

# Is attentional discounting in value-based decision making magnitude sensitive?

Angelo Pirrone & Fernand Gobet

Centre for Philosophy of Natural and Social Science  
London School of Economics and Political Science  
London, UK

## Abstract

Choices in value-based decision making are affected by the magnitude of the alternatives (i.e., the summed values of the options). Magnitude sensitivity has been instrumental in discriminating between computational models of choice. Smith and Krajbich (2019) have shown that the attentional drift-diffusion model (aDDM) can account for magnitude sensitivity. This is because the discount parameter on the value of the nonfixated alternative ensures faster choices for high-magnitude alternatives, even in the case of high-magnitude equal alternatives compared to low-magnitude equal alternatives. Their result highlights the importance of visual fixations as a mechanism for magnitude sensitivity. This rationale relies on the untested assumption that the discount parameter is constant across magnitude levels. However, the discount parameter could vary as a function of the magnitude of the alternatives in unpredicted ways; this would suggest that the ability of the aDDM to account for magnitude sensitivity has been misinterpreted by previous research. Here, we reanalyse previous datasets and we directly test whether attentional discounting scales with the magnitude of the alternatives. Our analyses show that attentional discounting does not vary with magnitude. This result further strengthens the aDDM and the role that visual fixations could play as an explanation of magnitude sensitivity.

*Keywords:* value-based decision making, discount, magnitude sensitivity, visual fixations

---

Corresponding authors {a.pirrone, f.gobet}@lse.ac.uk. The authors declare that there is no conflict of interest regarding the publication of this article. The datasets can be downloaded at the link made available in Smith and Krajbich (2019). The MATLAB code and JASP analyses presented in this article are available to download at <https://osf.io/kvhc4/>.

Funding from the European Research Council (ERC-ADG-835002—GEMS) is gratefully acknowledged.

## Introduction

### The attentional drift-diffusion model

Extensive research over the past decade has shown that perceptual and value-based decision making can be explained, at the computational level of analysis, by the same unifying mechanism (Krajbich, Hare, Bartling, Morishima & Fehr, 2015; Ratcliff, Smith, Brown & McKoon, 2016). In particular, for the simple case of two-alternatives forced choices, the accumulation-to-bound process (e.g., Bogacz, Brown, Moehlis, Holmes & Cohen, 2006; Krajbich et al., 2015; Ratcliff et al., 2016; Usher & McClelland, 2001) has been the basis of a number of computational models of choice that parsimoniously explain the time course of evidence accumulation and the full distribution of choices and reaction times across domains.

Of all the proposed evidence-accumulation models, the arguably most celebrated is the drift diffusion model (DDM Ratcliff, 1978; Ratcliff & McKoon, 2008), which has been adopted both in perceptual (Bogacz et al., 2006; Ratcliff et al., 2016) and value-based decision making (Krajbich, Armel & Rangel, 2010; Milosavljevic, Malmaud, Huth, Koch & Rangel, 2010). Typically, in the case of perceptual decision making (e.g., motion discrimination), participants integrate physical evidence over time, and in the case of value-based decision making (e.g., food choices), participants sequentially and stochastically extract information about the value of the alternatives. This information is then integrated over time in order to make a choice. However, note that there are also perceptual tasks that require a sequential and stochastic extraction of information, just like for value-based decisions (Tavares, Perona & Rangel, 2017).

Figure 1 shows a graphical representation of the DDM's behaviour over time. According to the DDM, participants integrate evidence over time until a decision threshold for one of the two alternatives is reached and a decision is made in favour of that alternative. In Figure 1, each threshold (i.e., each alternative) is reached twice over the course of four separate decisions; the time it takes to reach a threshold is the decision time and it is determined

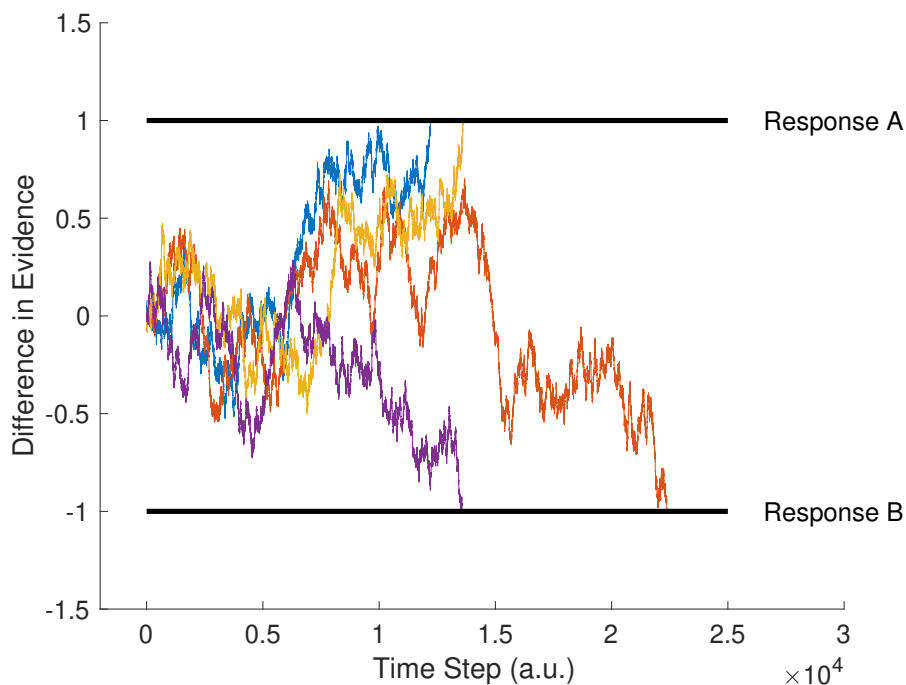


Figure 1.

