief based on prior experience. Notably, risk-taking propensity (γ) and risk preference (ρ) are strongly and negatively correlated, which is contrary to our intuition.

To understand the negative correlation between risk-taking propensity and risk preference, we conducted a simulation analysis. Specifically, we generated simulation data of 30 trials for each value of the risk preference between the minimum value and the maximum value estimated from the real data. Other parameter values except the risk preference were fixed to the posterior means of the group parameters. Then, we calculated the adjusted average number of pumps based on the simulation data. Figure 4 clearly shows that the adjusted average number of pumps is negatively correlated with the risk preference. As the name implies, individuals with higher risk preference tend to seek more reward even though the probability of getting that reward is lower. However, the simulation result shows that the participants with higher risk preference press fewer number of pumps, which indicates they take less risk to get the same amount of reward. In other words, the risk preference of the EU model acts as 'risk aversion' in the BART paradigm. It is because in BART, a loss event is rare but the loss amount is large. In Equation (9), the subjective utility for pump decreases as the risk preference increases because the loss amount is much larger than the reward amount. The reduced subjective utility for pump decreases the probability that the participant will pump the balloon, so the average number of pumps decreases. This simulation result is consistent with a previous study using the Columbia Card Task (Pedroni et al., 2018). In contrast, the risk-taking propensity obviously has a positive correlation with the adjusted average number of pumps because the risk-taking propensity is proportional to the optimal number of pumps according to the Equation (3). Thus, the risk-taking propensity and the risk preference have a negative correlation.

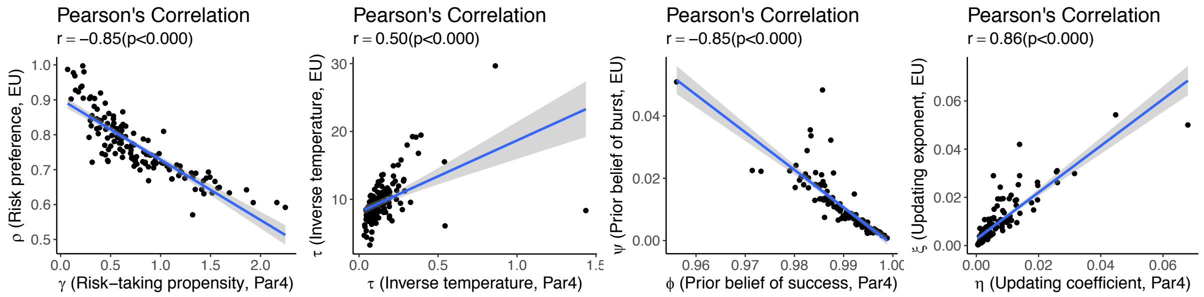

Figure 3. Correlations between the corresponding parameter pairs of the models. The blue lines indicate regression lines of each graph. Shaded regions indicate 95% confidence interval.

