Full title Measuring Utility with Diffusion Models

Short title Measuring Utility with Diffusion Models

Renato Berlinghieri[1]†, Ian Krajbich [2,3,4]*†, Fabio Maccheroni[5]†, Massimo Marinacci[5]†, Marco Pirazzini[6]†

[1] Laboratory for Information and Decision Systems, Massachusetts Institute of Technology.

[2] Department of Psychology, The Ohio State University.

[3] Department of Economics, The Ohio State University.

[4] Department of Psychology, University of California, Los Angeles.

[5] Department of Decision Sciences, Bocconi University.

[6] Department of Computer Science, Yale University.

∗ Corresponding author - krajbich@ucla.edu

† These authors contributed equally to this work.

Abstract

The diffusion model (DDM) is a prominent account of how people make decisions. Many of these decisions involve comparing two alternatives based on differences of perceived stimulus magnitudes, such as economic values. Here, we propose a consistent estimator for the parameters of a DDM in such cases. This estimator allows us to derive decision thresholds, drift rates, and subjective percepts (i.e., utilities in economic choice) directly from the experimental data. This eliminates the need to measure these values separately or to assume specific functional forms for them. Our method also allows one to predict drift rates for comparisons that did not occur in the dataset. We apply the method to two datasets, one comparing probabilities of earning a fixed reward and one comparing objects of variable reward value. Our analysis indicates that both datasets conform well to the DDM. Interestingly, we find that utilities are linear in probability and slightly convex in reward.

Teaser Subjective representations of value are linear in reward probabilities and convex in rewards.

1 Introduction

There is a growing consensus that many decisions are made using a process where decision makers accumulate noisy evidence about the options until the net evidence exceeds a predetermined threshold [1, 2, 3]. This decision process is captured by the diffusion decision model [4], also known as the drift diffusion model (DDM). The DDM explains both choice and reaction time (RT) data on tasks including perceptual decisions [5] – e.g., recognition memory, brightness discrimination, dot motion direction – and economic (value-based) decisions – e.g., consumer choice [3, 6, 7, 8, 9, 10], risky choice [11, 12, 13, 14, 15, 16], intertemporal choice [17, 18, 13, 15], social preferences [19, 20, 15].

The DDM provides a mapping from subjective evidence to choice probabilities and RT distributions. It is, however, agnostic about the mapping from objective characteristics of the options/stimuli to subjective evidence. This poses a problem for modelers who do not know what decision makers perceive, only what they are shown. To get around this issue, modelers typically take one of two approaches.

The ideal approach is non-parametric [21]. Here, modelers identify trials that they think are roughly equivalent, bin them together into different *conditions*, and then use the DDM to estimate the average evidence (i.e., drift rate) for each condition separately [22]. This requires researchers to make assumptions about what should or should not affect drift rate and needs large amounts of data.

Other researchers sometimes assume functional forms of how stimulus features should affect drift rate [3]. In economics, this amounts to imposing a particular utility function. This can be problematic as it may not be obvious what kind of function is appropriate. For example, numerosity representation can be logarithmic or linear depending on whether there is a single array or a comparison of two arrays [23]. Ideally, one would not need to make such assumptions, but rather let the data speak for itself.

In contrast to the above approaches, we propose a more fundamental take on the problem. Baldassi et al. [24] provide a way to test the hypothesis that subjects follow a DDM in which the drift rate is based on the difference of perceived stimulus magnitudes - a difference-based DDM. We build on their ideas by constructing consistent estimators for DDM parameters, which allow us to derive the subjective evidence from behavior rather than assuming it. This is a semi-parametric approach in that we assume a difference-based DDM and the restrictions that come along with it, but we put no other restrictions on the drift-rate function.

With these estimators we are able to obtain drift rates for each pair of options and test how those rates relate to the stimulus features. An added bonus of this approach is that it allows us to derive decision thresholds and subjective evidence directly from the data, rather than using computationally heavy, model-fitting procedures. Our estimator can be calculated in microseconds, rather than the hours or days that it can take to fit the DDM to a group of subjects.

We test this approach using two choice tasks taken from Cavanagh et al. [25] and Shevlin et al. [26]. In both choice tasks, subjects were presented with two options and asked to choose the preferred one. In the [25] task, every option had the same two outcomes (win or lose) but varying probabilities. In the [26] task, every option had different, but deterministic, cash values. In the [25] task, option values were retrieved from memory, while in the [26] task they were constructed on the spot.

We selected these two tasks because subjects were trying to maximize along a single dimension (probability or reward amount) and we know the options' objective values along those dimensions. This allows us to estimate the mapping from objective stimuli to subjective evidence. While these two tasks both involve economic choice, our theory applies to any choice task where the drift rate depends on perceived differences in stimulus magnitudes of any kind. Our theory also applies to more standard economic choice with options that vary along multiple dimensions but that can be summarized using a single subjective value. While there are surely some differences between subjective- and objective-value tasks, there are also many commonalities. Both tasks have been shown to recruit the reward network [27, 28] and are well explained by difference-based DDM's [29]. Also, many models in economics make no distinction between objective and subjective values.

To preview the results, we find that in both datasets our estimates yield the expected relationships between drift rates, choice probabilities, and RT. More interestingly, we find remarkably linear relationships between the estimated drift rates – i.e., the subjective evidence – and the objective features of the stimuli, namely the differences in probabilities (Cavanagh) and in cash rewards (Shevlin). In terms of the subjective evidence for individual options, we find that they are on average linear in probabilities and approximately linear, but slightly convex, in reward size. Our estimator displays several advantages over existing methods for estimating utilities: it is more accurate than logistic regression, it is much faster than standard DDM fitting methods, and it is on more solid theoretical ground than either.

2 Results

2.1 Difference-based DDMs

Let A be a choice set consisting of at least three alternatives, with typical elements a, b, c, and d. A DDM is a model of binary comparison between pairs of alternatives a and b. According to this model, noisy evidence about alternatives is accumulated at average rate $\mu_{a,b} = -\mu_{b,a}$, known as the *drift rate*, until an alternative is selected when the evidence in its favor attains a posited threshold level $\lambda > 0$, called *decision threshold* or *barrier*.

Specifically:

- The *net evidence* in favor of a against b is given, at each $t > 0$, by
$$Z_{a,b}(t) = \mu_{a,b}\, t + \sigma B(t) \tag{1}$$
where B is a standard Brownian motion and σ is a noise coefficient (see, e.g., Oksendal [30]).
- Comparison ends when the Brownian motion $Z_{a,b}$ reaches either the barrier λ or $-\lambda$; so, the *decision time* is the random variable
$$\mathrm{DT}_{a,b} = \min\{t : Z_{a,b}(t) = \lambda \text{ or } Z_{a,b}(t) = -\lambda\}$$
Here we equate DT and RT, a simplifying assumption that we relax in the supplementary material.
- When comparison ends, at random time $\mathrm{DT}_{a,b}$, the agent selects a if the upper barrier λ has been reached, and selects b otherwise; so, the *decision rule* is the random variable
$$\mathrm{DO}_{a,b} = \begin{cases} a & \text{if } Z_{a,b}(\mathrm{DT}_{a,b}) = \lambda \\ b & \text{if } Z_{a,b}(\mathrm{DT}_{a,b}) = -\lambda \end{cases}$$

As the parameters λ, μ, and σ are unique up to a common positive scalar multiple, noise can be normalized by setting $\sigma = \sqrt{2}$. Thus, a DDM is uniquely identified by two parameters λ and μ. The *barrier parameter* λ is a number, while the *drift parameter* $\mu : A \times A \to \mathbb{R}$ is an antisymmetric function that associates to each pair (a, b) of alternatives the corresponding drift rate $\mu_{a,b}$. These parameters allow us to express choice probabilities, log-odds, and mean decision times by

$$P_{a,b} = \mathbb{P}[\mathrm{DO}_{a,b} = a] = \frac{1}{1 + e^{-\lambda\mu_{a,b}}} \tag{2}$$

$$\ell_{a,b} = \ln\frac{P_{a,b}}{1 - P_{a,b}} = \lambda\mu_{a,b} \tag{3}$$

$$\overline{\mathrm{DT}}_{a,b} = \mathbb{E}[\mathrm{DT}_{a,b}] = \lambda^2\varphi(\ell_{a,b}) \tag{4}$$

where $\varphi(x) = x^{-1}\tanh(x/2)$ for all real numbers x.