Graphical representation of DDM's time course, simulated over four different runs. A decision for alternative A or B is made once a decision bound is reached. The decision is a function of the rate at which net evidence is accumulated, the distance between the two boundaries, the noise in the evidence accumulation and the starting point of evidence accumulation. Additional components are the non-decision time and the across-trial variability in parameters. In the aDDM, evidence accumulation is a function of the value of the fixated item, minus the discounted value of the nonfixated item.

by (i) a parameter known as the *drift rate*, which captures the strength at which evidence is integrated; for easier discriminations the drift rate is higher and a threshold is reached faster, on average, compared to difficult discriminations; (ii) the *boundary separation*, the distance between the two boundaries for a decision; as the boundary separation increases, more evidence accumulation is needed in order to make a choice; hence, on average, slower reaction times are predicted; and (iii) the noise in the evidence accumulation; higher noise in the evidence accumulation can contribute to faster, more inaccurate decisions.

The reaction time associated with a decision is a sum of the decision time, plus a parameter known as *non-decision time*, which captures any additional time not included in the decision time, such as the time necessary to encode the stimulus, and the time needed

in order to execute a motor response (e.g., pressing a button in order to signal the choice).

Further parameters, such as the bias towards a response, are part of the model (bias is usually fit or set to zero). In Figure 1, the evidence accumulation starts at the same distance between the two boundaries. However, it is possible that participants start to integrate evidence nearer to the boundary of one of the two alternatives; hence, they will be faster, on average, to reach the boundary for that alternative compared to the opposite boundary. Other parameters are the across-trials variabilities in drift-rate, non-decision time and bias – across-trial variabilities are either fit or more commonly set to zero (Ratcliff, Voskuilen & Teodorescu, 2018).

Seminal work by Krajbich et al. (2010) has shown that, in value-based choices, visual fixations play a key role in the evidence accumulation process. While, in the canonical DDM, the slope of evidence accumulation is a function of the difference in evidence between two alternatives, in the attentional drift-diffusion model (aDDM) proposed by Krajbich et al. (2010), the slope of evidence accumulation is determined by the difference between the value of the fixated and the discounted value of the nonfixated item.

That is, when participants look at the alternative on the left, the decision variable evolves according to

$$V_t = V_{t-1} + d(r_{left} - \theta r_{right}) + \varepsilon_t \quad (1)$$

while when participants look at the alternative to the right the decision variable evolves according to

$$V_t = V_{t-1} + d(\theta r_{left} - r_{right}) + \varepsilon_t \quad (2)$$

where  $V_t$  is the value of the decision variable at time  $t$ ,  $r_{left}$  and  $r_{right}$  are the values of the alternatives on the left and on the right,  $d$  is the speed at which evidence is integrated (i.e., an integration constant),  $\theta$  is the discount factor between 0 and 1 on the nonfixated item (if  $\theta = 0$  the nonfixated item is discounted completely while if  $\theta = 1$  the aDDM reduces

to the canonical DDM) and  $\varepsilon_t$  is white Gaussian noise with mean 0 and variance  $\sigma$  sampled at each time step.

The dynamics of the aDDM elegantly explain the tendencies of participants to choose the last-fixated alternative and the most-looked at alternative (Krajbich et al., 2010). The aDDM has been further extended to consumer choices (Krajbich, Lu, Camerer & Rangel, 2012) and to choices with multiple alternatives (Krajbich & Rangel, 2011).

### Estimating the parameters of the aDDM

Estimating the parameters of the aDDM requires computationally expensive grid search techniques for which no consensus exists regarding key aspects, such as how to use the eye-tracking data, how to bin RTs and how many simulations to run (for a discussion, see Smith, Krajbich & Webb, 2019).

Fortunately, Smith et al. (2019) recently proposed a simple method that allows one to quickly estimate the key parameter of the aDDM,  $\theta$ . Recall that  $\theta$  is the parameter that captures the weight of discount on the nonfixated item; this parameter varies in the range 0 to 1. When the value of  $\theta$  is zero, the value of the nonfixated alternative is discounted completely, while when it is one, the value of the nonfixated item is not discounted and hence the aDDM reduces to a canonical DDM in which visual fixations play no role. Previous research (Krajbich et al., 2010; Krajbich et al., 2015; Smith & Krajbich, 2019) has shown that the value of  $\theta$  is about 1/3.