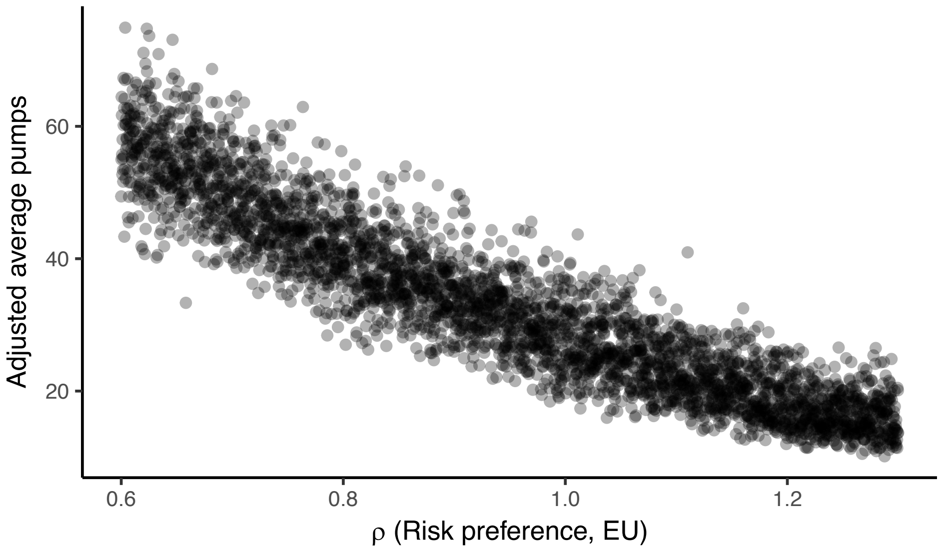

Figure 4. Simulated adjusted average number of pumps according to the risk preference. To find out the influence of risk preference on the adjusted average number of pumps, other parameter values except risk preference were fixed to the posterior mean values of the group parameters. Then, we generated simulation data of 30 trials for each value of risk preference as changing the risk preference. Lastly, we calculated the adjusted average number of pumps based on the simulation data. As the graph shows, the adjusted average number of pumps is negatively correlated with the risk preference.

3.3 Group difference

As a way of testing construct validity of the EU model, we applied the EU model to healthy and substance-dependent populations. We examined the group differences of three groups (healthy controls, heroin dependent, and amphetamine dependent; see below for the details) with respect to their behavioral performance and the parameter estimates of the EU model (we also tested the Par4 model).

3.3.1 Behavioral Performance

The heroin group displayed a marginally lower adjusted BART score (95% HDI: -9.73 ~ 0.629, mean= -4.59; with 95.9% of the posterior samples were smaller than 0) than the amphetamine group. The result suggests that heroin users might show lower risk-taking than amphetamine users during the BART. See Supplementary material for detail information of the behavioral performance and the group difference of the behavioral performance (Figure S1 and Figure S2).

3.3.2 Model Parameters

We estimated parameters of the EU model and the Par4 model for each group separately to compare the parameter values between the groups. Figure 5 shows the posterior distributions of the group parameters for each group with the EU model. The heroin group displayed credibly higher risk preference (ρ) than the healthy control group (95% HDI: 0.158 ~ 0.467, mean: 0.311) and the amphetamine group (95% HDI: 0.211 ~ 0.584, mean: 0.401). Also, the heroin group displayed credibly lower loss aversion (λ) than the healthy control group (95% HDI: -4.23 ~ -1.44, mean: -2.82) and the amphetamine group (95% HDI: -7.60 ~ -1.62, mean: -4.29). Additionally, the heroin group displayed credibly higher updating exponent (ξ) than the amphetamine group (95% HDI: 4.58e-4 ~ 0.0140, mean: 6.03e-3), and lower inverse temperature (τ) than the healthy control group (95% HDI: -2.20 ~ -0.0865, mean: -1.14). Figure 6 shows the posterior distributions of the group parameters for each group with the Par4 model. The heroin group displayed credibly lower risk-taking propensity (γ) than the amphetamine group (95% HDI: -0.280 ~ -0.012, mean: -0.151). Also, the heroin group displayed higher inverse temperature (τ) than the amphetamine group (95% HDI: 0.019 ~ 0.070, mean: 0.044).