Next we introduce the class of DDMs that are relevant in difference-based decisions.

Definition 1. *A DDM with drift parameter μ is* difference-based *if, and only if, there exists a function $u : A \to \mathbb{R}$, called* utility*, such that*
$$\mu_{a,b} = u(a) - u(b) \tag{5}$$
for all a and b in A.

A difference-based DDM presupposes a transformation u of physical stimuli – with $u(a)$ being the perceived magnitude of stimulus a – such that the rate of evidence accumulation $\mu_{a,b}$ depends only upon perceived magnitude differences $u(a) - u(b)$. The function u represents sensation intensities in perceptual discrimination tasks and subjective values in economic tasks. The utility terminology that we adopt reflects our original emphasis on value differences, where the difference-based DDM is known as a value-based DDM.

The next simple characterization of difference-based DDMs highlights the properties that will guide our estimation exercise. In reading it, recall that a function $f : A \times A \to \mathbb{R}$ is *cyclic* when it satisfies the triangle equality
$$f_{a,b} = f_{a,c} + f_{c,b} \tag{6}$$
for all alternatives a, b, and c. For an antisymmetric function $f : A \times A \to \mathbb{R}$, the function $\tilde{f} : A \times A \to \mathbb{R}$ given by
$$\tilde{f}_{a,b} = \frac{1}{|A|}\sum_{c\in A}(f_{a,c} + f_{c,b})$$
is the least square cyclic approximation of f (see Proposition 5 of the supplementary material). For instance, $\tilde{\ell}$ is the least squares cyclic approximation of the log-odds function ℓ.

Proposition 1. *For a DDM with drift parameter μ, the following conditions are equivalent:*

(i) the DDM is difference-based;

(ii) the drift parameter μ is cyclic;

(iii) the log-odds function ℓ coincides with $\tilde{\ell}$;

(iv) for all $a \neq b$,

$$\lambda = \sqrt{\frac{1}{|A|(|A|-1)} \sum_{c \neq d} \frac{\overline{\mathrm{DT}}_{c,d}}{\varphi\left(\tilde{\ell}_{c,d}\right)}} \qquad \textit{and} \qquad \mu_{a,b} = \frac{\tilde{\ell}_{a,b}}{\sqrt{\frac{1}{|A|(|A|-1)} \sum_{c \neq d} \frac{\overline{\mathrm{DT}}_{c,d}}{\varphi\left(\tilde{\ell}_{c,d}\right)}}} \tag{7}$$

Thus, only the drift parameters that satisfy the triangle equality (6), i.e. cyclicity, admit a difference-based representation. The next corollary shows that the relation between drifts rates and utilities is indeed explicit. When μ is cyclic, we can then retrieve u from choice probabilities and mean decision times via (7).

Corollary 2. *If a DDM is difference-based with drift parameter μ, then, given any d in A, the function defined by*

$$u(a) = \mu_{a,d} \qquad \forall a \tag{8}$$

is, up to an additive constant, the only utility u which realizes (5) for μ.

It is important to remark that the choice of the reference alternative d in (8) is immaterial since our result implies that a change in the reference alternative can only shift u by an additive constant.

In contrast, if a drift parameter μ is not cyclic, there is no function $u : A \to \mathbb{R}$ such that (5) holds. In this case, the functions defined by (8) for different reference alternatives d might rank alternatives in different ways. Thus, they cannot be interpreted as utilities (see Example 1 in the supplementary material).

2.2 Value-consistent estimators

Consider an analyst who observes agents' pairwise choices several times, say n. The available observations produce *empirical mean decision times* $\overline{\mathrm{DT}}^n_{a,b}$, *empirical choice frequencies* $P^n_{a,b}$, and *empirical log-odds* $\ell^n_{a,b}$, for all $a \neq b$ in A. Baldassi et al. [24] have shown that through these statistics the analyst can test the hypothesis that data are generated by a difference-based DDM. Here, we assume that this hypothesis has not been rejected and we consider the problem of building an estimator of the parameters of the difference-based DDM which generates the data.

Definition 2. *An estimator $(\hat{\lambda}^n, \hat{\mu}^n)$ of a DDM (λ, μ) is:*

- statistically consistent *if, and only if, $(\hat{\lambda}^n, \hat{\mu}^n) \to (\lambda, \mu)$ almost surely as $n \to \infty$.*
- value-consistent *if, and only if, $\hat{\mu}^n$ is cyclic for all n.*

The previous analysis demonstrates that estimators that are not value-consistent are not conceptually appropriate for difference-based DDMs. Specifically, when an estimator is not value-consistent, it may be the case that the estimated value $\hat{\mu}^n_{a,b}$ of $\mu_{a,b}$ is different from the sum $\hat{\mu}^n_{a,c} + \hat{\mu}^n_{c,b}$ of the estimated values of $\mu_{a,c}$ and $\mu_{c,b}$, for some alternatives a, b and c. But, this means that the estimated DDM$(\hat{\lambda}^n, \hat{\mu})$ is not difference-based, contradicting the hypothesis that the DDM being estimated is difference-based. In other words, when an estimator is value-inconsistent there is no utility that represents the estimated drifts, so it does not produce the parameters of a difference-based DDM.

For instance, starting from equation (3) and assuming for simplicity that the analyst knows that $\lambda = 1$, she could estimate μ by replacing (theoretical) choice probabilities with empirical choice frequencies. This idea leads to the plug-in estimator (see, e.g., Wasserman [31])

$$\hat{\mu}^n_{a,b} = \ell^n_{a,b} = \ln \frac{P^n_{a,b}}{1 - P^n_{a,b}} \tag{9}$$

This estimator is statistically consistent but value-inconsistent as empirical log-odds are in general not cyclic (see Example 2 in the supplementary material).

To address this issue, we consider the *empirical adjusted log-odds* function $\tilde{\ell}^n : A \times A \to \mathbb{R}$ given by

$$\tilde{\ell}^n_{a,b} = \frac{1}{|A|}\sum_{c\in A}\left(\ell^n_{a,c} + \ell^n_{c,b}\right) = \frac{1}{|A|}\sum_{c\in A}\left(\ln\frac{P^n_{a,c}}{1-P^n_{a,c}} + \ln\frac{P^n_{c,b}}{1-P^n_{c,b}}\right) \tag{10}$$

for all a and b in A. This function, proposed by Baldassi et al. [24], is the least square cyclic approximation of ℓ^n and can be computed by using empirical choice frequencies. We can then plug in empirical mean decision times $\overline{\mathrm{DT}}^n_{a,b}$ and empirical adjusted log-odds $\tilde{\ell}^n_{a,b}$ into equation (7) to obtain a plug-in estimator.

Definition 3. *The* cyclic estimator $(\tilde{\lambda}^n, \tilde{\mu}^n)$ *of a difference-based DDM is given by*

$$\tilde{\lambda} = \sqrt{\frac{1}{|A|(|A|-1)}\sum_{c\neq d}\frac{\overline{\mathrm{DT}}^n_{c,d}}{\varphi\left(\tilde{\ell}^n_{c,d}\right)}} \qquad \textit{and} \qquad \tilde{\mu}_{a,b} = \frac{\tilde{\ell}^n_{a,b}}{\sqrt{\frac{1}{|A|(|A|-1)}\sum_{c\neq d}\frac{\overline{\mathrm{DT}}^n_{c,d}}{\varphi\left(\tilde{\ell}^n_{c,d}\right)}}} \qquad \forall a \neq b \tag{11}$$

The next proposition shows that this estimator has the desired properties and yields an immediate utility estimation.