Webb (2019) and Smith et al. (2019) demonstrated that the probability to choose option  $i$  is

$$P_i = Prob[(E_i X_i - E_j X_j)d + (E_j X_i - E_i X_j)\gamma > \eta_i - \eta_j] \quad (3)$$

where  $i$  and  $j$  are the two options,  $E_i$  is the fraction of time spent looking at option  $i$ ,  $X_i$  is the value of option  $i$ , and  $\eta_i$  is a random component. Performing a simple logit regression allows one to estimate  $\theta = \gamma/d$  (Smith et al., 2019), where  $\gamma$  and  $d$  are regression coefficients.

Smith et al. (2019) have shown, through Monte Carlo simulations on synthetic data and by estimating  $\theta$  from previous studies, that this method allows fast and reliable estimates virtually instantaneously compared to time-consuming grid search techniques over combinations of likely parameters.

If needed, modellers can then use the estimate of  $\theta$  from Equation 3 to estimate the remaining parameters using a grid search; this enables a considerable reduction of the combinatorial space of possible models that the grid search has to consider.

### **Magnitude sensitivity**

Traditionally, only the difference between alternatives was taken under consideration in two-alternatives decision making and magnitude was deemed as ‘irrelevant’ (Ashby, Jekel, Dickert & Glöckner, 2016; Bogacz et al., 2006). Recently, it has been proposed and demonstrated that the overall magnitude of the alternatives affects decision making too, both in perceptual and value-based decision making, for organisms at different levels of biological complexity, from unicellular organisms to monkeys and humans (Bose, Pirrone, Reina & Marshall, 2020; Dussutour, Ma & Sumpter, 2019; Hunt et al., 2012; Kvam & Pleskac, 2016; Pais et al., 2013; Pirrone, Azab, Hayden, Stafford & Marshall, 2018; Pirrone, Stafford & Marshall, 2014; Pirrone, Wen & Li, 2018; Ratcliff et al., 2018; Steverson, Chung, Zimmermann, Louie & Glimcher, 2019; Teodorescu, Moran & Usher, 2016).

Magnitude sensitivity predicts faster choices that sacrifice accuracy (i.e., the most valuable choice is chosen less often) for conditions of higher magnitude and it allows long-term reward maximisation over accuracy, since it avoids decision deadlocks for high magnitude near-equal alternatives. This criterion has been deemed evolutionary advantageous under specific decision making settings (for a discussion, see Pirrone, Azab et al., 2018; Pirrone et al., 2014; Teodorescu et al., 2016) that characterise decision making under ecological constraints, such as foraging scenarios and decision making under time pressure or in the presence of competitors. For such scenarios, sacrificing accuracy results in reducing costs of delays associated with long deliberations over high magnitude near-equal value alternat-

ives. Further explanations (Ratcliff et al., 2018; Teodorescu et al., 2016) do not align with the evolutionary arguments regarding maximisation of value; rather, magnitude sensitivity is considered a result of additional noise in the decision process that accompanies greater signal strength.

### Research questions

Smith and Krajbich (2019) have recently proposed that the aDDM can naturally account for magnitude sensitive reaction times, compared to modifications of the aDDM such as the one proposed by Cavanagh, Wiecki, Kochar and Frank (2014). The aDDM can account for magnitude sensitivity thanks to  $\theta$ , the discounting parameter. Consider the case of equal alternatives; as reported by Smith and Krajbich (2019), a choice between two alternatives having value of 5 will be made faster compared to the case in which two alternatives have value of 1. This is because the discount factor on the nonfixated item ensures that the net difference between the value of the alternatives will have, on average, a higher value for higher magnitude equal alternatives and cross a decision boundary faster. The authors tested this hypothesis using six previous datasets and showed that the aDDM can account for magnitude sensitivity; a result that cannot be accounted by other fixation-dependent models of decision making (Cavanagh et al., 2014).

However, the rationale proposed by Smith and Krajbich (2019) relies on the untested assumption that the  $\theta$  parameter, and its behavioural counterpart – the fixation time advantage, are constant across conditions of different magnitude. To the best of our knowledge, this assumption has not been tested previously.

The focus of our investigation is exploratory, as opposed to confirmatory: by reanalysing datasets from previous studies, we investigate whether visual discounting ( $\theta$  in the aDDM) varies as a function of magnitude.

The aDDM predicts a support for the null hypothesis regarding a varying effect of fixation time advantage (the most looked at alternative is more likely to be chosen) across magnitude levels. That is, the aDDM predicts that the effect of fixation time advant-

age is constant (that is, does not differ) at different magnitude levels. For this reason, it is important to evaluate whether the data provide support for the null hypothesis. Frequentist/orthodox analyses do not allow one to quantify the support in favour of the null hypothesis, and for this reason we decided to adopt Bayesian analyses using JASP (JASP Team, 2020).

Across four datasets, our results show that  $\theta$  and the effect of fixation bias advantage remain constant as the magnitude of the alternatives increases. That is, the value of the discount on the nonfixated item (and the associated parameter  $\theta$ ) does not differ between high and low magnitude conditions.

## Methods

### Data

We used data from four binary-choice food experiments; all datasets came from Ian Krajbich's lab and were made publicly available on OSF (for links, see Smith & Krajbich, 2019).

