

Intertemporal Choice as Discounted Value Accumulation

Christian A. Rodriguez*, Brandon M. Turner, Samuel M. McClure

Department of Psychology, Stanford University, Stanford, California, United States of America

Abstract

Two separate cognitive processes are involved in choosing between rewards available at different points in time. The first is temporal discounting, which consists of combining information about the size and delay of prospective rewards to represent subjective values. The second involves a comparison of available rewards to enable an eventual choice on the basis of these subjective values. While several mathematical models of temporal discounting have been developed, the reward selection process has been largely unexplored. To address this limitation, we evaluated the applicability of the Linear Ballistic Accumulator (LBA) model as a theory of the selection process in intertemporal choice. The LBA model formalizes the selection process as a sequential sampling algorithm in which information about different choice options is integrated until a decision criterion is reached. We compared several versions of the LBA model to demonstrate that choice outcomes and response times in intertemporal choice are well captured by the LBA process. The relationship between choice outcomes and response times that derives from the LBA model cannot be explained by temporal discounting alone. Moreover, the drift rates that drive evidence accumulation in the best-fitting LBA model are related to independently estimated subjective values derived from various temporal discounting models. These findings provide a quantitative framework for predicting dynamics of choice-related activity during the reward selection process in intertemporal choice and link intertemporal choice to other classes of decisions in which the LBA model has been applied.

Citation: Rodriguez CA, Turner BM, McClure SM (2014) Intertemporal Choice as Discounted Value Accumulation. PLoS ONE 9(2): e90138. doi:10.1371/journal.pone.0090138

Editor: Jean Daunizeau, Brain and Spine Institute (ICM), France

Received: October 3, 2013; **Accepted:** January 28, 2014; **Published:** February 28, 2014

Copyright: © 2014 Rodriguez et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by National Science Foundation's (NSF) Graduate Research Fellowship Program (GRFP), the NSF's Integrative Graduate Education and Research Traineeship (IGERT) program, and National Institutes of Health (F32GM103288). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: carodrig@stanford.edu

Introduction

In order to choose between rewards available at different points in time it is often necessary to evaluate the tradeoff between the size of potential rewards and the corresponding delays until their receipt. For example, deciding whether to save or spend a certain amount of money requires determining whether ensuring greater future wealth is worth delaying the pleasure of spending and consuming now. When engaged in this form of decision making, a class of decisions known as intertemporal choice, humans and other species discount the value of rewards in proportion to the delay at which they are available. Moreover, the behavior observed in intertemporal choice experiments reveals preferences consistent with a steep reduction in the value of rewards delayed from the present moment but more modest discounting of rewards delayed from future time points [1]. This property is particularly evident as a greater reluctance to forego immediate for delayed rewards compared with when both outcomes are delayed, a tendency that manifests itself in impulsivity and a predilection for procrastination. Several mathematical models have been shown to account for this pattern of delay discounting [2]. However, subjective valuation is only one of the cognitive processes involved in intertemporal choice behavior [3,4].

In addition to representing the value of delayed rewards, intertemporal choices require comparing alternatives and selecting among them. One proposal for how delayed rewards might be compared and selected is through a process of sequential sampling of discounted values [3]. Similar processes are commonly assumed

to underlie perceptual judgments based on sensory evidence [5]. This hypothesis suggests that there exists a direct connection between choices made on the basis of discounted values and other choices which have been argued to derive from sequential sampling processes. However, the hypothesis that a sequential sampling process underlies intertemporal decision-making has not been empirically tested. Therefore, our primary goal is to determine whether intertemporal choice behavior can be explained by a sequential sampling process based on discounted value.

There are several computational models that employ sequential sampling mechanisms to explain choice behavior (cf. [6–8]). A major accomplishment of all of these models is their ability to provide a process-level account of how experimental manipulations such as time pressure and stimulus ambiguity simultaneously affect response times (RT) and error rates. While many of these models might be able to explain intertemporal choice behavior, we used the Linear Ballistic Accumulator (LBA) model [8] in our analyses. The LBA model incorporates the fundamental features of all sequential sampling models, including trial-to-trial variability in the rate of evidence accumulation, a decision criterion, and constants to account for perception and motor execution times. The major advantage of the LBA model is its analytical tractability, which facilitates testing several versions of the model to determine which combination of parameters best accounts for intertemporal choice behavior. We show that the LBA model provides an excellent description of the relationship between

choice outcomes and RT and that best-fitting model parameters can be directly related to subjective values.

Materials and Methods

Subjects

Fifty healthy adults participated in this study (28 females, ages 19–46 years, mean 24.36 years). All subjects gave written informed consent. Stanford University's Institutional Review Board approved the study. One subject was excluded because the behavior did not allow us to estimate reliable temporal discounting parameters. Another three subjects were excluded because of data collection problems. Data from a total of forty-six subjects were analyzed (28 females, ages 19–46 years, mean 24.26 years).

Temporal discounting model and task design

The experiments were conducted over two sessions. The purpose of the first session was to estimate each individual's discount rate using a hyperbolic discounting model. For half of our subjects ($n=23$) the second session consisted of an electroencephalography (EEG) experiment. For the other half the second session consisted of a functional magnetic resonance imaging (fMRI) experiment. The analyses reported below were obtained from the behavior observed during these EEG and fMRI sessions.

We assumed that the subjective value of delayed rewards was discounted according to

$$V_D = \frac{r}{1+kt} \quad (1)$$

where r is the magnitude of a reward offered at delay t . The individually-determined parameter k is the discount factor [9]. While subjects completed the first session, we used a stair-stepping procedure to approximate k . All choices required participants to select between a delayed reward (of amount r available at delay t) and a fixed immediate reward of \$10. For any choice, indifference between the immediate and delayed options implies a discount rate of $k=(r-10)(10t)^{-1}$. We refer to this implied equivalence point as k_{eq} ; our procedure amounted to varying k_{eq} systematically until indifference was reached. Specifically, we began with $k_{eq}=0.02$. If the delayed offer was chosen, k_{eq} was decreased by a step size of $\alpha=0.01$ for the next trial. Otherwise, k_{eq} increased by the same amount. At every second choice reversal, occurring within five consecutive trials, the step size was reduced by 5%. A total of 60 trials were completed. We placed no limits on the time subjects could take to respond, and presented both offers on the screen, as "\$10 now" on the left side, and "\$ r in t days" on the right.