Proposition 3. *The cyclic estimator is both statistically and value-consistent. Moreover, given any d in A, the function defined by*

$$\tilde{u}^n(a) = \tilde{\mu}^n_{a,d} \qquad \forall a \tag{12}$$

is, up to an additive constant, the only function that satisfies

$$\tilde{\mu}^n_{a,b} = \tilde{u}^n(a) - \tilde{u}^n(b) \qquad \forall a, b$$

The resulting utility estimation builds on the theoretical foundation given by the triangle equality as well as on the joint use of empirical choice frequencies and empirical decision times, as suggested by the approaches of Clithero [8] and Webb [32]. Again, as observed immediately after Corollary 2, the choice of the reference alternative d in (12) is immaterial because of the cyclicity of $\tilde{\mu}^n$.

We close with an important remark. The computation of the cyclic estimator seems to require, *prima facie*, a complete choice dataset in which all possible pairwise comparisons of distinct alternatives in A are observed. Yet, in the supplementary material we provide a simple iterative scheme that extends the estimator to incomplete datasets by leveraging on the premise that observations are generated by an underlying difference-based DDM and an estimator for such a DDM must satisfy the triangle equality (6).

2.3 Empirical analysis

In this section we complete the analysis of the cyclic estimator by applying it to behavioral data from two independent experiments, one with comparisons between all pairs of options (complete) and the other with missing comparisons (incomplete). In both cases we obtain very good fits and a "surprise." The surprise is that the utilities that we estimate mostly conform with those predicted by the most basic decision theoretic model of (expected) payoff maximization. The goodness of fit confirms the efficiency of estimating rather than postulating the drift parameters.

Datasets The first dataset comes from Cavanagh et al. [25] and is complete. In this study, participants chose between all pairs of six alternatives, eight times each, for a total of 120 trials per session. The alternatives were represented by Hiragana characters, each with a different probability (0.2, 0.3, 0.4, 0.6, 0.7, 0.8) of yielding a constant reward. These probabilities were *a priori* unknown to the participants but learned through experience in a prior training phase. Subjects completed two sessions, each with a different set of six characters.

The second dataset comes from Shevlin et al. [26] and is incomplete. In this study, participants chose between hundreds of unique alternatives. Each alternative was a 3×2 array of colored squares, with each color worth a different amount of money. There were 12 possible colors, each one worth a different amount of money; the values increased linearly across the color spectrum. The value of each array was the sum of the values of its six colored squares; this resulted in 27 possible array values. These arrays were divided into three value tiers: low, medium, and high value. In each trial, participants chose between two alternatives in the same tier. Like the Cavanagh et al. [25] study, the rewards for each color were *a priori* unknown to the participants but learned in a prior training phase. Before exclusions (as implemented in Shevlin et al. [26]), there were 135 trials per participant in the data we analyzed, or 45 trials per tier.

Both studies consisted of two parts: a *training phase* where the participants learned either the probabilities (Cavanagh) or the rewards (Shevlin), and a *test phase* where they chose from pairwise combinations of the alternatives. Training

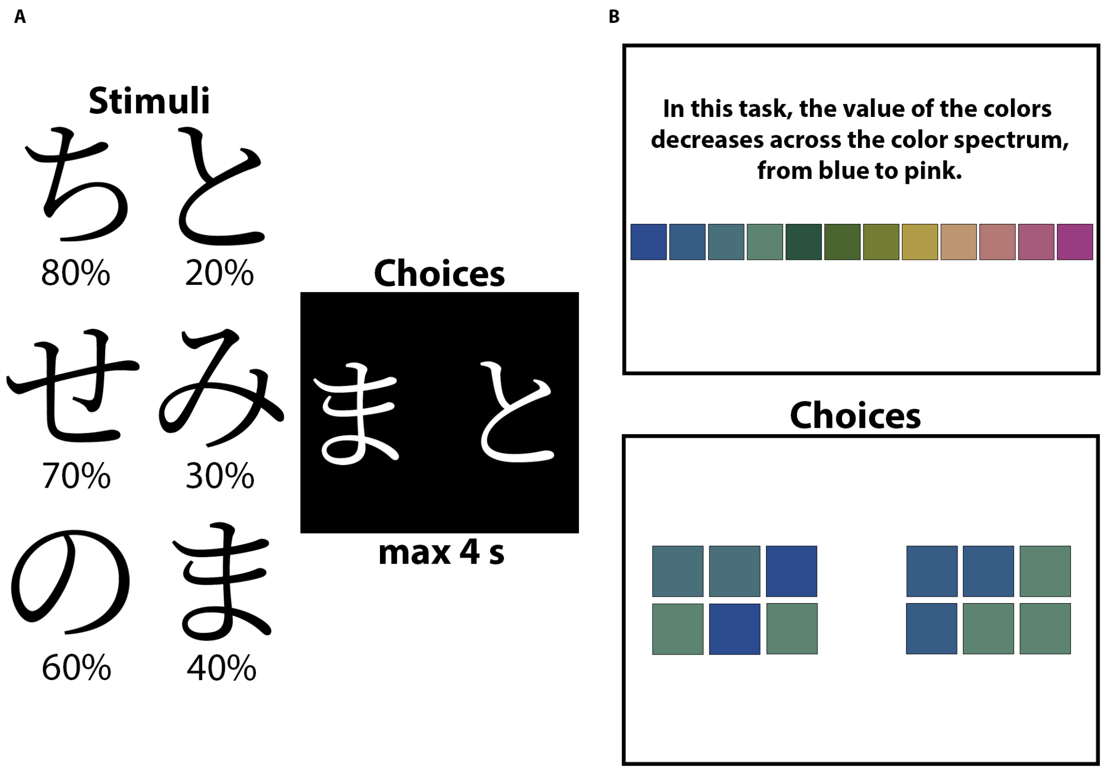

Figure 1: **Description of the experiments.** (A) In the Cavanagh et al. [25] task, participants chose between pairs of six Hiragana characters that were trained to be associated with reward probabilities as indicated on the left. (B) In the Shevlin et al. [26] task, participants chose between 3×2 arrays of colored squares. Each color represents an integer point value from 1 to 12. The point values increased from blue to pink or pink to blue; this was counterbalanced across participants.

stopped when there was sufficient discrimination between correct and incorrect choices. In the test phase, participants chose between pairs of alternatives without receiving any feedback, so that choice was purely based on what had previously been learned. Here, we only analyze the test-phase trials. See Figure 1 for further details on the tasks in these two experiments.

In summary, we study two tasks where participants learned the value of different alternatives. We chose these datasets because they present the proof of two concepts: the cyclic estimator works for both complete and incomplete datasets, and for both probabilistic and deterministic rewards. They also allow us to test how the utilities reflect the objective values.

Analysis The cyclic estimator – equations (11) and (12) and their generalizations provided in the supplementary material – allows us to directly estimate the drift rates (utility differences) and boundary separation (threshold) of a DDM from observable choice and RT data. In the supplementary material (Figs. S3-S4) we provide evidence that the cyclic estimator accurately recovers drift rates and boundary separations in synthetic data simulated from a DDM with typical parameters. Next, we investigate the performance of the cyclic estimator on the empirical data.