These datasets were analysed in Smith and Krajbich (2019) in order to show the ability of the aDDM to account for magnitude sensitivity. The four experiments from which the datasets were collected are variations of the classical experimental paradigm in value-based decision making: participants rate images of food and, subsequently, they are presented with binary choices between pairs of food images previously rated. At the end of the experiment, participants are rewarded with one of the items (selected at random) that they had chosen and they also receive a payment for their participation. During the binary-choice experiment, participants' eye movements are recorded.

Note that Smith and Krajbich (2019) also included two further datasets in which participants made choices between learned, abstract stimuli that were rewarded probabilistically. It is debatable whether these choices are *value-based* (for a discussion, see Pirrone & Gobet, 2020); furthermore, one dataset (Dataset 5) did not include a condition with equal alternatives, which is condition of interest for some of our analyses. For these reasons, we



did not include these two studies in our reanalyses.

**Data Set 1.** These data come from Data Set 1 reported in Smith and Krajbich (2019), and are taken from the influential study presented in Krajbich et al. (2010): 39 participants made 100 choices between pairs of images of food items previously rated. Only items that were rated between 0 and 10 were included in the binary-choice experiment; the maximum allowed difference between the items was 5.

**Data Set 2.** These data come from Data Set 2 reported in Smith and Krajbich (2018): 44 participants made 200 choices between pairs of images of food items previously rated. Only items that were rated between 1 and 10 were included in the binary-choice experiment; the maximum allowed difference between the items was 5.

**Data Set 3.** These data come from Data Set 3 reported in Smith and Krajbich (2019), taken from Chen and Krajbich (2016): 44 participants made 200 choices between pairs of images of food items previously rated. Only items that were rated between 0 and 10 were included in the binary-choice experiment; the maximum allowed difference between the items was 3.

**Data Set 4.** These data come from Data Set 4 reported in Smith and Krajbich (2019) and subsequently reported in Gwinn and Krajbich (2020): 36 participants made 200 choices between pairs of images of food items previously rated between 1 and 10. This study was included because it allowed us to test the effect of magnitude on  $\theta$  while controlling for value difference, given that the maximum allowed difference between the items was 1.

## Analyses and Results

In order to test whether attentional discounting varies as a function of magnitude, we performed both behavioural analyses and analyses based on the aDDM in which we estimated  $\theta$  for low and high magnitude levels.

### Behavioural analyses

If there is a relationship between attention and magnitude on choice (e.g., no or lower relationship between attention and choice for high overall-value choices), then the effect of

Table 1

*Classification scheme of Bayes Factor. Taken from Lee and Wagenmakers (2014), adjusted from Jeffreys (1998).*

Bayes Factor	Interpretation
>100	Extreme evidence for $H_0$
30 – 100	Very strong evidence for $H_0$
10 – 30	Strong evidence for $H_0$
3 – 10	Moderate evidence for $H_0$
1 – 3	Anecdotal evidence for $H_0$
1	No evidence
1/3 – 1	Anecdotal evidence for $H_1$
1/3 – 1/10	Moderate evidence for $H_1$
1/10 – 1/30	Strong evidence for $H_1$
1/30 – 1/100	Very strong evidence for $H_1$
<1/100	Extreme evidence for $H_1$

fixation time advantage is expected to vary at different magnitude levels, even for equal alternatives – that is, alternatives having a null difference.

We decided to use equal alternatives only since those conditions allow to estimate the effects of fixation advantage and magnitude on choice, in the absence of additional decisional confounds due to value difference. The focus on equal alternatives has characterised previous studies about magnitude sensitivity (Bose et al., 2020; Pirrone, Azab et al., 2018; Pirrone & Gobet, 2020; Pirrone et al., 2014; Pirrone, Wen & Li, 2018).

In the following, we report the Bayes Factor (BF) for exclusion which quantifies the amount of evidence for the null hypothesis relative to the alternative hypothesis (Lee & Wagenmakers, 2014). For BF, we used the classification scheme adopted by JASP that is adjusted from Jeffreys (1998), as reported in Table 1.

The number of trials per magnitude level did not allow an estimation for each individual and magnitude level, hence we opted for a median split analysis that allowed us to categorise each trial as low or high magnitude. We also calculated the fixation bias advantage, defined as whether participants during a trial fixated longer to the left or the right alternative.

Magnitude (low vs high) and fixation advantage (left vs right) were entered in a re-

peated measures Bayesian ANOVA in which study (i.e., dataset 1, 2, 3 or 4) was entered as a between subject factor. Bayesian model averaging (Hoeting, Madigan, Raftery & Volinsky, 1999) is a method that, averaging across all combinations of models that could have generated the data, provides the likelihood for exclusion (or inclusion) of a specific factor as a candidate for data explanation. This method allows concise information regarding the factors in cases in which there is a high number of combinations of models that could have generated the data.

Table 2 shows that fixation advantage is the only factor having a decisive role, since the BF is so low that it approaches 0 (i.e., the BF for inclusion approaches infinite). Interestingly, the probability to choose left was not affected by magnitude and the effect of fixation advantage did not vary across magnitude levels; as reported in Table 2, the BF for the interaction of magnitude and fixation advantage provides strong evidence for the support of the null hypothesis.

