# Magnitude-sensitive reaction times reveal non-linear time costs in multi-alternative decision-making

**James A. R. Marshall**[1*], **Andreagiovanni Reina**[1,2], **Angelo Pirrone**[3†‡]

***For correspondence:** james.marshall@sheffield.ac.uk (JARM)

[1]Department of Computer Science, University of Sheffield, UK; [2]IRIDIA, Université Libre de Bruxelles, Belgium; [3]Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, UK

## Abstract

Optimality analysis of value-based decisions in binary and multi-alternative choice settings predicts that reaction times should be sensitive only to differences in stimulus magnitudes, but not to overall absolute stimulus magnitude. Yet experimental work in the binary case has shown magnitude sensitive reaction times, and theory shows that this can be explained by switching from linear to geometric time costs, but also by nonlinear subjective utility. Thus disentangling explanations for observed magnitude sensitive reaction times is difficult. Here for the first time we extend the theoretical analysis of geometric time-discounting to ternary choices, and present novel experimental evidence for magnitude-sensitivity in such decisions. We consider the optimal policies for all possible combinations of linear and geometric time costs, and linear and nonlinear utility; interestingly, geometric discounting emerges as the predominant explanation for magnitude sensitivity.

## Introduction

While the normative, optimal policy, approach to understanding decision-making is now well established for perceptual decisions (*e.g.* ***Bogacz et al.*** (***2006***)), it has only recently been applied to value-based decisions (***Fudenberg et al., 2018***; ***Tajima et al., 2016***, ***2019***); such decisions differ from perceptual decisions because decision makers are rewarded by the value of the selected option, rather than whether or not they selected the best option (*e.g.* ***Pirrone et al.*** (***2014***); ***Tajima et al.*** (***2016***, 2019); ***Krajbich et al.*** (***2010***)). Recently researchers have analysed multi-alternative value-based decision-making (***Tajima et al., 2019***), building on earlier work in optimal decision policies for binary value-based choices (***Tajima et al., 2016***). Through sophisticated analysis based on the standard tool for solving such decision problems, stochastic dynamic programming (***Mangel and***

***Clark, 1988***; ***Houston and McNamara, 1999***), the authors also present neurally-plausible decision mechanisms that may implement or approximate the optimal decision policies (***Tajima et al., 2016***, ***2019***). These policies turn out to be described by rather simple and well-known decision mechanisms, such as drift-diffusion models with decision thresholds that collapse over time for the binary case (***Tajima et al., 2016***), and nonlinear time-varying thresholds that interpolate between best-vs-average and best-vs-next in the multi-alternative case (***Tajima et al., 2019***).

Interestingly, the theoretically optimal policy for the binary decision case (***Tajima et al., 2016***) is inconsistent with empirical observations of magnitude-sensitive reaction-times (***Teodorescu et al.*** (***2016***); ***Pirrone et al.*** (***2018a***); ***Steverson et al.*** (***2019***); ***Zajkowski et al.*** (***2019***); ***Turner et al.*** (***2019***), but see ***Bhui*** (***2019***)), unless assumptions are made that subjective utilities for decision-makers are nonlinear, or decisions are embedded in a fixed-length time period with known or learnable distributions of trial option values, so that a variable opportunity cost arises from decision time (***Tajima et al., 2016***). Furthermore, even single-trial dynamics lead to magnitude sensitive reaction times (***Pirrone et al., 2018b***).