Critically, our use of the hyperbolic discounting model to summarize behavior in this first experimental session had no bearing on the modeling results that follow. We used the hyperbolic model because it provided a good fit to behavior with a single parameter (k) summarizing preferences. Fits of this model were used solely to generate choices for the second experimental session. Alternative delay discounting functions that may or may not provide better fits to behavior would have a subtle impact on the choice set (dollar amounts of choice options) for the second study, but no impact on the model fitting that is the primary aim of the current study.

After completing the first session, we fit a softmax decision function to participants' choices. Intuitively, this procedure allowed us to determine how consistently participants selected the option with greater subjective value. Practically, we fit the softmax to better equate choices during the second session, across

participants. In particular, our aim was to equate the relative impact of delayed rewards, across subjects, with respect to actual choice outcomes (i.e. the likelihood of selecting the delayed option). Best fitting softmax functions were estimated by maximizing the likelihood of observed choices. We assumed that the likelihood of selecting a delayed reward (P_D) was given by

$$P_D = \frac{1}{1+e^{-m(V_D-V_I)}} \quad (2)$$

where V_D is given by Equation 1, $V_I = \$10$ (i.e., the fixed-value of the immediate reward also given by the right side of Equation 1) and m describes a subject's sensitivity to changes in V_D .

We used individually determined values of k and m to generate choices for the second session. At every trial, t was randomly selected from a range of 30–45 days. We then calculated and offered an amount r that would give P_D of 0.1, 0.3, 0.5, 0.7, or 0.9 (Figure 1a-b). The EEG group completed 30 trials at every P_D level, except at $P_D=0.5$, for which they completed 60 trials. The fMRI group completed 40 trials at every P_D level, except at $P_D=0.5$, for which they completed 80 trials. Non-uniform trial distributions as a function of P_D were introduced to allow us to study the effects of choice difficulty on EEG and fMRI measures, with equal numbers of trials at each difficulty level. We report the results of these analyses elsewhere. Trial types were randomized and counterbalanced over two blocks for the EEG group and over four blocks for the fMRI group. We also counterbalanced the mapping between choices and button presses for every subject. During the first half of the second session, approximately half of subjects (13 in EEG, 11 in fMRI) indicated choices of the delayed reward by pressing a button with their left index finger and immediate choices by pressing a different button with their right index finger. The other subjects indicated their choices by the inverse left-right mapping. All subjects switched the initial response mapping during the second half of the session.

To ensure reliable neural measures, we used a sequential presentation of delay and amount during the second session (Figure 1c). During pilot studies we found that a simultaneous presentation of delay and amount caused participants to sequentially fixate the information, producing excessive EEG artifacts. Having the information presented sequentially allowed subjects to maintain central fixation during the task, avoiding these artifacts. As we show below, this sequential presentation of delayed reward information had no adverse effects on behavior. We maintained the same sequential presentation during the fMRI study for the purpose of facilitating direct comparisons and pooling of behavioral data. We report RT as measured from the onset of the decision period, 1000 ms into the trial. The duration of the decision period was fixed at 4000 ms. When subjects made choices in less than 4000 ms the amount information disappeared and the screen remained blank until 4000 ms elapsed. Trial length was thus fixed at 5000 ms. We discarded any trial in which a response was made in less than 200 ms or fell outside of the decision period. To optimize experimental time and separability of neural signals across trials for both groups, we introduced a long inter-trial-interval for the fMRI group (between 4–10 s), whereas the inter-trial-interval was shorter for the EEG group (100–350 ms). In exchange for participation subjects received \$10 cash and an additional amount, determined by their choice in a randomly selected trial, taken from either the first or second sessions.

Model specification and fitting

Figure 2 provides an illustrative diagram of our LBA model of intertemporal choice. To provide a formal description of the

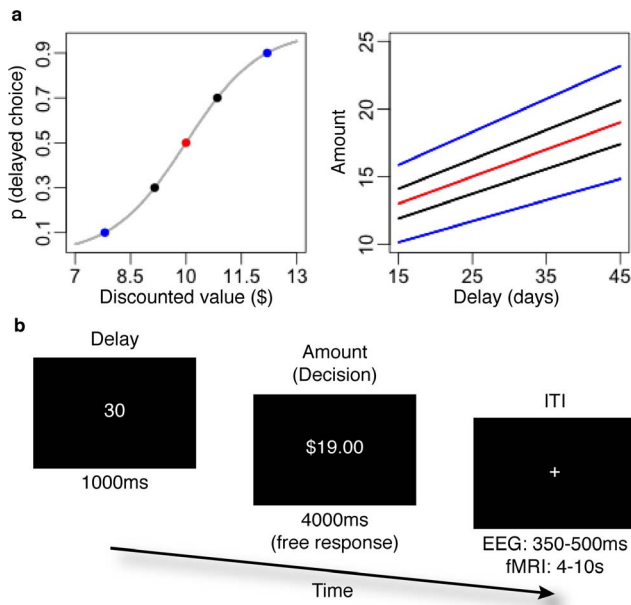


Figure 1. Experimental design. (a) Delayed reward offers corresponded with one of five different levels of discounted value. Each level of discounted value corresponded to one of five probabilities of choosing the delayed reward: 0.1, 0.3, 0.5, 0.7 or 0.9. (b) Every delay could be combined with any of five different amounts to yield a different discounted value and probability of choosing the delayed reward. (c) Delay and amount information was presented sequentially. Delays were presented first for 1000 ms. Amounts were presented second, replacing the presentation of the delay and remaining on the screen for a maximum of 4000 ms. After every trial, a fixation cross was presented on the center of the screen for a randomly chosen inter-trial-interval in the order of hundreds of milliseconds during the EEG experiment and several seconds during the fMRI experiment. doi:10.1371/journal.pone.0090138.g001