The first behavioral quantity predicted by the DDM is the choice probability for each alternative pair (psychometric function). The estimated model seems to capture the psychometric behavior of the subjects from both experiments with great precision (Figure 2)

The second behavioral quantity predicted by the model is the mean DT for each binary choice (chronometric function). Before investigating the chronometric function goodness of fit, we need to deal with a systematic bias, pointed out in the literature (see, e.g., [33, 25, 26]), that high overall-value decisions ("win-win" comparisons) are faster than predicted by the vanilla DDM. In the Cavanagh et al. [25] task, these are the three comparisons between $\{0.8, 0.7\}$, $\{0.8, 0.6\}$ and $\{0.7, 0.6\}$. In the Shevlin et al. [26] task, these are the two comparisons between the highest valued arrays within each of the three value tiers. In the plots we highlight these higher value trials. With this caveat, the chronometric function predicted by the estimated model fits the behavioral data from both experiments nicely, as Figure 3 shows.

In the supplementary material (Fig. S5) we additionally present fits of the DDM to quantiles of the RT distributions, conditional on correct and error responses.

Cavanagh et al. [25] model the DDM using the *difference of reward probabilities* – or, equivalently, of *expected payoffs* – of the alternatives as drift for the binary comparisons. Analogously, Shevlin et al. [26] model the DDM using the *difference of monetary payoffs* as drift. We do not make these assumptions. To compare the two approaches, in Figure 4, we examine the relationship between the objective differences in reward probabilities/monetary payoffs that they use as drifts (Δv) and our *differences in estimated utility* (Δu). Remarkably, the data show a linear relationship between these two quantities in the Cavanagh data and a slightly convex relationship in the Shevlin data.

This linear relationship between differences in payoffs and differences in estimated utilities is perhaps the most striking empirical finding of the paper. This discovery, made possible by the introduction of the cyclic estimator, suggests that the subjects in the experiment actually learned the (expected) payoffs and maximized them. This optimizing behavior was assumed as an hypothesis by Cavanagh et al. [25] and Shevlin et al. [26]. Our finding thus supports their assumptions.

We derive absolute utilities from relative utilities by fixing the lowest payoff alternative as a reference (Figure 5). In the Cavanagh et al. [25] task the relation between utilities and true probabilities is strikingly linear. This is consistent with the expected utility maximization hypothesis. In the Shevlin et al. [26] task the relation is piecewise linear, with the slope slightly but significantly increasing with objective value. This increasing slope is consistent with the findings of Shevlin et al. [26]. We established this with a linear regression of the utilities on the value range (0 to 8) within each value condition interacted with the value condition (low $= -1$, medium $= 0$, high $= 1$). The interaction between value and condition was significantly positive ($t(23) = 3.32$, $p = 0.003$).

We can also compare the utilities from the cyclic estimator to the utilities obtained from other methods. The traditional approach is to estimate the utilities using a logistic regression on the choice data, with dummy variables for each option. A newer and more computationally involved approach is to run an analogous regression for drift rates within the DDM. This is accomplished using the software package HDDM, which is a Bayesian hierarchical method that uses group-level choice and RT data to fit the DDM [34]. Comparing the cyclic estimator to the logistic and HDDM estimators on the empirical data, we find that there is good agreement between all three methods, but that the cyclic estimator seems to be more linear, i.e., less prone to overfitting, than the logistic estimator (Fig. 6).

Finally, we also analyzed the data from individual subjects in each study. As expected, there is heterogeneity in the linearity of the relation between utilities and true values, particularly in the Cavanagh et al. [25] data.

The individual-level analyses reveal an advantage of our cyclic estimator over other estimators like logistic regression and HDDM. If we assume that the correct estimates should be linear, then we can calculate the mean absolute error of the utility estimates for the different estimators. Our cyclic estimator has a significantly lower mean absolute error than a logistic regression estimator for the Shevlin et al. [26] dataset ($p = 10^{-10}$, $p = 10^{-12}$, $p = 10^{-13}$) though not for the Cavanagh et al. [25] dataset ($p = 0.18$). And while the cyclic estimator does not outperform the HDDM estimator in mean absolute error, it uses a lot less data (single subject and choice data only) and runs in a matter of microseconds. See the supplementary material (Figs. S1-S2) for more detail.

3 Discussion

In this paper we used a principled approach to consistently estimate the parameters of a difference-based DDM. This allowed us to investigate how reward probabilities and amounts affect evidence accumulation (i.e., drift rates). We observed almost linear relations between differences in utilities and differences in both reward probabilities and amounts. More specifically, we observed that utilities are linear in probability and slightly convex in reward amount.

Our approach yields trial-level drift rates that show a consistent relationship with choice probabilities and RTs, as expected with the DDM. Throughout the paper, we have equated RT and DT. In reality, RT usually consists of DT as well as a non-decision time. In the supplementary material (Fig. S6) we show that accounting for non-decision time does indeed improve the quality of our fits. We also present an estimator that, while untested, might provide an alternative way to estimate non-decision times; we leave the validation of that estimator to future research. It is important to note that fitting the DDM often includes other parameters, such as starting point and across-trial variability in drift rate and starting point. Without these parameters, our estimators may suffer on datasets with substantial non-decision times, response biases, or slow/fast errors. In this article we assume that the modeler has concluded that a simple difference-based DDM is appropriate for their data. Our theory cannot be used to estimate other models or to compare them to the DDM. If the modeler wants to establish whether the DDM is appropriate for their data, they can use the test proposed in Baldassi et al. [24].

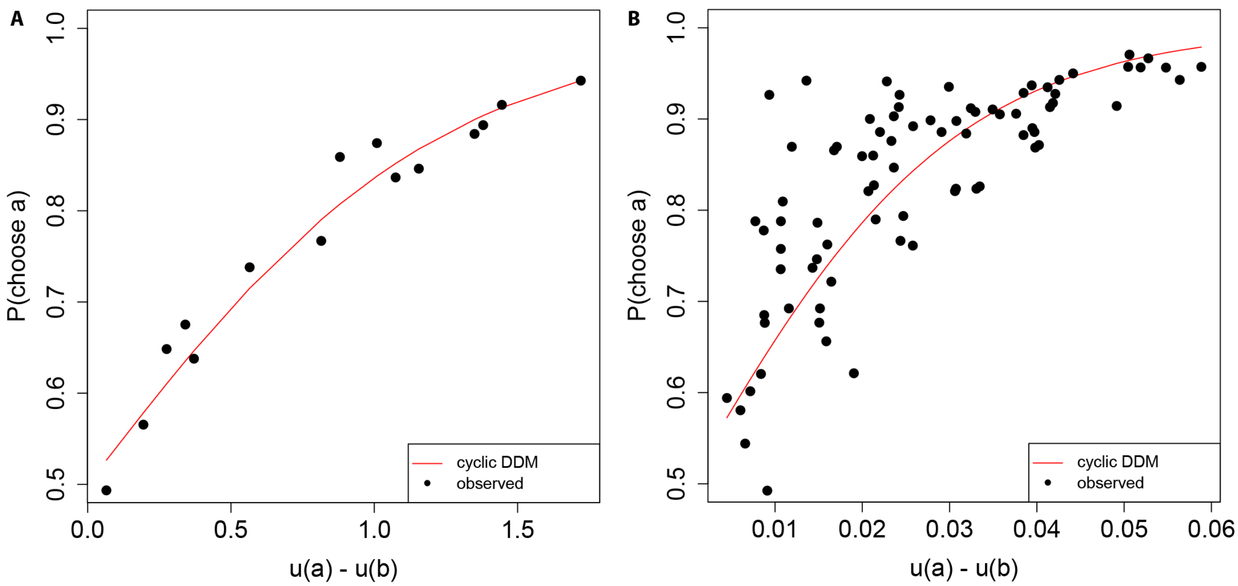

Figure 2: **Observed and estimated choice probabilities.** The probability of choosing Option A over Option B as a function of the difference in their estimated utilities, for (A) the Cavanagh et al. [25] data, and (B) the Shevlin et al. [26] data. The estimated utilities are calculated from equations (10), (11), and (12), using all comparisons $\{a, c\}$ and $\{c, b\}$. The choice probabilities are either the average empirical probabilities from all direct comparisons $\{a, b\}$ (black dots), or the DDM predicted probabilities calculated from Equation (2) (red line) using the estimated parameters.