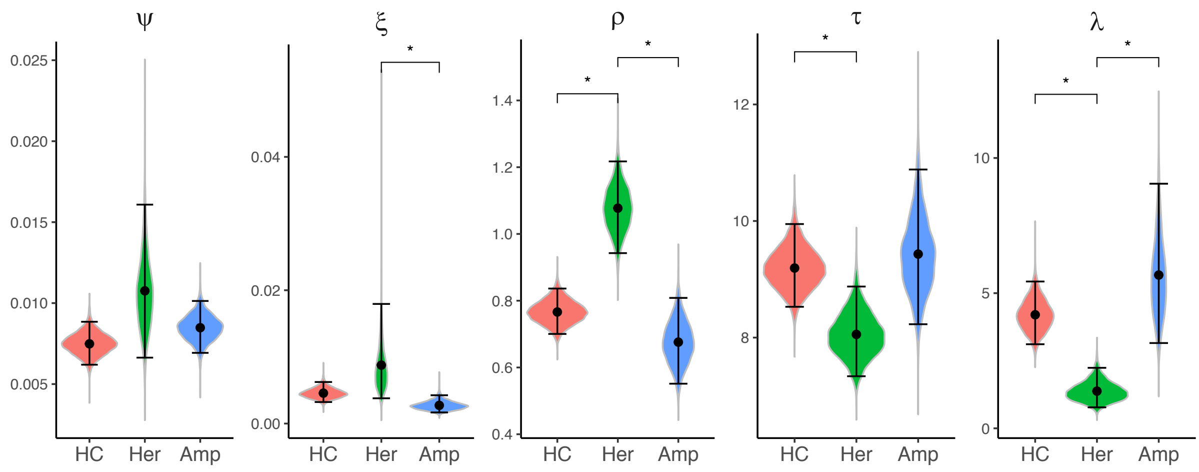

Figure 5. Posterior distributions of the group parameters with the EU model. Tick marks on bottom and top of each graph indicate 95% highest density intervals (HDIs). Points on the middle of each graph indicate mean values. Asterisks indicate credible differences. ψ: prior belief of burst, ξ: updating exponent, ρ: risk preference, τ: inverse temperature, λ: loss aversion. HC: healthy control group, Her: heroin group, Amp: amphetamine group.

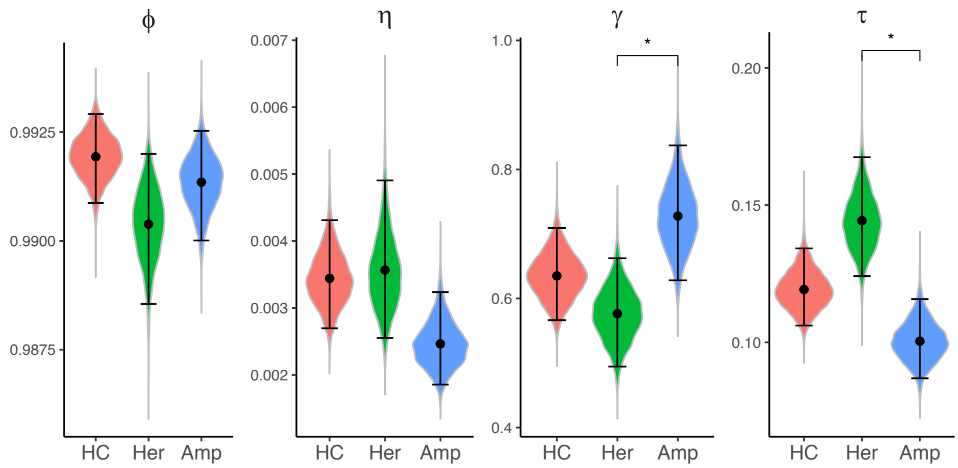

Figure 6. Posterior distributions of the group parameters with the reparameterized version of the four-parameter model (Par4 model). Tick marks on bottom and top of each graph indicate 95% highest density intervals (HDIs). Points on the middle of each graph indicate mean values. Asterisks indicate credible differences. ϕ: prior belief of success, η: updating coefficient, γ: risk-taking propensity, τ: inverse temperature. HC: healthy control group, Her: heroin group, Amp: amphetamine group.

The results of the behavioral performance and the model parameters are consistent. Among the three groups, the differences between the heroin group and the amphetamine group were the most noticeable. The heroin group displayed a marginally lower adjusted BART score, lower risk-taking propensity (γ), and higher risk preference (ρ). Given that the risk preference (ρ) acts as risk aversion in the BART (Figure 4), these results consistently show that heroin users show lower risk-taking than amphetamine users during the BART.

The group difference results also show that the model parameters of the EU model provide additional information about underlying cognitive processes, which cannot be obtained from behavioral performance indices. Specifically, the group differences in model parameters of the EU model show that the heroin group displayed not only higher risk preference (ρ) but also lower loss aversion (λ) than the amphetamine group, which we cannot infer from behavioral performance alone. It is of note that the reduced loss aversion of the heroin group is consistent with the result of a previous study of decision-making using the Iowa Gambling Task (Ahn et al., 2014).

4. Discussion

The main focus of this study is on the development of a novel BART model that outperforms existing models. For this purpose, we proposed a reparameterized version of the four-parameter model (the Par4 model) and developed a new BART model (the EU model) based on the reparameterized version. To evaluate the EU model, we compared it with the Par4 model. The model comparison results suggest that the EU model shows better prediction performance across all populations and better parameter recovery than the Par4 model. To validate the EU model, we calculated the correlations between corresponding parameter pairs for the Par4 and the EU models. All of the corresponding parameter pairs had strong correlations. However, contrary to our expectations, there was a negative correlation between risk-taking propensity and risk preference. We used simulation analysis to explore the reasons for this relationship. As a way of testing construct validity of the EU model, we analyzed differences among substance-using populations in behavioral performance and model parameters of the Par4 and the EU models. The group differences in behavioral performance and model parameters of the Par4 and the EU models were consistent with each other. The results of the correlations and the group differences provide supporting evidence to validate the EU model. The overall results suggest that the EU model should be the most prioritized computational model for the BART.