model, we denote the RT on the i th trial for the j th subject in the v th value condition as $RT_{i,j,v} \in (0, \infty)$, and the corresponding choice as $C_{i,j,v}$ where $C_{i,j,v} \in \{I, D\}$. I and D are the immediate and delayed rewards respectively. The model assumes that evidence for I and D is accumulated independently in separate accumulators. Both accumulators begin with some choice bias, which is provided as independent amounts of starting point evidence $\{a_I, a_D\}$, sampled from a common uniform distribution $\mathcal{U}[0, A]$. Evidence then increases through time at rates $\{d_I, d_D\}$, which are sampled from independent normal distributions with means $\{\mu_I, \mu_D\}$. Mean accumulation rates vary across value conditions, but the standard deviation σ is the same for I and D . Therefore, $d_I \sim \mathcal{N}(\mu_{v,I}, \sigma)$ and $d_D \sim \mathcal{N}(\mu_{v,D}, \sigma)$. Each accumulator gathers evidence until either one reaches a response threshold b . The observed RT is the sum of the decision time, plus some extra time τ , which accounts for non-comparison and selection processes, such as temporal discounting and motor execution. Letting $\{a_I, a_D\} = \mathbf{a}$ and $\{d_I, d_D\} = \mathbf{d}$, the RT in any given trial is given by

$$RT = \min_{\{I, D\}} \left(\frac{b-a}{d} \right) + \tau. \tag{3}$$

The model provides a closed-form and joint account of RT and choice probability across value conditions by specifying “defective” probability density functions (PDF) for I and D in terms of

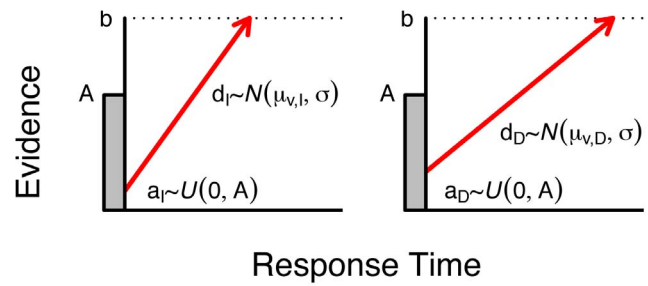


Figure 2. Illustrative diagram of the Linear Ballistic Accumulator model for intertemporal choice, where each response option is represented as a separate accumulator. Following the presentation of a stimulus and some non-decision time τ , information accumulates ballistically for each alternative. A decision is made that coincides with the accumulator that reaches the threshold b first. The model assumes trial-to-trial variation in both starting point and drift rate. doi:10.1371/journal.pone.0090138.g002

the parameters just described. These defective PDFs give the probabilities of each accumulator reaching the bound at time t . For our best fitting model, the full PDFs are given by

$$PDF_{v,I}(I, RT | \mu_{v,I}, \tau_v, a_I, b, \sigma) = f_I(RT - \tau)(1 - F_D(RT - \tau)), \text{ and} \\ PDF_{v,D}(D, RT | \mu_{v,D}, \tau_v, a_D, b, \sigma) = f_D(RT - \tau)(1 - F_I(RT - \tau)), \tag{4}$$

where $f(RT - \tau)$ and $F(RT - \tau)$ are the PDF and cumulative density functions of each accumulator (see [8] for details).

We estimated LBA model parameters using a hierarchical Bayesian procedure. This procedure offers two advantages over conventional maximum likelihood methods, providing measures of uncertainty for every parameter estimate and allowing the sharing of information across subjects (e.g., [10,11]), which improves fitting accuracy[12–14]. We assume that the data for each subject is characterized by an individual set of LBA model parameters θ , and that these subject-specific parameters are constrained by a set of group-level parameters ϕ , which characterize the central tendency and dispersion of θ across subjects. The procedure first samples the posterior distributions for every subjects’ θ and uses these estimates to derive the posterior distribution of ϕ . On every subsequent iteration, the posterior estimates of ϕ are used to constrain the sampling of possible values of θ for every subject. We specified mildly informative priors for θ , based on empirical evidence from previous fits of the LBA model using the hierarchical Bayesian procedure [15]. For ϕ , we specified a conjugate relationship between prior and posterior (see, e.g., [16]). Assuming a conjugate relationship at the group-level allowed us to derive exact conditional posterior distributions, so that we could perform the estimation of all of the parameters simultaneously, based on a single sample of subject-level parameters. The joint posterior distribution estimated by this procedure is given by:

$$p(\theta, \phi | C, RT) \propto p(\phi) p(\theta | \phi) \mathcal{L}(C, RT | \theta) \tag{5}$$

where $p(\phi)$ is the prior distribution for ϕ , $p(\theta | \phi)$ is the prior distribution for θ given ϕ , and

$$\mathcal{L}(C, RT | \theta) = \prod_i \prod_j \prod_v PDF_{v,C_{i,j,v}}(C_{i,j,v}, RT_{i,j,v} | \theta_j)$$

is the likelihood function of the data under the LBA model (given by Equation 4).

To satisfy scaling conditions, we imposed a constraint such that the drift rates sum to one (i.e., $\mu_{v,I} + \mu_{v,D} = 1$). Consequently, it is sufficient to only estimate the drift rate for the delayed reward. For the subject-specific parameters, we first transformed the parameters so that they had continuous, infinite support (i.e., can take on any real value). Thus, for parameters bounded by zero, we applied a log transformation, whereas for the drift rates – which were bounded by zero and one – we used a logit transformation. Following these transformations, we specified the following priors for θ :

$$\text{logit}(\mu_{v,D,j}) \sim \mathcal{N}(\mu_{\mu}^{(v)}, \mu_{\sigma}^{(v)}),$$

$$\log(\tau_{v,j}) \sim \mathcal{N}(\tau_{\mu}^{(v)}, \tau_{\sigma}^{(v)}),$$

$$\log(A_j) \sim \mathcal{N}(A_{\mu}, A_{\sigma}),$$

$$\log(b_j) \sim \mathcal{N}(b_{\mu}, b_{\sigma}), \text{ and}$$

$$\log(\sigma_j) \sim \mathcal{N}(\sigma_{\mu}, \sigma_{\sigma}).$$

To obtain the desired conjugate relationship between prior and posterior at the level of ϕ , we specified the following priors for the group-level means:

$$\mu_{\mu}^{(v)} \sim \mathcal{N}(0.75, 0.5),$$

$$\tau_{\mu}^{(v)} \sim \mathcal{N}(-1, 0.5),$$

$$A_{\mu} \sim \mathcal{N}(1.5, 0.8),$$

$$b_{\mu} \sim \mathcal{N}(1.5, 0.8), \text{ and}$$

$$\sigma_{\mu} \sim \mathcal{N}(0.75, 0.5),$$

and the following priors for the group-level standard deviations,

$$\mu_{\sigma}^{(v)} \sim \Gamma^{-1}(4, 10),$$

$$\tau_{\sigma}^{(v)} \sim \Gamma^{-1}(4, 10),$$

$$A_{\sigma} \sim \Gamma^{-1}(4, 10),$$

$$b_{\sigma} \sim \Gamma^{-1}(4, 10), \text{ and}$$

$$\sigma_{\sigma} \sim \Gamma^{-1}(4, 10),$$

where $\Gamma^{-1}(a, b)$ denotes the inverse gamma distribution with shape parameter a , and scale parameter b . This particular choice of a and b for the priors produces a skewed distribution with an approximate 95% credible set of (1.14, 9.05), and an expected value of 3.32. These choices reflect our *a priori* beliefs: we did not expect the between-subject variability to be less than 1, and felt that larger values would become increasingly less likely to account for these data.

While our prior selections were informed by other similar modeling applications (see, e.g., [15]), we remained conservative in our choices to avoid undue parameter constraint, because our experimental task was considerably different from prior research using the hierarchical version of the LBA model.

We used Gibbs sampling to estimate parameters at the group-level [16], and differential evolution with Markov chain Monte Carlo to estimate parameters at the subject-level (DE-MCMC; [15, 17]). For the subject level estimates, we used 24 chains and obtained 5,000 samples after a burn-in period of 5,000 samples. We then thinned the chains to reduce autocorrelation by retaining every fourth sample. Thus, our estimates of the joint posterior distribution of LBA model parameters are based on 30,000 samples. The burn-in period allowed us to converge quickly to the high-density regions of the posterior distribution, while the rest of the samples allowed us to improve the reliability of the estimates.

To find the optimal number of parameters needed to account for intertemporal choice behavior, we tested a variety of model variants where different sets of parameters were assumed to vary across value conditions. We fit a total of eight variants, following a model building approach based on the Bayesian predictive information criterion (BPIC; [18]). Table 1 shows the model variants we fit (left column) with the particular constraints that were imposed (right column) along with the resulting BPIC values obtained (middle column). We started with the simplest possible model and added parameters only if they improved model fits on the basis of BPIC. The most basic model (M1) only allowed the mean drift rates $\{\mu_I, \mu_D\}$ to vary across value conditions. Another four models freed each of the remaining parameters (τ, A, b and σ), independently, across value conditions. Because the model that freed τ (M2) was superior to M1, we considered three additional models that freed $\{\mu_I, \mu_D\}$ and τ , together with each of the remaining parameters independently. None of these three models improved fits, indicating that no additional parameter combinations needed to be tested. We did not consider any models that freed parameters other than μ between I and D because we found no *a priori* justification for them.

Results

Model fits

Table 1 shows BPIC results for all the models tested. The best overall model, albeit by a small margin, was M2, which allowed mean drift rates (μ) and non-decision times (τ) to vary across experimental conditions. Figure 3 shows the quality of the fits obtained with this model. The match between the data and the model predictions is clear in each of the defective PDFs and histograms shown on the top row. These fits speak to the LBA model's ability to simultaneously account for observed RT distributions and choice probabilities during intertemporal choice.

The bottom row of Figure 3 shows the model fits with the RT distributions for both accumulators on the same axis to better

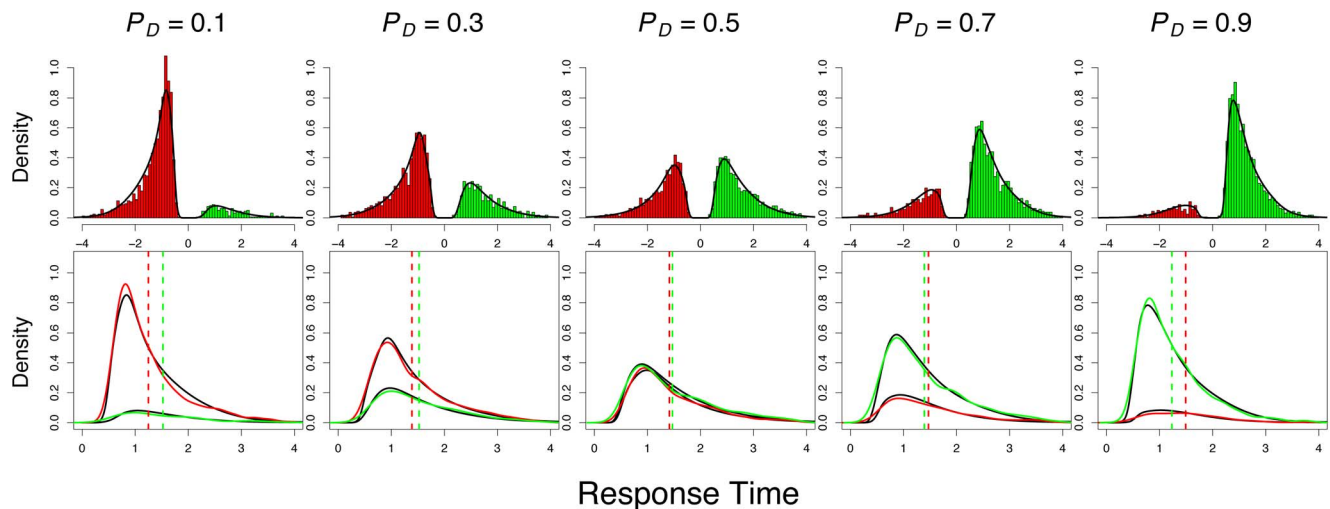


Figure 3. A comparison of model fits to empirical data. The top row shows the aggregated posterior predictive distribution (densities) overlaid on the aggregated empirical data (histograms). The response time distribution for the immediate reward is plotted on the left (i.e., with a negative axis; red), whereas the delayed reward is plotted on the right (green). The choice probability can be inferred by comparing the relative heights of the two distributions. The bottom row shows the same distributions as overlapping density functions with corresponding colors. The model fits are shown as black densities. The median response times for the empirical data are shown as the dashed vertical lines with corresponding colors.
doi:10.1371/journal.pone.0090138.g003