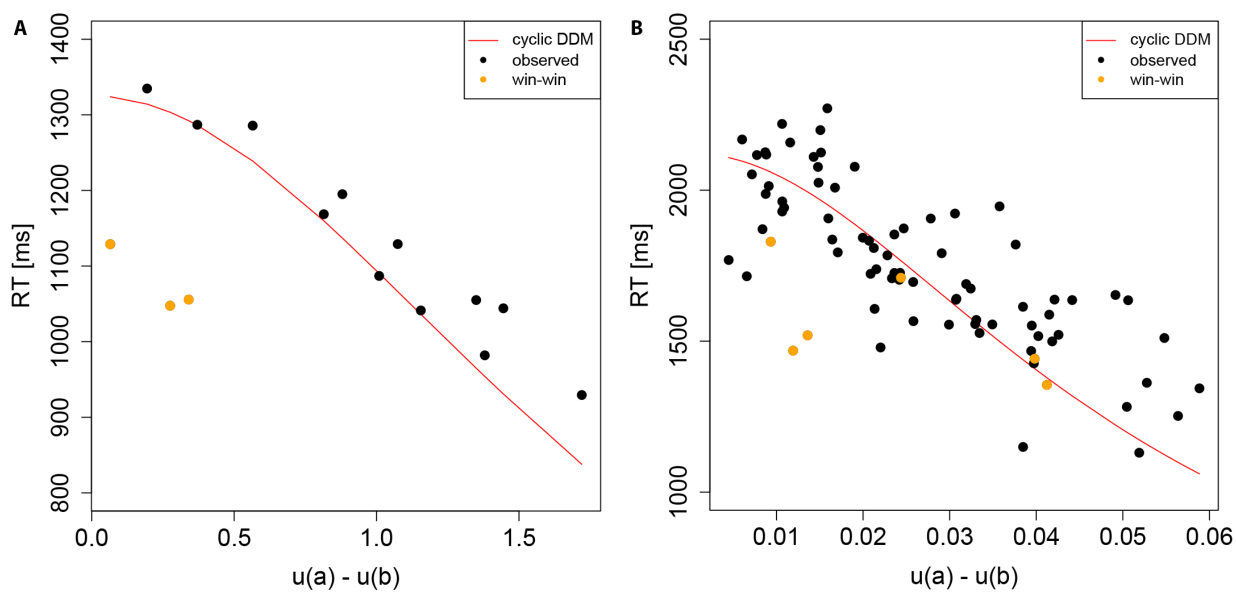

Figure 3: **Observed and estimated mean response times.** The mean response time (RT; in milliseconds) in choices between Option A and Option B as a function of the difference in their estimated utilities, for (A) the Cavanagh et al. [25] data, and (B) the Shevlin et al. [26] data. The estimated utilities are calculated from equations (10), (11), and (12), using all comparisons $\{a, c\}$ and $\{c, b\}$. The RTs are either the average RT from all direct comparisons $\{a, b\}$ (black dots), or the DDM predicted decision times calculated from Equation (4) (red line) using the estimated parameters. The orange dots indicate "win-win" trials; trials with the highest valued pairs of alternatives.

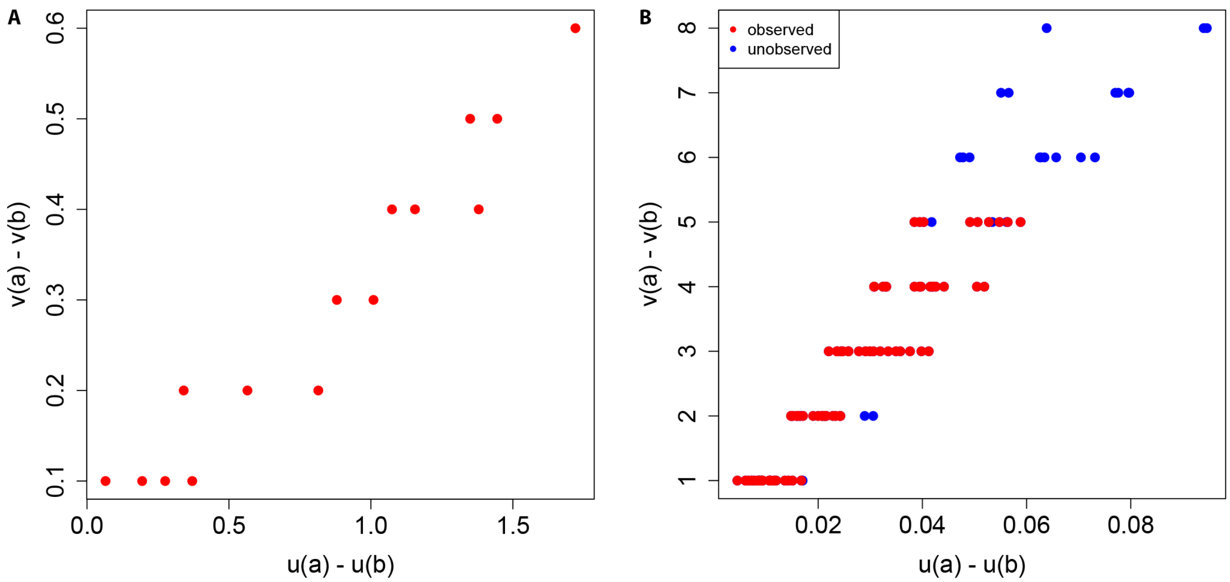

Figure 4: **Differences in objective value and estimated utility.** The difference in objective value (v) as a function of the difference in estimated utility (u) for each pair of options in (A) the Cavanagh et al. [25] data and (B) the Shevlin et al. [26] data. The estimated utilities are calculated from equations (10), (11), and (12), using all comparisons $\{a, c\}$ and $\{c, b\}$. The true values are the probabilities of reward in the Cavanagh et al. [25] task and the point values in the Shevlin et al. [26] task. This figure contains additional points (in blue) for comparisons $\{a, b\}$ that were not in the experiment, but for which we could still estimate Δu had they occurred, based on other comparisons.

Our method is appealing because it is free from both a priori behavioral and statistical assumptions (aside from the assumption that data are generated by a difference-based DDM), but still fits both choice probabilities and RTs well. Rather than estimating drift rates for a set of conditions, which are based on assumptions about which trials are similar, our approach estimates a unique drift rate for each comparison. Moreover, it is able to do so out-of-sample, using other comparisons in the dataset. One can even extrapolate the drift rates for comparisons not present in the data, as we did for the data of Shevlin et al. [26].

The ability to estimate drift rates for individual comparisons is a major advantage, particularly in economic choice, where drift rates depend on subjective values. Modelers typically need to measure the stimulus features in their experiments. In perceptual decisions this is not too burdensome: typically, these features are objective and controlled by the experimenter (e.g., number, size, brightness). In economic decisions, however, the features are often subjective and a product of the subjects' preferences. As a result, many of these experiments involve lengthy rating/evaluation tasks – in addition to the comparison tasks – to measure the subjective values of the stimuli [35, 9, 29, 36]. Moreover, these value measurements are at odds with a "preference discovery" interpretation of the value-based DDM in which decision makers accumulate noisy evidence about the options' subjective values that, *ex ante*, they only imperfectly know [37]. By using our method, researchers might be able to avoid collecting independent evaluations of all the stimuli in their experiments. However, it is important to note that in order for this method to work, one must have choice probabilities (not 0 or 1) for at least some comparisons. In other words, some comparisons must be presented multiple times within an experiment, or data must be aggregated across subjects facing overlapping comparisons.