Previous analyses made an assumption that appears widespread in psychology and neuroscience, that decision-makers should optimise their Bayes Risk from such decisions (***Tajima et al., 2016***, ***2019***); this is equivalent to maximising the expected value of decisions in which there is a linear cost for the time spent deciding (***Bogacz et al., 2006***; ***Pirrone et al., 2014***). For a lab subject in a pre-defined and known experimental design this may appear appropriate, for example because there may be a fixed time duration within which a number of decision trials will occur and the subject can learn the value distribution of the trials (e.g. ***Bogacz et al.*** (***2006***); ***Pirrone et al.*** (***2014***)). However, an alternative and standard formulation of the Bellman equation, the central equation in constructing a dynamic program, accounts for the cost of time by discounting future rewards geometrically, so a reward one time step in the future is discounted by rate $\gamma < 1$, two time steps in the future by $\gamma^2$, and so on (see Materials and Methods). This is a standard assumption in behavioural ecology (***Mangel and Clark, 1988***; ***Houston and McNamara, 1999***), in which discounting the future means that future rewards are not guaranteed but are uncertain, due to factors such as interruption, consumption of a food item by a competitor, mortality, and so on. Thus discounting the future represents the inherent uncertainty that animals must make decisions under in their natural environments, in which their brains evolved. The appropriate discount is then the probability that future rewards are realised, hence geometric discounting is optimal since probabilities multiply. Indeed there is extensive evidence of such reward discounting in humans and other animals (*e.g.* ***Sellitto et al.*** (***2010***), although this frequently suggests hyperbolic rather than geometric discounting, a fact that in itself merits an explanation based on optimality theory (***McNamara and Houston, 2009***)).

Rederiving optimal policies to account for geometric (***Marshall, 2019***) or general multiplicative (***Steverson et al., 2019***) costs of time qualitatively changes them in the binary decision case, in-

troducing magnitude-sensitive reaction times (***Marshall, 2019***; ***Steverson et al., 2019***). However, disentangling these from nonlinear subjective utility is challenging, and cannot be excluded as an explanation for previous results (***Teodorescu et al., 2016***; ***Pirrone et al., 2018a***; ***Steverson et al., 2019***; ***Zajkowski et al., 2019***; ***Turner et al., 2019***; ***Bhui, 2019***; ***Smith and Krajbich, 2019***; ***Pirrone and Gobet, 2021***).

Here for the first time we extend the theoretical and experimental study of magnitude-sensitivity to three-alternative decisions. We first present evidence for magnitude-sensitive reaction times in three-way equal-alternative decisions. We then present optimal policy analyses and novel numerical simulations for such ternary decisions. Importantly, for a wide variety of utility functions, strong magnitude-sensitivity is only observed when there is a multiplicative cost for time, rather than the previously assumed linear time cost. Thus magnitude-sensitivity is revealed as genuinely diagnostic for multiplicative time costing, as all other assumptions either do not generate this phenomenon, or can be discounted.

## Results

### Multi-Alternative Decisions in Human Psychophysical Trials are Magnitude-Sensitive

Here we provide strong empirical evidence for magnitude sensitivity with multiple alternatives, using an experimental paradigm similar to the one used to show magnitude sensitivity for two-alternative decision making (***Teodorescu et al., 2016***; ***Pirrone et al., 2018a***). Details of the experiment (methods, participants, etc.) are reported in Materials and Methods. Participants had to choose which of three grey patches was brighter in an online experiment. Although the experiment included conditions for which a brighter alternative existed, conditions of interest were equal alternatives of different magnitude, that is, conditions for which the three patches had the same brightness that could vary across magnitude conditions. Equal alternatives allow us to test hypotheses regarding magnitude sensitivity, by keeping differences in evidence constant (***Pirrone et al., 2018a***,b; ***Pirrone and Gobet, 2021***).

As previously done for binary decisions (***Pirrone et al., 2018a***,b), here we focused our analyses exclusively on equal alternatives. For the analyses, we did not censor any datapoints.