Table 2

*Analysis of effects for a repeated measures Bayesian ANOVA on choice (probability to choose left) with median split magnitude (low vs high), fixation advantage (left vs right) and study as between subject factor.  $P(\text{incl})$  is the sum of the prior probabilities of all models that include the effect;  $P(\text{incl}|\text{data})$  is the sum of the posterior probabilities of all models that include the effect;  $BF_{\text{excl}}$  is the Bayes Factor for the exclusion of an effect. Priors were set to default values in JASP, to meet formal desiderata.*

Effects	P(incl)	P(incl data)	$BF_{\text{excl}}$
magnitude	0.737	0.131	18.554
advantage	0.737	1	0
study	0.737	0.019	140.989
magnitude $\times$ advantage	0.316	0.017	26.533
magnitude $\times$ study	0.316	$7.063 \cdot 10^{-5}$	6534.228
advantage $\times$ study	0.316	$6.463 \cdot 10^{-4}$	713.682
magnitude $\times$ advantage $\times$ study	0.053	$4.039 \cdot 10^{-8}$	$1.375 \cdot 10^6$

Figure 2 shows a graphical summary of results. In line with results from the Bayesian ANOVA, there is strong evidence for an effect of fixation advantage: when participants look at the left alternative for longer, they are more likely to choose that alternative; conversely, when they fixate the alternative on the right longer, they are more likely to choose right.

The effect of fixation advantage is constant across magnitude levels and can be seen not to differ between low and high magnitude conditions.

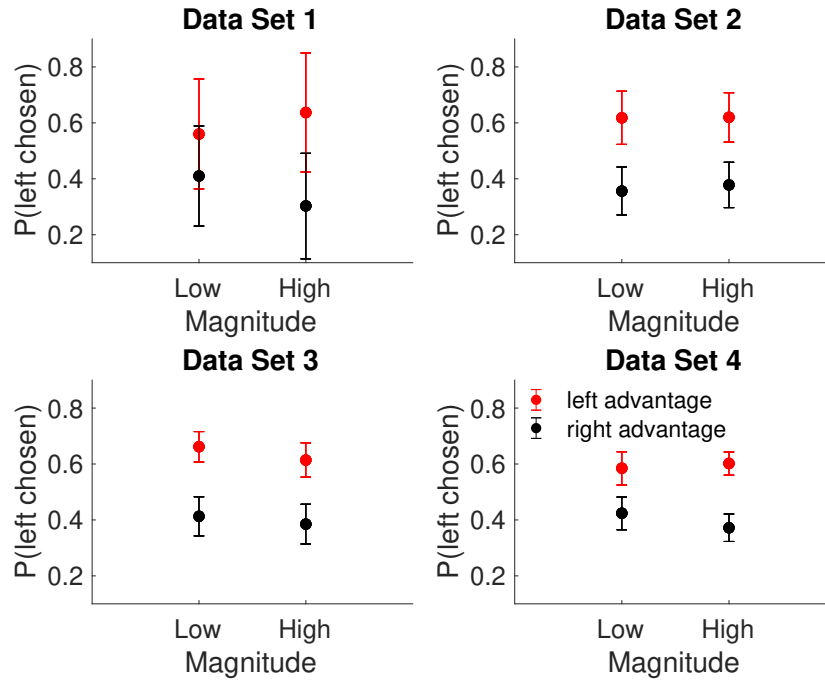


Figure 2.

Probability to choose left as a function of magnitude of the alternatives (low vs high) and fixation advantage (left or right). Bars are 95% Bayesian credible intervals.

### aDDM analyses: Estimating $\theta$

Subsequently, we estimated  $\theta$  using Equation 3 for each high vs low magnitude level (median split) for each participant for each study. No datapoint was excluded from these analyses. We estimated  $d$  and  $\gamma$  separately for trials in which trial magnitude  $\mu_i$  was lower than the median magnitude of the dataset  $M$ , and for trials in which trial magnitude was higher or equal to the median magnitude of the dataset.

That is, we estimated  $d$  and  $\gamma$  from the following logistic regression in which choice  $P_{left}$  is expressed as probability of choosing left:

$$P_{left} \sim 1 + d|(\mu_i < M) + d|(\mu_i \geq M) + \gamma|(\mu_i < M) + \gamma|(\mu_i \geq M)$$

Note that estimating  $d$  and  $\gamma$  similarly to what done in Smith and Krajbich (2019):

$$P_{left} \sim 1 + d|(v_{left} < median(v_{left})) + d|(v_{left} \geq median(v_{left})) + \gamma|(v_{left} < median(v_{left})) + \gamma|(v_{left} \geq median(v_{left}))$$

yielded practically identical results.

We estimated  $\theta$  for low magnitude conditions as  $\theta_{low} = \gamma|(\mu_i < M)/d|(\mu_i < M)$  and for high magnitude conditions as  $\theta_{high} = \gamma|(\mu_i \geq M)/d|(\mu_i \geq M)$ .

Comparing the two separate estimates of  $\theta$  allowed us to test whether this parameter varied across low and high magnitude levels, Table 3.

Since the BF for the effect of magnitude provides strong support for the null hypothesis, our analyses show that  $\theta$  does not vary between high and low magnitude levels in any of the four studies. This outcome is not surprising given the result reported above of no difference in fixation advantage between high and low magnitude conditions.