Our empirical analysis of the Cavanagh et al. [25] dataset revealed a linear function from reward probabilities to drift rates, on average. This contrasts with some of the literature on non-linear probability weighting, where explicit probabilities are generally inverse-S weighted (overweighting of small probabilities and underweighting of large probabilities) as in Prospect Theory [38], or where learned probabilities are generally S weighted [39] (overweighting of large probabilities and underweighting of small probabilities). In most of this work, decision makers face trade-offs between probability and reward size. Thus, it is possible that the linear relationship that we identified is due to the fixed rewards in the Cavanagh et al. [25] task (or perhaps the extensive training). It is also worth noting that the functions were typically not linear at the individual level (see supplementary material, Fig. S1). This is consistent with the literature using this task [40].

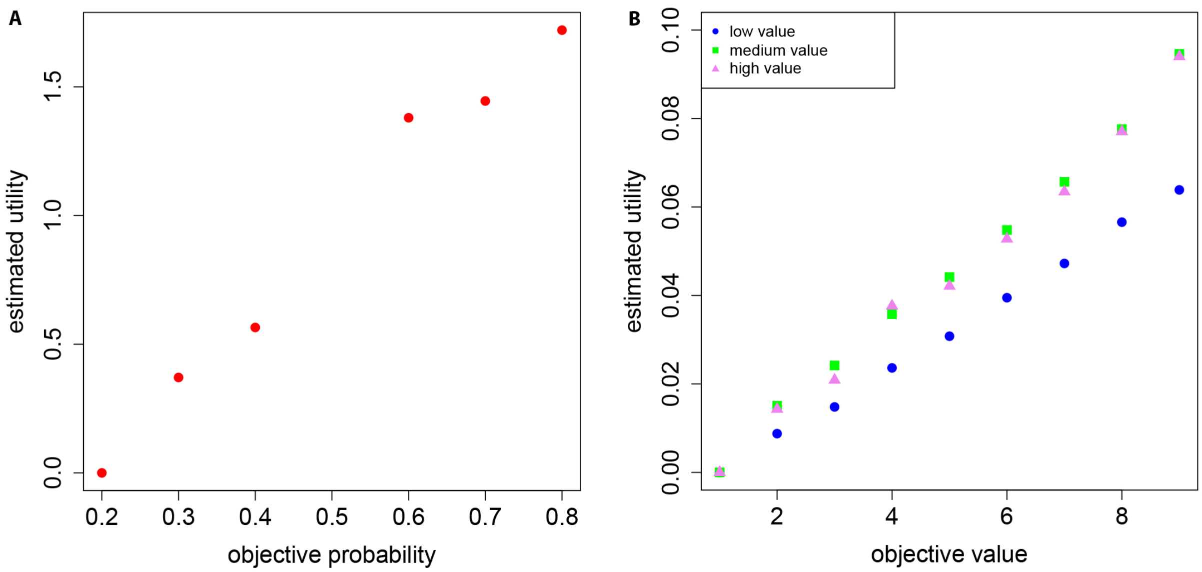

Figure 5: **Estimated utility functions.** Utility vs. probability for (A) the Cavanagh et al. [25] data, and vs. value for (B) the Shevlin et al. [26] data. Analogous to Figure 4, the estimated utilities are calculated from equations (10), (11), and (12), using all comparisons $\{a, c\}$ and $\{c, b\}$, while the objective probabilities are the probabilities of reward in the Cavanagh et al. [25] task and the objective values are the point values in the Shevlin et al. [26] task. In the Shevlin et al. [26] task we plot the data for the three different value tiers separately. Because utilities are on an arbitrary scale, we fix the lowest utility within each set to zero.

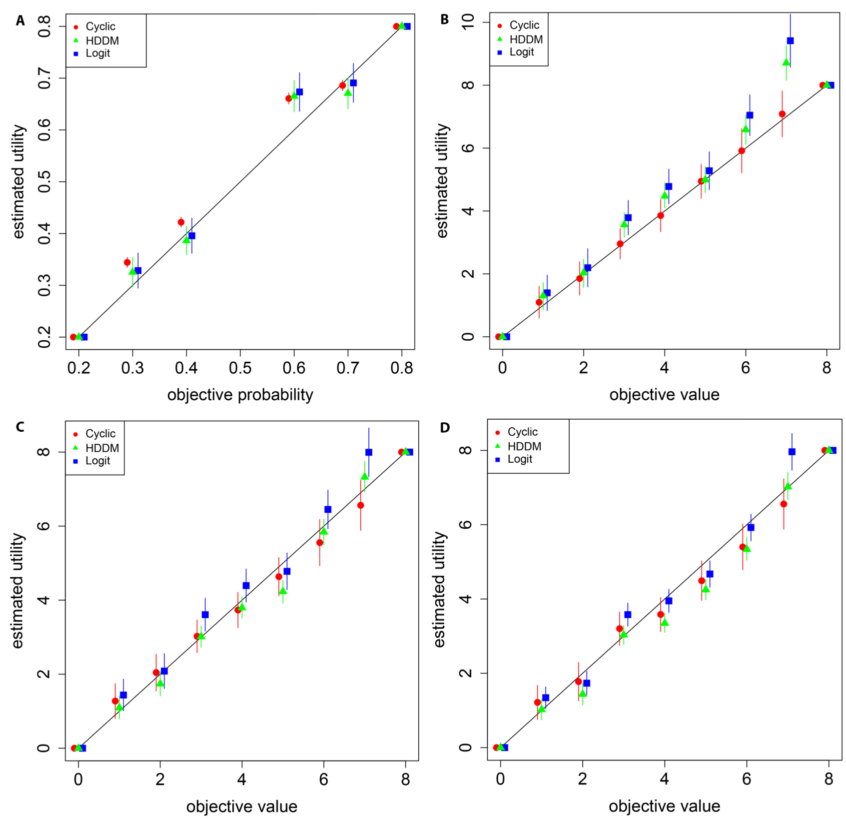

Figure 6: **Model comparison of the estimated utility functions.** Utility vs. probability for the (A) Cavanagh et al. [25] data, and vs. value for the (B-D) Shevlin et al. [26] data in Low, Medium, and High value conditions respectively. These plots display a linear relationship (diagonal black line) as well as the fits from our cyclic estimator, HDDM, and a logistic regression. Because utility is on an arbitrary scale, we forced the lowest and highest utilities to be accurate and then measured the relative locations of the remaining estimates. For the logit and cyclic estimators, the bars are 95% confidence intervals; for the HDDM the bars are 95% highest density intervals.

Our findings are also consistent with the previous findings in Shevlin et al. [26]: the slightly convex utility function is consistent with their findings that drift rates are higher at high values. At the same time, when focusing within a value level (low, medium, or high), we find a quite linear function from objective value differences to drift rate.

There is one notable way in which the data deviate from our estimated DDM. Comparisons between the highest-value alternatives are systematically faster than expected. This is a well-documented phenomenon [25, 41, 42]. There are several explanations for this phenomenon, including increased dopamine activity for high-value options and greater drift-rate variability due to attentional shifts between high-value options [43, 42]. These are deviations from the standard "plain vanilla" DDM that affect decision times more than choice probabilities (see the supplementary material, Fig. S11). For this reason, they are highlighted only in the figures concerning RT, where they are labeled as "win-win comparisons."

Our empirical tests of the cyclic estimator were limited to objective-value tasks where options differ along a single stimulus dimension, namely reward size and probability. We do not believe that this limits our contribution, as the conditions apply to any difference-based DDM, and we know that difference-based DDMs work well in subjective-value tasks as well. The reason that we chose to focus on objective-value tasks is because there is a ground truth that we can compare our utility estimates to. In subjective-value tasks there is no ground truth. If we found non-linear relations between estimated and reported utilities, we would not know if that was an issue with the estimator or with the reports. So, while we advocate for using our estimator to infer utilities in subjective-value tasks, we did not think it was appropriate to use such tasks to validate the estimator in this article. That is a next step.