An important finding of this study is that it suggests a way to improve parameter recovery. We showed that reparameterizing parameters associated with more than one role into parameters with unique roles helps the model recover accurate parameter values. Adequate parameter recovery is a fundamental assumption and necessary for adequately analyzing parameters of a computational model and it is noteworthy that we can improve parameter recovery by reparameterization alone. At the same time, it should be noted that the information criteria such as AIC, BIC, and LOOIC for the reparameterized version and the original model are more or less the same. It suggests that the reparameterized version doesn't have additional explanatory power compared with the original model. The results demonstrate that parameter recovery and post-hoc fits measured with information criteria reflect different aspects of computational models, and we need to use both methods for comprehensive evaluation.

When we analyzed the group difference between substance-dependent populations by using behavioral performance and the model parameters of the Par4 and the EU model, there were notable group differences between heroin users and amphetamine users. This result is consistent with the results of previous studies showing that opiates (heroin) and stimulants (amphetamine) addiction are behaviorally and neurobiologically distinct (Badiani, Belin, Epstein, Calu, & Shaham, 2011) and related to different dopamine modulation mechanisms (Kreek et al., 2012). Notably, the group differences in model parameters of the EU model show reduced loss aversion in heroin users compared to amphetamine users, which is consistent with the result of the previous study using the Iowa Gambling Task (Ahn et al., 2014). Together, these results suggest that reduced loss aversion may be a computational marker for heroin dependence.

There are several advantages of the EU model compared to existing models. Importantly, the EU model utilizes a general RL framework for understanding the learning process of the BART. The weight updating framework used in the EU model may be an alternative to other well-established models such as the Rescorla-Wagner model (Rescorla & Wagner, 1972) to quantify learning situations. Furthermore, with the EU model that contains two risk-taking related parameters, risk preference (ρ) and loss aversion (λ), it is possible to separately identify sensitivity to the actual value change and tendency to avoid loss. As a result, the risk preference (ρ) and the loss aversion (λ) for the EU model uniquely detects the difference in risk-taking between the healthy group and the heroin group (Figure 5). In contrast, the Par4 model only contains one parameter, risk-taking propensity (γ), related to risk-taking, which is proportional to the optimal number of pumps.

In summary, we proposed a novel BART model that outperforms the existing model and validated the model by analyzing correlations between the model parameters and group differences between different substance-dependent populations. The new model provides a general RL framework for understanding the learning process which is generalizable to other learning tasks and identifies the cognitive components related to risk-taking in the BART. We expect that the weight updating framework may be applicable to other computational models.

Declarations of interest: none

Acknowledgements

We thank the Seoul Science High School (SSHS) students for assistance. Also, we would like to thank all volunteers for their participation in this study. We express our gratitude to Georgi Vasilev, Kiril Bozgunov, Elena Psederska, Dimitar Nedelchev, Rada Naslednikova, Ivaylo Raynov, Emiliya Peneva and Victoria Dobrojalieva for assistance with recruitment and testing of study participants. Research reported in this publication was supported in part by the National Institute on Drug Abuse and Fogarty International Center under award number R01DA021421 to J.V, and the Basic Science Research Program through the National Research Foundation (NRF) 795 of Korea funded by the Ministry of Science, ICT, & Future Planning (NRF-2018R1C1B3007313 and NRF-2018R1A4A1025891) to W.-Y.A.

References

Ahn, W. Y., Busemeyer, J. R., Wagenmakers, E. J., & Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cognitive science, 32*(8), 1376-1402.

Ahn, W. Y., Dai, J., Vassileva, J., Busemeyer, J. R., & Stout, J. C. (2016). Computational modeling for addiction medicine: from cognitive models to clinical applications. In *Progress in brain research* (Vol. 224, pp. 53-65): Elsevier.

Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Comput Psychiatr, 1*, 24-57. doi:10.1162/CPSY_a_00002

Ahn, W. Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics, 4*(2), 95-110. doi:10.1037/a0020684

Ahn, W. Y., Vasilev, G., Lee, S.-H., Busemeyer, J. R., Kruschke, J. K., Bechara, A., & Vassileva, J. (2014). Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. *Frontiers in Psychology, 5*, 849.