illustrate the relationship between choice probability and RT. As net value (i.e. $|V_D - V_I|$) increases, choices for the reward of less subjective value are slower relative to choices for the reward of greater value. This finding is illustrated by the increased separation of RT medians as the probability of choosing the delayed reward deviates further from $P_D=0.5$ (Figure 3). We confirmed the reliability of this pattern in the data by analyzing RT medians for choices that were consistent versus inconsistent with estimated subjective values. Specifically, we performed a rank-test on RT medians for consistent and inconsistent choices for all value conditions for which $P_D \neq 0.5$ and confirmed that inconsistent responses were slower relative than consistent responses in all conditions where $P_D \neq 0.5$ ($p = 5.883 \times 10^{-12}$). A similar relationship between RT and choice probability is commonly observed during perceptual decision making under stressed accuracy conditions. As choice probabilities deviate from $P_D=0.5$, the

means of the drift rate distributions ($\{\mu_I, \mu_D\}$) grow further apart (cf. [19,20]). Recall that $\mu_I = 1 - \mu_D$. However, subjects maintain an elevated accumulation bound (b) relative to the starting points ($\{a_I, a_D\}$). As a result, choices for the reward of less subjective value only occur in the improbable trials where the drift rate for the highest valued reward is unusually low, the drift rate for the lowest valued reward is unusually high, and subjects require more accumulated information before a decision can be made. If the starting points were large relative to the decision bound we would observe the opposite interdependence of RT and choice probabilities. Inconsistent choices would be faster than consistent choices, because fast errors occur when the initial choice bias drives the accumulation close to the decision bound before much evidence influences the decision. This value accumulation mechanism can explain why our model fitting results indicated that variability in b or A was not required to provide a good fit for these data (i.e., M1 and M2 performed better than M3, M4, M6, and M7).

Table 1. Mean Bayesian predictive information criterion fit statistics for each model variant we tested (standard deviations of the BPIC values computed across chains appear in parentheses).

Model	BPIC (std. dev.)	Constraint
M1	20101.37 (19.02)	μ
M2	20090.73 (26.91)	μ, τ
M3	20168.13 (71.25)	μ, A
M4	20197.20 (58.93)	μ, b
M5	20135.33 (46.86)	μ, σ
M6	20138.64 (98.01)	μ, τ, A
M7	20111.70 (44.94)	μ, τ, b
M8	20153.45 (28.34)	μ, τ, σ

For each model, the third column indicates the set of parameters assumed to vary across value conditions.

doi:10.1371/journal.pone.0090138.t001

Non-decision time

The best fitting model, M2, specifies a total of 13 subject-specific parameters, four more than the next best, and simplest model, M1. The four additional parameters modeled differences in non-decision time (τ) by value condition (P_D). To evaluate whether there was indeed systematic variance in non-decision time, we first inspected group-level estimates of τ , shown in the left panel of Figure 4. These parameter estimates showed a positive quadratic pattern centered at $P_D=0.5$. To test the quadratic relationship between τ and value, we performed a mixed-effects regression analysis with the nlme package in R (Jose Pinheiro et al., 2013), specifying subjects as random effects, and the regressor $(P_D - .5)^2$ as a predictor of subject-specific maximum *a posteriori* (MAP) estimates of τ . The results corroborated a positive quadratic relationship between τ estimates and value ($t(183) = 3.506$, $p = 6 \times 10^{-04}$), suggesting that there is an increase in valuation and/or motor-execution times as net value increases.

In the LBA model, τ functions as an offset term that captures differences in condition-wise RT that are not captured by the other parameters. The obvious empirical statistics related to average RT differences are condition-wise median and minimum RT. We therefore next tested whether (1) τ estimates were related to either median or minimum RT, and (2) whether minimum and/or median RT differed by value condition as suggested by the positive quadratic relationship between τ estimates and value.

The middle and right panels of Figure 4 plot subject-specific MAP estimates of τ against minimum and median RT, respectively. We conducted two mixed-effects regressions (using subjects as random effects) to determine whether τ estimates were related to minimum or median RT at each value condition. As hypothesized, τ estimates showed a significant linear relationship with minimum RT ($\beta=0.137$, $t(183)=9.716$, $p<1\times 10^{-16}$) and also a significant linear relationship with median RT ($\beta=0.029$, $t(183)=4.599$, $p<1\times 10^{-16}$).

Given these results, we next sought to determine whether RT differed across value conditions in the same manner as did estimates of τ . To test this hypothesis, we ran two additional mixed-effects regressions using the quadratic regressor $(P_D - .5)^2$ as a predictor of minimum and median RT (with subjects again as random effects). Recall that τ estimates showed a *positive* quadratic relationship with value. This relationship with value was not evident in analyses of minimum or median RT. Specifically, minimum RT did not show a significant quadratic relationship ($t(183) = -0.403$, $p=0.688$), and median RT showed a significant *negative* relationship with value ($t(183) = -6.169$, $p<1\times 10^{-16}$). We conclude from these results that neither minimum nor median RT alone can explain the positive quadratic relationship between τ and value. Taken together, our results suggest that the additional degrees of freedom in M2 allowed the model to capture within-subject changes in minimum RT and residual variance of median RT across value conditions.

Drift rates and value

To obtain a more precise characterization of M2 as a mechanistic theory of discounted value accumulation, we examined the relationship between independently estimated accumulation rates and discounted values. We first tested whether there

were systematic differences in group-level estimates of μ_D as a function of P_D . Group-level means of μ_D increased as a function of P_D . Specifically, we ran a mixed-effects regression of subject-specific MAP estimates of μ_D on P_D (using subjects as random-effects). This test revealed a significant positive linear relationship ($\beta=0.124$, $t(183)=30.587$, $p<1\times 10^{-16}$; Figure 5, left plot).