The use of DDM and other sequential sampling models has a long history in statistics [44] and cognitive psychology [4]. There is also a growing literature applying such models to economic choice, most notably with decision field theory [14]. Recent work in economics has explored the optimality of sequential sampling models and their link to strength of preference [45, 46, 8, 47, 48, 15, 49, 37, 50]. One advantage of the DDM is that it jointly predicts choice probabilities and RT. Thus, the analysts' access to RT data should improve their ability to infer preferences or beliefs [46, 8, 15, 51, 32, 52, 49], as we also observed here with the HDDM estimator. The DDM can also be used to decompose the decision process into evaluations of the alternatives, prior biases, speed-accuracy motivation, and so on [53, 54]. In more recent years, these models have become very popular in cognitive neuroscience and in decision neuroscience [55, 1]. They can also account for eye-tracking and brain imaging data [56, 3, 57, 58, 35, 20, 59, 60]. Our ability to estimate trial-level drift rates provides researchers with more precise inputs for modeling such data [61]. We hope that these findings fuel further work on the DDM in economic choice (and beyond) and that these methods will be used to better understand the relation between objective and subjective evidence.

4 Materials and Methods

4.1 Cavanagh task

The first dataset comes from Cavanagh et al. [25]. The study included 20 subjects (12 male, mean age of 20 years) from the Brown University community, who were rewarded with either 20 USD or extra course credit.

Subjects performed the task twice using different non-overlapping character sets. The mapping from characters to probabilities was randomized across subjects. The data were combined across the two sessions.

Each session began with a training phase in which the probabilities were learned through reinforcement. During the training phase subjects were presented with one of three stimulus pairs with the following reinforcement probabilities 0.8/0.2, 0.7/0.3, or 0.6/0.4. Subjects went through one to six blocks of training (60 stimuli each) until they made the correct choice on 0.65, 0.6, and 0.5 of the trials in the 0.8/0.2, 0.7/0.3, and 0.6/0.4 pairs, respectively. Subjects who did not satisfy all these requirements by the sixth block proceeded to the test phase regardless.

In the test phase, subjects chose between all pairs of six alternatives, eight times each, for a total of 120 trials per session, or 240 trials total. Trials began with a fixation cross for 1,000 ms. Stimuli were presented for a maximum of 4,000 ms, and disappeared as soon as a choice was made. Participants then saw a blank black screen for 1,000 ms. There was no feedback in the testing phase.

4.2 Shevlin task

The second dataset comes from Shevlin et al. [26]. The study included 70 subjects from the Ohio State University community, who were rewarded with a show-up fee of 5 USD as well as a bonus payment based on their performance (30 points per 1 USD; mean earnings: 9.81 USD).

Subjects completed two phases of the study. In the first phase, participants viewed a spectrum of 12 distinct colors. Subjects were told that the value of these colors was either increasing or decreasing from left to right. Subjects then completed a training phase, in which they chose between two arrays, each composed of six colored squares. The values for each colored square ranged from 1 to 12. After each choice, subjects saw the values of both arrays (equal to the sum of the colored squares), as well as their total earnings. Subjects first completed a block of 30 trials. If they reached or surpassed 0.70 accuracy, they proceeded to phase 2; otherwise, they completed another block of 30 trials. This process was repeated until each subject achieved 0.70 accuracy or completed six training blocks.

In phase 2, subjects faced 270 binary-choice trials. Trials were constructed by first creating stimulus pairs for the middle-value condition and then subtracting or adding a constant value of 4 to every square. The arrays were constructed so that the colored squares were never all the same color, but there were no other restrictions. Additionally, the value difference in each trial was always between 1 and 5 points.

In the original study there were blocks of 15 trials, some with trials all from the same tier and some with trials from different tiers (five from each). Here we only analyze the latter blocks as they are the most natural. In the supplementary material (Figs. S7-S10) we also examine the other blocks.

In order to properly constrain subjects' earnings, they started out with a deficit of 10,500 points. Subjects who ended the study with a negative balance only received the show-up fee. Subjects were informed of this deficit and the conversion rate at the beginning of the study.

References

[1] Michael N Shadlen and Daphna Shohamy. Decision making and sequential sampling from memory. *Neuron*, 90(5):927–939, 2016.

[2] Jerome R Busemeyer, Sebastian Gluth, Jörg Rieskamp, and Brandon M Turner. Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in Cognitive Sciences*, 23(3):251–263, 2019.

[3] Ian Krajbich, Carrie Armel, and Antonio Rangel. Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10):1292–1298, 2010.

[4] Roger Ratcliff. A theory of memory retrieval. *Psychological Review*, 85(2):59–108, 1978.

[5] Roger Ratcliff, Philip L Smith, Scott D Brown, and Gail McKoon. Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4):260–281, 2016.

[6] Milica Milosavljevic, Jonathan Malmaud, Alexander Huth, Christof Koch, and Antonio Rangel. The drift diffusion model can account for value-based choice response times under high and low time pressure. *Judgment and Decision Making*, 5(6):437–449, 2010.

[7] Geoffrey Fisher. An attentional drift diffusion model over binary-attribute choice. *Cognition*, 168:34–45, 2017.

[8] John A Clithero. Improving out-of-sample predictions using response times and a model of the decision process. *Journal of Economic Behavior & Organization*, 148:344–375, 2018.

[9] Pradyumna Sepulveda, Marius Usher, Ned Davies, Amy A Benson, Pietro Ortoleva, and Benedetto De Martino. Visual attention modulates the integration of goal-relevant evidence and not value. *Elife*, 9:e60705, 2020.

[10] Khai Xiang Chiong, Matthew Shum, Ryan Webb, and Richard Chen. Combining choice and response time data: A drift-diffusion model of mobile advertisements. *Management Science*, 2023.

[11] Moshe Glickman, Orian Sharoni, Dino J Levy, Ernst Niebur, Veit Stuphorn, and Marius Usher. The formation of preference in risky choice. *PLoS Computational Biology*, 15(8):e1007201, 2019.

[12] Veronika Zilker and Thorsten Pachur. Nonlinear probability weighting can reflect attentional biases in sequential sampling. *Psychological Review*, 129(5):949–975, 2022.

[13] Wenjia Joyce Zhao, Adele Diederich, Jennifer S Trueblood, and Sudeep Bhatia. Automatic biases in intertemporal choice. *Psychonomic Bulletin & Review*, 26(2):661–668, 2019.

[14] Jerome R Busemeyer and Adele Diederich. Survey of decision field theory. *Mathematical Social Sciences*, 43(3):345–370, 2002.

[15] Arkady Konovalov and Ian Krajbich. Revealed strength of preference: Inference from response times. *Judgment and Decision Making*, 14(4):381–394, 2019.

[16] Feng Sheng, Arjun Ramakrishnan, Darsol Seok, Wenjia Joyce Zhao, Samuel Thelaus, Puti Cen, and Michael Louis Platt. Decomposing loss aversion from gaze allocation and pupil dilation. *Proceedings of the National Academy of Sciences*, 117(21):11356–11363, 2020.

[17] Junyi Dai and Jerome R Busemeyer. A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General*, 143(4):1489–1514, 2014.

[18] Dianna R Amasino, Nicolette J Sullivan, Rachel E Kranton, and Scott A Huettel. Amount and time exert independent influences on intertemporal choice. *Nature Human Behaviour*, 3(4):383–392, 2019.

[19] Ian Krajbich, Todd Hare, Bjoern Bartling, Yosuke Morishima, and Ernst Fehr. A common mechanism underlying food choice and social decisions. *PLoS Computational Biology*, 11(10):e1004371, 2015.

[20] Cendri A Hutcherson, Benjamin Bushong, and Antonio Rangel. A neurocomputational model of altruistic choice and its implications. *Neuron*, 87(2):451–462, 2015.