Table 3

*Analysis of effects for the difference between high and low magnitude  $\theta$  across the four studies.*

Effects	P(incl)	P(incl data)	BF <sub>excl</sub>
magnitude	0.600	0.125	10.539
study	0.600	0.349	2.802
magnitude $\times$ study	0.200	0.003	76.851

**A cautionary tale: Individual differences matter.** In this section, we report the result from a first analysis that we had performed by using the aggregated trials for each study and collapsing those across magnitude. Although this analysis itself is conceptually/statistically inappropriate, we report it as it highlights the importance of accounting for individual differences when estimating model parameters and the risks of drawing conclusions when data are aggregated across participants.

We estimated  $\theta$  using Equation 3 for each magnitude level (summed value of the two items) of the aggregated datasets of all the participants for each study, without allowing for individual differences.

The estimated  $\theta$  for each of the three data sets, as a function of magnitude of the alternatives, is reported in Figure 3. This Figure seems to suggest that the discount para-

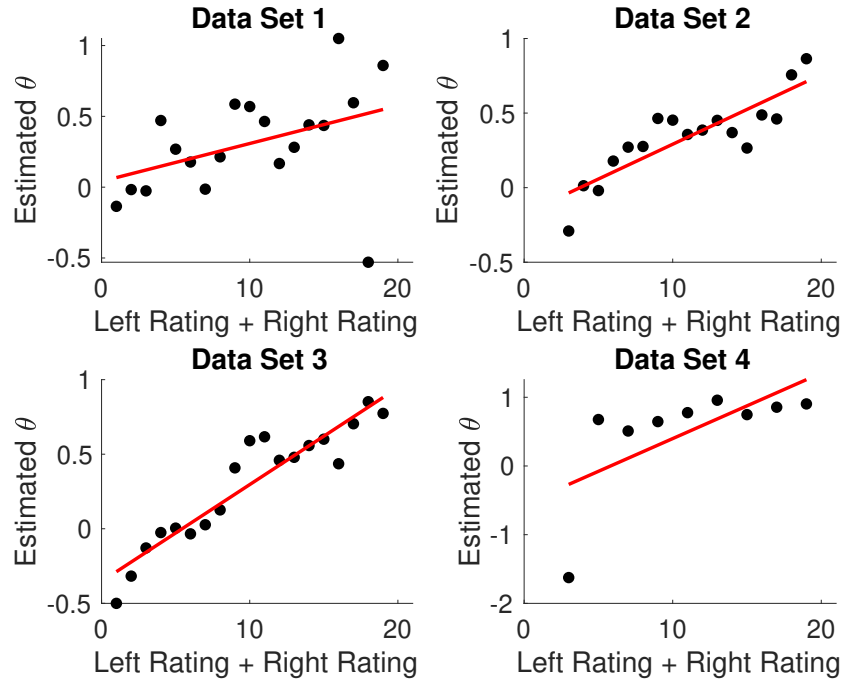


Figure 3.

Results of the analyses using the aggregated data. For each of the four data sets, we report in black the value of the estimated  $\theta$  as a function of overall magnitude of the alternatives (i.e., left rating + right rating). The red line is a linear regression fitted to the data. In all datasets,  $\theta$  increases (i.e., the discount decreases) as a function of overall magnitude.

meter scales with the magnitude of the alternatives. That is, high magnitude conditions are less discounted compared to low magnitude conditions; hence, the amplifying effect of gaze on value is counterbalanced by lower discount levels for high magnitude conditions (note: as theta increases, the discount decreases). It is interesting to note that at first we had taken these results as strong evidence that attentional discounting does decrease with magnitude in value-based decision making. However, in light of the analyses reported in the previous sections (in which we accounted for participant differences), this result highlights the importance of taking individual differences into account in order to generate valid conclusions from data, especially in a field such as value-based decision making in which strong inter-individual differences are well known and documented.

Preliminary analyses showed that aDDM simulations with static  $\theta$  across magnitude levels but varied across individuals result in correlations between  $\theta$  and magnitude when