Ahn, W. Y., & Vassileva, J. (2016). Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence. *Drug alcohol dependence, 161*, 247-257.

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199-213): Springer.

Aklin, W. M., Lejuez, C., Zvolensky, M. J., Kahler, C. W., & Gwadz, M. (2005). Evaluation of behavioral measures of risk taking propensity with inner city adolescents. *Behaviour research therapy, 43*(2), 215-228.

Badiani, A., Belin, D., Epstein, D., Calu, D., & Shaham, Y. (2011). Opiate versus psychostimulant addiction: the differences do matter. *Nature reviews neuroscience, 12*(11), 685.

Berger, J. O. (2013). *Statistical decision theory and Bayesian analysis*: Springer Science & Business Media.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: decomposing performance on the Bechara gambling task. *Psychological assessment, 14*(3), 253.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., . . . Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software, 76*(1).

Chung, D., Christopoulos, G. I., King-Casas, B., Ball, S. B., & Chiu, P. H. (2015). Social signals of safety and risk confer utility and have asymmetric effects on observers' choices. *Nature neuroscience, 18*(6), 912.

Collins, A. G., Ciullo, B., Frank, M. J., & Badre, D. (2017). Working memory load strengthens reward prediction errors. *Journal of Neuroscience, 37*(16), 4332-4342.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204-1215.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). Bayesian data analysis, 2nd edn. Texts in Statistical Science. In: Boca Raton, London, NewYork, Washington DC: Chapman & Hall, CRC.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science, 7*(4), 457-472.

Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage, 62*(1), 154-166.

Haines, N., Vassileva, J., & Ahn, W. Y. (2018). The Outcome-Representation Learning Model: A Novel Reinforcement Learning Model of the Iowa Gambling Task. *Cognitive science, 42*(8), 2534-2561.

Heathcote, A., Brown, S. D., & Wagenmakers, E.-J. (2015). An introduction to good practices in cognitive modeling. In *An introduction to model-based cognitive neuroscience* (pp. 25-48): Springer.

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I* (pp. 99-127): World Scientific.

Kreek, M. J., Levran, O., Reed, B., Schlussman, S. D., Zhou, Y., & Butelman, E. R. (2012). Opiate addiction and cocaine addiction: underlying molecular neurobiology and genetics. *The Journal of clinical investigation, 122*(10), 3387-3393.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General, 142*(2), 573.

Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical Bayesian models. *Journal of Mathematical Psychology, 55*(1), 1-7.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., . . . Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied, 8*(2), 75-84. doi:10.1037/1076-898x.8.2.75

Lejuez, C. W., Simmons, B. L., Aklin, W. M., Daughters, S. B., & Dvir, S. (2004). Risk-taking propensity and risky sexual behavior of individuals in residential substance use treatment. *Addictive behaviors, 29*(8), 1643-1647.

Meredith, M., & Kruschke, J. K. (2018). Bayesian Estimation Supersedes the t-Test.

Pedroni, A., Rieskamp, J., Pachur, T., Frey, R., Westfall, J. E., & Figner, B. (2018). A Prospect Theory Account of the "Hot" Columbia Card Task.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review, 85*(2), 59.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research theory, 2*, 64-99.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics, 6*(2), 461-464.

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences, 106*(13), 5035-5040.

van Ravenzwaaij, D., Dutilh, G., & Wagenmakers, E.-J. (2011). Cognitive model decomposition of the BART: Assessment and application. *Journal of Mathematical Psychology, 55*(1), 94-105. doi:10.1016/j.jmp.2010.08.010

van Ravenzwaaij, D., & Oberauer, K. (2009). How to use the diffusion model: Parameter recovery of three methods: EZ, fast-dm, and DMAT. *Journal of Mathematical Psychology, 53*(6), 463-473.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics Computing, 27*(5), 1413-1432.

Wagenmakers, E.-J., Van Der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic bulletin review, 14*(1), 3-22.

Wallsten, T. S., Pleskac, T. J., & Lejuez, C. W. (2005). Modeling behavior in a clinically diagnostic sequential risk-taking task. *Psychol Rev, 112*(4), 862-880. doi:10.1037/0033-295X.112.4.862

Worthy, D. A., & Maddox, W. T. (2014). A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to WK Estes. *Journal of Mathematical Psychology, 59*, 41-49.

Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron, 75*(3), 418-424.