[21] Mohammed Abdellaoui. Parameter-free elicitation of utility and probability weighting functions. *Management science*, 46(11):1497–1512, 2000.

[22] Roger Ratcliff and Philip L Smith. A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111(2):333–367, 2004.

[23] Roger Ratcliff and Gail McKoon. Modeling numerosity representation with an integrated diffusion model. *Psychological Review*, 125(2):183–217, 2018.

[24] Carlo Baldassi, Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, and Marco Pirazzini. A behavioral characterization of the drift diffusion model and its multialternative extension for choice under time pressure. *Management Science*, 66(11):5075–5093, 2020.

[25] James F Cavanagh, Thomas V Wiecki, Angad Kochar, and Michael J Frank. Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, 143(4):1476–1488, 2014.

[26] Blair RK Shevlin, Stephanie M Smith, Jan Hausfeld, and Ian Krajbich. High-value decisions are fast and accurate, inconsistent with diminishing value sensitivity. *Proceedings of the National Academy of Sciences*, 119(6):e2101508119, 2022.

[27] Oscar Bartra, Joseph T McGuire, and Joseph W Kable. The valuation system: a coordinate-based meta-analysis of bold fmri experiments examining neural correlates of subjective value. *Neuroimage*, 76:412–427, 2013.

[28] John A Clithero and Antonio Rangel. Informatic parcellation of the network involved in the computation of subjective value. *Social cognitive and affective neuroscience*, 9(9):1289–1302, 2014.

[29] Stephanie M Smith and Ian Krajbich. Mental representations distinguish value-based decisions from perceptual decisions. *Psychonomic Bulletin & Review*, 28(4):1413–1422, 2021.

[30] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[31] Larry Wasserman. *All of nonparametric statistics*. Springer Science & Business Media, 2006.

[32] Ryan Webb. The (neural) dynamics of stochastic choice. *Management Science*, 65(1):230–255, 2019.

[33] Michael J Frank. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Networks*, 19(8):1120–1136, 2006.

[34] Thomas V Wiecki and Michael J Frank. A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychological Review*, 120(2):329–355, 2013.

[35] Rafael Polanía, Ian Krajbich, Marcus Grueschow, and Christian C Ruff. Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron*, 82(3):709–720, 2014.

[36] Joshua Hascher, Nitisha Desai, and Ian Krajbich. Incentivized and non-incentivized liking ratings outperform willingness-to-pay in predicting choice. *Judgment & Decision Making*, 16(6):1464–1484, 2021.

[37] Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, and Aldo Rustichini. Multinomial logit processes and preference discovery: inside and outside the black box. *Review of Economic Studies, forthcoming.*, 2022.

[38] Daniel Kahneman and Amos Tversky. On the interpretation of intuitive probability: A reply to jonathan cohen. *Cognition*, 7(4):409–411, 1979.

[39] Ralph Hertwig, Greg Barron, Elke U Weber, and Ido Erev. Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15(8):534–539, 2004.

[40] Michael J Frank, Lauren C Seeberger, and Randall C O'reilly. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*, 306(5703):1940–1943, 2004.

[41] Angelo Pirrone, Habiba Azab, Benjamin Y Hayden, Tom Stafford, and James AR Marshall. Evidence for the speed–value trade-off: Human and monkey decision making is magnitude sensitive. *Decision*, 5(2):129–142, 2018.

[42] Blair RK Shevlin and Ian Krajbich. Attention as a source of variability in decision-making: Accounting for overall-value effects with diffusion models. *Journal of Mathematical Psychology*, 105:102594, 2021.

[43] Roger Ratcliff and Michael J Frank. Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models. *Neural Computation*, 24(5):1186–1229, 2012.

[44] A Wald. Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, 16(2):117–186, 1945.

[45] Michael Woodford. Stochastic choice: An optimizing neuroeconomic model. *American Economic Review*, 104(5):495–500, 2014.

[46] Federico Echenique and Kota Saito. Response time and utility. *Journal of Economic Behavior & Organization*, 139:49–59, 2017.

[47] Drew Fudenberg, Philipp Strack, and Tomasz Strzalecki. Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108(12):3651–3684, 2018.

[48] Satohiro Tajima, Jan Drugowitsch, and Alexandre Pouget. Optimal policy for value-based decision-making. *Nature Communications*, 7(1):1–12, 2016.

[49] Carlos Alós-Ferrer, Ernst Fehr, and Nick Netzer. Time will tell: Recovering preferences when choices are noisy. *Journal of Political Economy*, 129(6):1828–1877, 2021.

[50] Rahul Bhui. Testing optimal timing in value-linked decision making. *Computational Brain & Behavior*, 2(2):85–94, 2019.

[51] Cary Frydman and Gideon Nave. Extrapolative beliefs in perceptual and economic decisions: Evidence of a common mechanism. *Management Science*, 63(7):2340–2352, 2017.

[52] Cary Frydman and Ian Krajbich. Using response times to infer others' private information: an application to information cascades. *Management Science*, 68(4):2970–2986, 2022.

[53] Corey N White and Russell A Poldrack. Decomposing bias in different types of simple decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(2):385–398, 2014.

[54] Nitisha Desai and Ian Krajbich. Decomposing preferences into predispositions and evaluations. *Journal of Experimental Psychology: General*, 151(8):1883–1903, 2022.

[55] Rafal Bogacz, Eric-Jan Wagenmakers, Birte U Forstmann, and Sander Nieuwenhuis. The neural basis of the speed–accuracy tradeoff. *Trends in Neurosciences*, 33(1):10–16, 2010.

[56] Ulrike Basten, Guido Biele, Hauke R Heekeren, and Christian J Fiebach. How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences*, 107(50):21767–21772, 2010.

[57] Todd A Hare, Wolfram Schultz, Colin F Camerer, John P O'Doherty, and Antonio Rangel. Transformation of stimulus value signals into motor commands during simple choice. *Proceedings of the National Academy of Sciences*, 108(44):18120–18125, 2011.

[58] Sebastian Gluth, Jörg Rieskamp, and Christian Büchel. Deciding when to decide: time-variant sequential sampling models explain the emergence of value-based decisions in the human brain. *Journal of Neuroscience*, 32(31):10686–10698, 2012.

[59] Brandon M Turner, Leendert Van Maanen, and Birte U Forstmann. Informing cognitive abstractions through neuroimaging: the neural drift diffusion model. *Psychological Review*, 122(2):312–336, 2015.

[60] M Andrea Pisauro, Elsa Fouragnan, Chris Retzler, and Marios G Philiastides. Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous eeg-fmri. *Nature Communications*, 8(1):1–9, 2017.

[61] Sebastian Gluth and Nachshon Meiran. Leave-one-trial-out, loto, a general approach to link single-trial parameters of cognitive models to neural data. *ELife*, 8:e42607, 2019.

[62] János Aczél. *Lectures on functional equations and their applications*. Academic Press, 1966.

[63] Eric-Jan Wagenmakers, Han LJ Van Der Maas, and Raoul PPP Grasman. An ez-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, 14(1):3–22, 2007.

[64] J McK Cattell. The time of perception as a measure of differences in intensity. *Philosophische Studien*, 19:63–68, 1902.

Acknowledgments

We thank Minhee Yoo for research assistance and James Cavanagh, Thomas Wiecki, Angad Kochar, and Michael J. Frank for sharing their data.

Funding

National Science Foundation awards 2148982 and 1554837, and Cattell Sabbatical Fund (IK); MIUR PRIN grant 2017CY2NCA (FM).

Author contributions

All authors have contributed equally to this work.

Competing interests

Authors declare that they have no competing interests.

Data and materials availability

The datasets and code for the cyclic estimator are available on the Open Science Framework at: https://osf.io/qh6ag/. All other data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.