Next, we tested for a relationship between observed choice probabilities for the delayed reward and MAP estimates of μ_D and μ_I at the level of individual subjects. Specifically, we hypothesized that drift rates (μ) should be related to subjective value through a linear transform, with a slope parameter to account for differences in scale (i.e. μ_D and μ_I are restricted to be between 0 and 1 but V_D and V_I are in dollars with a mean of \$10) and an offset parameter to account for differences in drift rate and value means. We further reasoned that if drift rates were directly related to discounted subjective value then drift rates ought to be related to choice probability in the same way that differences in value are related to choice probability. Based on fits of the hyperbolic temporal discounting model (Equation 1) to choice outcomes, we already knew that a sigmoidal relationship (Equation 2) existed between subjective value (i.e. $\Delta V = V_D - V_I$) and choice probabilities (i.e. P_D). If modeled drift rates had the same relationship then we would expect a similar relationship between P_D , μ_D , and μ_I . However, μ_D and μ_I were not independent in our model specification. They were restricted such that $\mu_D + \mu_I = 1$. Thus, the difference in drift rates, $\Delta\mu = \mu_D - \mu_I$ reduces to a linear transformation of μ_D : $\Delta\mu = 2\mu_D - 1$. We therefore tested whether a sigmoidal relationship exists between subject- and condition-specific P_D and a linear transform of μ_D :

$$P_D = \frac{1}{1 + e^{-(\beta_1 + \beta_2 \mu_D)}}, \quad (6)$$

where β_1 and β_2 are subject specific parameters.

We tested for evidence to support Equation 6 in two ways. First, we performed a mixed-effects logistic regression using μ_D to predict P_D , with subjects as random effects. This analysis revealed a significant fit ($\beta_1 = -4.782$, $\beta_2 = 9.556$, $z = 47.03$, $p < 1 \times 10^{-16}$). The sigmoidal relationship is also clearly evident in the center plot of Figure 5 which plots P_D against μ_D . Next, we

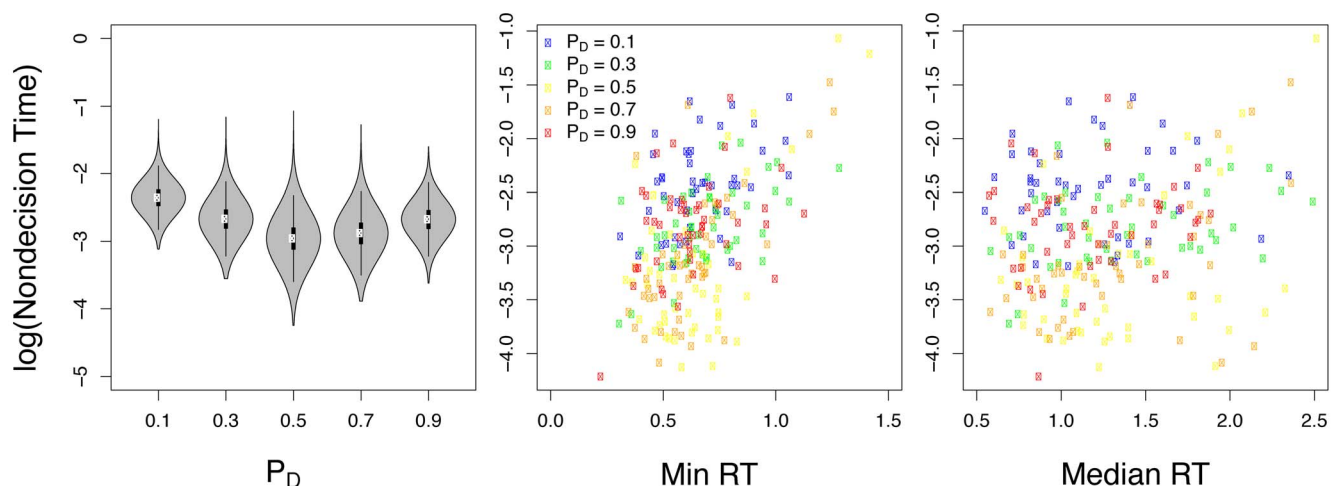


Figure 4. Relationships between model parameters, choice probability, and RT statistics. The left panel shows the estimated group level non-decision time parameter for each value condition. The middle and right panels show the maximum a posteriori (MAP) estimate for each subject's non-decision time parameter against their minimum and median response time, respectively. doi:10.1371/journal.pone.0090138.g004

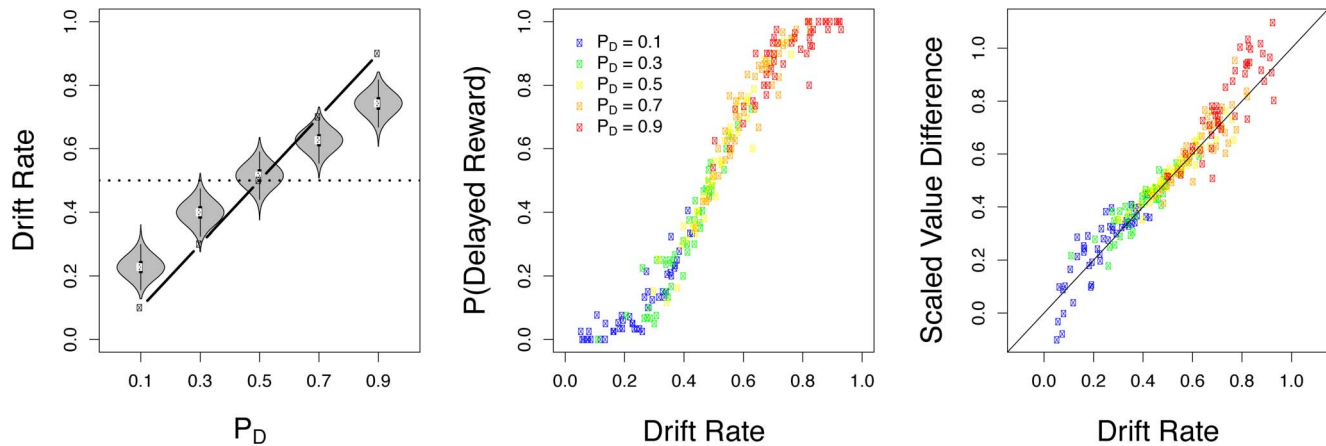


Figure 5. Relationships between model parameters, choice probability, and discounted value. The left panel shows the estimated group level drift rate for each value condition. The middle panel shows the maximum a posteriori (MAP) estimate for each subject's drift rate against observed choice probabilities for the delayed reward (P_D). The right panel shows the MAP estimate as a function of subject-specific discounted values for the delayed reward (V_D). doi:10.1371/journal.pone.0090138.g005