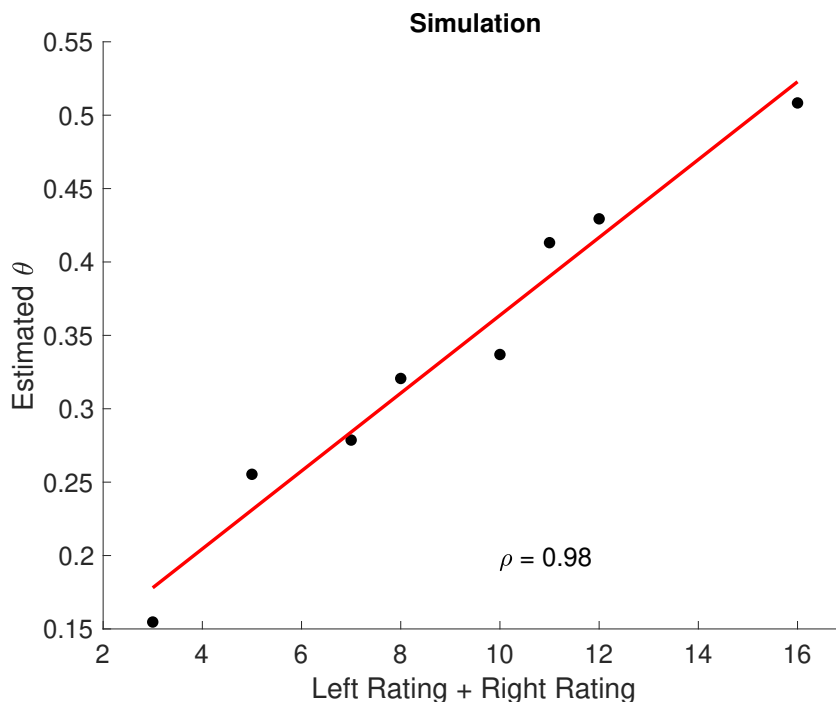


Figure 4.

Simulations of the aDDM using the aDDM Toolbox (<https://github.com/goptavares/aDDM-Toolbox>). 19 artificial participants were generated, each performing the same number and type of trials (i.e., value on the left vs right) and with constant (default) aDDM parameters across conditions – see data on OSF for details.  $\theta$  is the only parameter that varied across participants (from .05 to .95, in steps of .05 for a total of 19 possible combinations), but as all other parameters, it was kept constant for each participant across magnitude levels.  $\theta$  was estimated separately for each magnitude level using Equation 3 by combining all trials from all participants for each magnitude level; the result is an estimate of  $\theta$  that clearly increases as a function of the magnitude of the alternatives.

subject-level random effects are not accounted<sup>1</sup>, Figure 4. While it is out of the scope of the current work to investigate in detail the source of discrepancies between analyses with and without individual differences taken into account, this will be the focus of future ad-hoc work that will directly address differences in model parameter’s estimates with and without accounting for individual differences.

<sup>1</sup>We thank an anonymous reviewer of a significantly different previous version of this manuscript for firstly pointing this out.

### Discussion

In this paper, we have shown that attentional discounting, captured by the parameter  $\theta$  in the aDDM, is not a function of the magnitude of the alternatives. In previous studies, this parameter was kept constant across conditions; although this was not explicitly justified, we suspect that it was assumed that participants set  $\theta$  before the stimulus appears on screen; hence, it could not be adjusted during the course of a trial or at the beginning of a trial. This is the same rationale as that proposed for fixing boundary separation across conditions in cognitive modelling using DDMs (Bogacz et al., 2006; Ratcliff & McKoon, 2008; Ratcliff et al., 2016). However, this assumption cannot be made a priori and needed to be tested – which is what we have done in the current study. By way of example, a reasonable prediction could have been that high magnitude conditions are less discounted compared to low magnitude conditions; therefore, the amplifying effect of gaze on value could be counterbalanced by lower discount levels for high magnitude conditions. This would suggest that the ability of the aDDM to account for magnitude sensitivity has been overestimated by previous research. Finding that the discount of the nonfixated item is a function of the magnitude of the alternatives would also open up a number of timely research questions and the possibility of reconsidering conclusions of previous studies. Thus, we believe that it was important to directly test this hypothesis.

A number of magnitude sensitive accounts of decision making have been proposed by previous research; Teodorescu et al. (2016) have shown that both magnitude dependent processing noise (i.e., noise that scales with the magnitude of the alternatives, rather than being constant) and lateral inhibition between evidence accumulators (Usher & McClelland, 2001) provide a principled and quantitatively accurate account of magnitude sensitive reaction times. Note that the hypothesis of magnitude dependent noise received further experimental support in Brunton, Botvinick and Brody (2013), Louie, Khaw and Glimcher (2013), Lu and Doshier (2008). Similarly, other models in which magnitude sensitivity arises as a function of the accumulation dynamics or as a function of collapsing thresholds have been proposed (Bose et al., 2020; Pais et al., 2013). Interestingly, these models are agnostic



about visual fixations in their original formulations, and for these models magnitude sensitivity arises as a function of information processing and/or accumulation dynamics. Future research should integrate visual fixation effects within those magnitude sensitive accounts. While those models do not make any prediction regarding visual fixations, visual fixations play a key role in both perceptual (Tavares et al., 2017) and value-based decision making (Krajbich et al., 2010; Smith & Krajbich, 2019) that must be accounted for by any model of choice, unless the unreasonable assumption of stimuli being equally fixated at the same time is made.

It has been widely argued elsewhere (Bose et al., 2020; Pirrone, Azab et al., 2018; Ratcliff et al., 2018; Teodorescu et al., 2016) that magnitude sensitivity is a key factor that allows to overcome model mimicry between computational models of choice; only a limited number of models and mechanisms allow magnitude sensitivity. Here, we further strengthen the evidence that attentional discounting provides a principled account of magnitude sensitivity in value-based decision making. Future studies should directly compare this mechanism with competing ones, such as magnitude-dependent noise in the evidence of accumulation.