tested whether the relationship between P_D and ΔV (i.e., Equation 2) was directly related to the relationship between P_D and $\Delta\mu$ (i.e., Equation 6). If so, then the logistic function in both analyses should be equivalent and the following relationship should hold:

$$\mu_D = \frac{m(V_D - V_I) - \beta_1}{\beta_2}. \quad (7)$$

We estimated all of the parameters in Equation 7 from separate logistic regression analyses. Namely, β_1 and β_2 were obtained from fitting Equation 6, m derived from fitting Equation 2, and $V_D - V_I$ was obtained from best fits of Equation 1, all independently for every subject. In a group-level analysis, we used a mixed-effects regression with subjects as random effect and the right side of Equation 7 as the predictor. This analysis revealed a highly significant slope near unity ($\beta = 0.881$, $t(183) = 51.326$, $p < 1 \times 10^{-16}$). Together, these analyses indicated that there was a strong and direct relationship between drift rates and discounted value. Parameter estimates derived from fitting the LBA model to behavior therefore provided an independent means of estimating subjective values. Moreover, subjective values estimated from the LBA model corresponded closely with values estimated using a hyperbolic discounting model.

Generalizability of the relationship between drift rates and value

The previous analysis showed that a relationship existed between drift rates derived from LBA model fits and subjective value calculated based on a hyperbolic discount function. Of course, subjective value may actually be determined in a manner that differs in functional form from the hyperbolic equation (cf. [2]). Indeed, numerous functions have been proposed to account for delay discounting. In this final section, we aimed to show that drift rates derived from the LBA model are related to subjective value more generally; that is, that the relationship between drift rates and subjective value does not strictly depend on capturing subjective value using the hyperbolic discount function. To do so, we first fitted two additional discounting models to individual subjects' choices, substituting the right side of Equation 1 with exponential and "quasi-hyperbolic" value functions. For the

exponential discounting function, we assumed V_D to be given by:

$$V_D^e = re^{-\alpha t}, \quad (8)$$

where r is the delayed reward amount, α is the discount rate, and t is the delay. Similarly, for the quasi-hyperbolic discounting function, we assumed V_D to be given by:

$$V_D^{\beta\delta} = r\beta\delta^t, \quad (9)$$

where r is again the delayed reward amount, β is 1 when there is no delay or some fixed value between 0 and 1 when there is a delay, δ is between 0 and 1, and t is the delay (always greater than zero).

We then obtained estimates of $V_D - V_I$ using Equation 8 and Equation 9, as well as two independent estimates of m , one for each discounting function, from Equation 2, for every subject. Next, we ran mixed-effects regression analyses with subjects as random effect and the right side of Equation 7 as predictors of subject-specific drift rate estimates. The analysis using V_D^e revealed a significant slope near unity ($\beta = .938$, $t(183) = 25.662$, $p < 1 \times 10^{-16}$) and the analysis using $V_D^{\beta\delta}$ also revealed a significant positive slope ($\beta = .47$, $t(183) = 15.58$, $p < 1 \times 10^{-16}$). We therefore conclude that drift rates are related to subjective value independent of the specific functional form assumed for delay discounting.

Discussion

We have shown that intertemporal choice behavior is consistent with a process of discounted value accumulation instantiated by the LBA model. Our findings support the broader hypothesis that selecting among delayed rewards can be explained by a sequential sampling process that corresponds closely with mechanisms known to predict other types of choices (cf. [3]). Thus, perceptual and value-based decision making may depend on similar comparison and selection processes. It is interesting to speculate on whether this similarity reflects a direct correspondence between the cognitive and neural processes that support selection across diverse

domains or whether there is simply a common motif for action selection used in separate choice domains.

The LBA model we employed here has been used to explain neural activity during perceptual decision making (cf. [20–22]). Furthermore, sequential sampling processes such as that implemented by the LBA model provide a direct link between neural dynamics and decision making behavior. For example, evidence about visual motion is believed to be integrated in the lateral intraparietal (LIP) area, resulting in a progressive increase in LIP neuron firing rates that reflect the accumulation of sensory evidence and predict choice outcomes and response times [23,24]. Our results represent a first step in extending such findings from perceptual decision making tasks to generate quantitative predictions about discounted value accumulation in intertemporal choice. Moreover, our hierarchical LBA model fitting method might be particularly advantageous for studying the neural mechanisms of value accumulation when used in combination with the “joint modeling framework”, which was designed to simultaneously explain neuroimaging and choice data [25,26]. Using this framework, [25] have shown that it is possible to link neural and behavioral measures in a way that maps the mechanisms assumed by cognitive models directly to neural function. This approach allows for the specification of *a priori* predictions for how neural mechanisms should influence the modeled cognitive processes that presumably best explain behavior, providing a basis for hypothesis tests that are simultaneously informed by neural data, model parameters, and behavior.

Our results revealed a relationship between response time and choice probability, such that low probability choices are associated with increased response time. Similar results have been observed in previous studies using accumulation models to account for behavior in risk preference [27–29] and simple choice tasks [30–34]. Our observation that the LBA model can accommodate the relationship between response times and choice probability during intertemporal choice is thus consistent with previous findings and suggests that the LBA model might also be useful in accounting for behavior in other value-based decision domains.

Our best-fitting model included variability in drift rates and non-decision times across value conditions. This result violated our *a priori* expectation that drift rate variability across value conditions would be sufficient to account for our behavioral manipulation. Moreover, our results indicate that the model containing non-decision time variability performed only slightly better than the simplest model which was consistent with our theoretical expectation. Thus, from a purely theoretical standpoint, we favor the simplest model. However, for methodological consistency and empirical validity, we supported and analyzed the fits obtained from the best-fitting model. The BPIC statistic provides a measure of model quality that penalizes for the total number of parameters

in the model [18]. Relying on the BPIC statistic we corroborated our prediction that very few parameters needed to vary across conditions, but also found that the best model was not the simplest one. Future studies using the LBA could corroborate if in fact the simplest model generalizes better than the model with variability in non-decision time.