A number of models can account for magnitude sensitivity (e.g., Pais et al., 2013; Ratcliff et al., 2018; Teodorescu et al., 2016) but, to the best of our knowledge, only aDDMs make a fixation-dependent account of magnitude sensitivity. While, as we suggest, future studies should directly compare this mechanism with competing ones, such as magnitude-dependent noise in the evidence of accumulation, the focus of this article has been to test the assumption that fixation-dependent accounts can account for magnitude sensitivity because the effect of fixation is constant at different magnitude levels. This result is important in particular for aDDMs (as they are the key models for the effect of fixation on decision making), but it also has repercussions for any model of decision making that makes prediction about magnitude sensitivity on the basis of visual fixations.

## References

- Ashby, N. J., Jekel, M., Dickert, S. & Glöckner, A. (2016). Finding the right fit: A comparison of process assumptions underlying popular drift-diffusion models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(12), 1982.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765.
- Bose, T., Pirrone, A., Reina, A. & Marshall, J. A. (2020). Comparison of magnitude-sensitive sequential sampling models in a simulation-based study. *Journal of Mathematical Psychology*, *94*, 102298.
- Brunton, B. W., Botvinick, M. M. & Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science*, *340*(6128), 95–98.
- Cavanagh, J. F., Wiecki, T. V., Kochar, A. & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, *143*(4), 1476.
- Chen, W. J. & Krajbich, I. (2016). Pupil dilation and attention in value-based choice. *Unpublished Manuscript, The Ohio State University*.
- Dussutour, A., Ma, Q. & Sumpter, D. (2019). Phenotypic variability predicts decision accuracy in unicellular organisms. *Proceedings of the Royal Society B*, *286*(1896), 20182825.
- Gwinn, R. & Krajbich, I. (2020). Attitudes and attention. *Journal of Experimental Social Psychology*, *86*, 103892.
- Hoeting, J. A., Madigan, D., Raftery, A. E. & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial. *Statistical Science*, 382–401.
- Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. & Behrens, T. E. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*, *15*(3), 470–476.
- JASP Team. (2020). Jasp (version 0.12.2) [computer software], <https://jasp-stats.org/>.
- Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.

- Krajbich, I., Armel, C. & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298.
- Krajbich, I., Hare, T., Bartling, B., Morishima, Y. & Fehr, E. (2015). A common mechanism underlying food choice and social decisions. *PLoS Computational Biology*, *11*(10), e1004371.
- Krajbich, I., Lu, D., Camerer, C. & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, *3*, 193.
- 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*(33), 13852–13857.
- Kvam, P. D. & Pleskac, T. J. (2016). Strength and weight: The determinants of choice and confidence. *Cognition*, *152*, 170–180.
- Lee, M. D. & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Louie, K., Khaw, M. W. & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences*, *110*(15), 6139–6144.
- Lu, Z.-L. & Doshier, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review*, *115*(1), 44.
- Milosavljevic, M., Malmaud, J., Huth, A., Koch, C. & Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making*, *5*(6), 437.
- Pais, D., Hogan, P. M., Schlegel, T., Franks, N. R., Leonard, N. E. & Marshall, J. A. (2013). A mechanism for value-sensitive decision-making. *PloS One*, *8*(9).
- Pirrone, A., Azab, H., Hayden, B. Y., Stafford, T. & Marshall, J. A. (2018). Evidence for the speed–value trade-off: Human and monkey decision making is magnitude sensitive. *Decision*, *5*(2), 129.

- Pirrone, A. & Gobet, F. (2020). Value-based decision making in naturalistic vs laboratory settings. *Under Review*.
- Pirrone, A., Stafford, T. & Marshall, J. A. (2014). When natural selection should optimize speed-accuracy trade-offs. *Frontiers in Neuroscience*, 8, 73.
- Pirrone, A., Wen, W. & Li, S. (2018). Single-trial dynamics explain magnitude sensitive decision making. *BMC Neuroscience*, 19(1), 1–10.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59.
- Ratcliff, R. & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922.
- Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281.
- Ratcliff, R., Voskuilen, C. & Teodorescu, A. (2018). Modeling 2-alternative forced-choice tasks: Accounting for both magnitude and difference effects. *Cognitive Psychology*, 103, 1–22.
- Smith, S. M. & Krajbich, I. (2018). Attention and choice across domains. *Journal of Experimental Psychology: General*, 147(12), 1810.
- Smith, S. M. & Krajbich, I. (2019). Gaze amplifies value in decision making. *Psychological Science*, 30(1), 116–128.
- Smith, S. M., Krajbich, I. & Webb, R. (2019). Estimating the dynamic role of attention via random utility. *Journal of the Economic Science Association*, 5(1), 97–111.
- Steverson, K., Chung, H.-K., Zimmermann, J., Louie, K. & Glimcher, P. (2019). Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. *Scientific Reports*, 9(1), 1–14.
- Tavares, G., Perona, P. & Rangel, A. (2017). The attentional drift diffusion model of simple perceptual decision-making. *Frontiers in Neuroscience*, 11, 468.
- Teodorescu, A. R., Moran, R. & Usher, M. (2016). Absolutely relative or relatively absolute: Violations of value invariance in human decision making. *Psychonomic Bulletin & Review*, 23(1), 22–38.

- Usher, M. & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550.
- Webb, R. (2019). The (neural) dynamics of stochastic choice. *Management Science*, *65*(1), 230–255.