We showed that drift rates estimated with the model are directly related to discounted subjective values independently derived from behavioral models of intertemporal choice. The drift rate parameters of the LBA model therefore have a direct psychological interpretation and suggest a powerful means to estimate subjective values independent of assuming and fitting a specific form for temporal discounting (e.g. the hyperbolic model in Equation 1). In contrast, we are uncertain about how to interpret the variability in non-decision times across value conditions. On average, non-decision times decreased with increased difficulty. Moreover, although median RT showed a modest relationship with non-decision times, median RT increased with choice difficulty, reflecting a dependence on accumulation rates. Non-decision times also correlated strongly with minimum RT, which did not vary systematically across value conditions, but was highly variable across subjects. This suggests that our best-fitting model is reflecting the fact that minimum RT varies considerably across value conditions. It is unclear what to conclude from these findings. Our belief is that non-decision times capture idiosyncratic differences in choice strategies and valuation processes across subjects and that incorporating a parameter to absorb these trends improves model fits overall and the interpretability of drift rates more specifically.

In summary, we have demonstrated that an LBA model provides an excellent description of the choice process in intertemporal decision making. The model fits RT distributions, provides an explanation for interdependence between RT and choice probability, and can be interpreted in terms of value accumulation. These results validate the LBA model as a complementary tool to temporal discounting models for studying the cognitive and neural mechanisms of intertemporal choice. Because the LBA has been applied to a wide range of perceptual decision making tasks, our findings not only demonstrate that a general mechanism of evidence accumulation drives decision making but also support a common and analytically tractable framework for explaining it.

Author Contributions

Conceived and designed the experiments: CAR SMM. Performed the experiments: CAR. Analyzed the data: CAR BMT. Contributed reagents/materials/analysis tools: CAR BMT. Wrote the paper: CAR BMT SMM.

References

- Ainslie G (1991) Derivation of “rational” economic behavior from hyperbolic discount curves. *The American Economic Review* 81: 334–340.
- van den Bos W, McClure SM (2013) Towards a general model of temporal discounting. *Journal of the Experimental Analysis of Behavior* 99: 58–73.
- Kable JW, Glimcher PW (2009) The neurobiology of decision: Consensus and controversy. *Neuron* 63: 733–745.
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience* 9: 545–556.
- Gold JJ, Shadlen MN (2007) The neural basis of decision making. *Annual Review of Neuroscience* 30: 535–574.
- Ratcliff R (1978) A theory of memory retrieval. *Psychological Review* 85: 59–108.
- Usher M, McClelland JL (2001) On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review* 108: 550–592.
- Brown S, Heathcote A (2008) The simplest complete model of choice reaction time: Linear ballistic accumulation. *Cognitive Psychology* 57: 153–178.
- Mazur JE (1987) An adjusting procedure for studying delayed reinforcement. In: *Quantitative Analysis of Behavior: The Effects of Delay and Intervening Events on Reinforcement Value*, Hillsdale, New Jersey: Lawrence Erlbaum Associates. pp. 55–73.
- Shiffrin RM, Lee MD, Kim W, Wagenmakers EJ (2008) A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science* 32: 1248–1284.
- Lee MD, Wagenmakers EJ (2012) A course in Bayesian graphical modeling for cognitive science. Available: <http://www.ejwagenmakers.com/BayesCourse/BayesBookWeb.pdf>. Accessed 2012 Jan 1.
- Lee MD (2008) Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin and Review* 15: 1–15.

13. Rouder JN, Lu J (2005) An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin and Review* 12: 573–604.
14. Rouder JN, Lu J, Speckman P, Sun D, Jiang Y (2005) A hierarchical model for estimating response time distributions. *Psychonomic Bulletin and Review* 12: 195–223.
15. Turner BM, Sederberg PB, Brown S, Steyvers M (2013) A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods* 18: 368–384.
16. Gelman A, Carlin JB, Stern HS, Rubin DB (2004) *Bayesian Data Analysis*. New York, NY: Chapman and Hall.
17. ter Braak CJF (2006) A Markov chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing* 16: 239–249.
18. Ando T (2007) Bayesian predictive information criterion for the evaluation of hierarchical bayesian and empirical bayes models. *Biometrika* 94: 443–458.
19. Ratcliff R, Rouder JN (1998) Modeling response times for two-choice decisions. *Psychological Science* 9: 347–356.
20. Forstmann BU, Tittgemeyer M, Wagenmakers EJ, Derrfuss J, Imperati D, et al. (2011) The speed-accuracy tradeoff in the elderly brain: A structural model-based approach. *Journal of Neuroscience* 31: 17242–17249.
21. Forstmann BU, Dutilh G, Brown S, Neumann J, von Cramon DY, et al. (2008) Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences* 105: 17538–17542.
22. Forstmann BU, Anwander A, Schäfer A, Neumann J, Brown S, et al. (2010) Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences* 107: 15916–15920.
23. Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *Journal of Neurophysiology* 86: 1916–1936.
24. Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *The Journal of neuroscience* 22: 9475–9489.
25. Turner BM, Forstmann BU, Wagenmakers EJ, Brown SD, Sederberg PB, et al. (2013) A bayesian framework for simultaneously modeling neural and behavioral data. *NeuroImage*.
26. Turner BM (2013) Constraining cognitive abstractions through Bayesian modeling. In: Forstmann BU, Wagenmakers EJ, editors, *An introduction to model-based cognitive neuroscience*, New York, NY: Springer. p. 22.
27. Busemeyer JR (1985) Decision making under uncertainty: a comparison of simple scalability, fixed-sample, and sequential-sampling models. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11: 538.
28. Busemeyer JR, Townsend JT (1993) Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review* 100: 432.
29. Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MF, et al. (2012) Mechanisms underlying cortical activity during value-guided choice. *Nature neuroscience* 15: 470–476.
30. Krajbich I, Armel C, Rangel A (2010) Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience* 13: 1292–1298.
31. Krajbich I, Rangel A (2011) Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences* 108: 13852–13857.
32. Milosavljevic M, Malmaud J, Huth A, Koch C, Rangel A (2010) The drift diffusion model can account for value-based choice response times under high and low time pressure. *Judgement & Decision Making* 5: 437–449.
33. De Martino B, Fleming SM, Garrett N, Dolan RJ (2012) Confidence in value-based choice. *Nature neuroscience* 16: 105–110.
34. Towal RB, Mormann M, Koch C (2013) Simultaneous modeling of visual saliency and value computation improves predictions of economic choice. *Proceedings of the National Academy of Sciences* 110: E3858–E3867.