As shown in Figure 1, the data show strong magnitude sensitivity, given that choices for equal alternatives of higher magnitude conditions (higher brightness on a scale from 0 to 1 on Python) were made faster. Results from a Bayesian repeated-measures ANOVA using JASP (***JASP Team, 2020***) shows decisive evidence for the model including magnitude as an independent variable, since this is $9.138 \times 10^{23}$ times more likely to have generated the data compared to the null model, which includes no predictor variables at all. As the COVID-19 pandemic necessitated an online experiment we could not collect or control information on a number of possible confounds (viewer position, motivation, room luminosity, etc.), and there are multiple sources of unaccounted variability in our online experiment; however there is no *a priori* reason to expect these to act as consistent

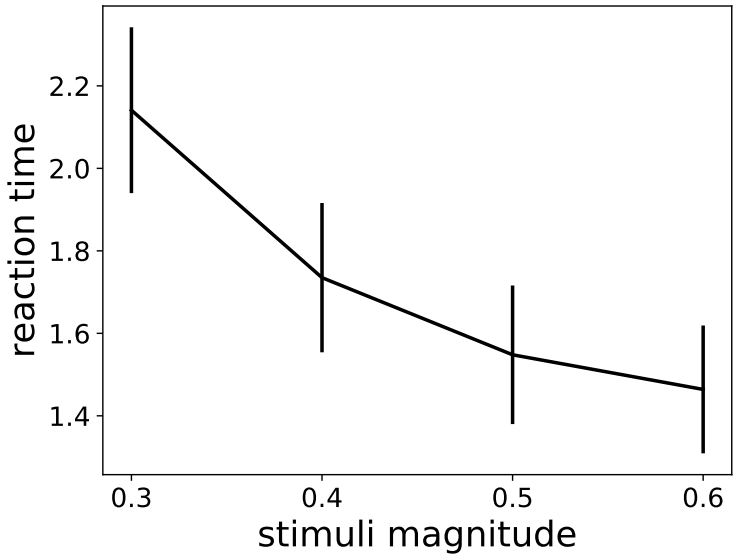


**Figure 1.** Empirical results from the behavioural online experiment. Decreasing reaction times as a function of the magnitude of the equal alternatives. X-axis presents mean brightness of equal alternatives (0.3, 0.4, 0.5, 0.6), on a scale of brightness from 0 to 1 in PsychoPy. Y-axis presents mean reaction times, in seconds. Bars show 95% credible intervals. All $N = 117$ participants experienced each equal alternative presentation ten times, interleaved with unequal alternative trials in pseudo-randomised order. A Bayesian ANOVA shows that a model with magnitude as an independent variable best explains the data, being $9.138{\cdot}10^{23}$ more likely to have generated the data compared to the null model.

confounds in the magnitude-sensitive reaction times observed. Furthermore, the very large sample size for our study ($N = 117$; compared to $N = 8$ and $N = 9$ for previous similar investigations (***Teodorescu et al., 2016***; ***Pirrone et al., 2018a***)) should minimise effects due to randomly-distributed confounds.

## Optimal Policies

For our theoretical analysis we begin by re-deriving optimal policies for decisions when the change is made from linear costing of time, or Bayes Risk, to geometric discounting of future reward. Note that geometric discounting of future rewards is similar to, but not the same as, non-linear utility. As remarked in the introduction above, for binary decisions magnitude-sensitive reaction times can be explained by optimal decision policies for either multiplicative (*e.g.* geometric) time discounting (***Marshall, 2019***; ***Steverson et al., 2019***) or nonlinear subjective utility with linear time costs (***Tajima et al., 2016***). In the multi-alternative case, on the other hand, the picture is more nuanced; moving from linear costing of time to geometric discounting of future rewards changes complicated time-dependent non-linear decision thresholds ((***Tajima et al., 2019***) Fig. 7) into either simple linear ones that collapse over time for lower-value option sets (Fig. 2), or nonlinear boundaries that evolve over time similarly to the Bayes Risk-optimising case for higher-value option sets (***Marshall*** (***2019***); Fig. 2). As Tajima *et al.* note, the simpler linear decision boundaries implement the 'best-vs-average' decision strategy, whereas the more complex boundaries interpolate between 'best-vs-average' and 'best-vs-next' decision strategies (***Tajima et al., 2019***); interestingly simply moving to nonlinear

subjective utility with linear time costs simplifies the decision strategy to the 'best-vs-next' strategy (Fig. 2; see ***Tajima et al.*** (***2019***), Fig. 6C).

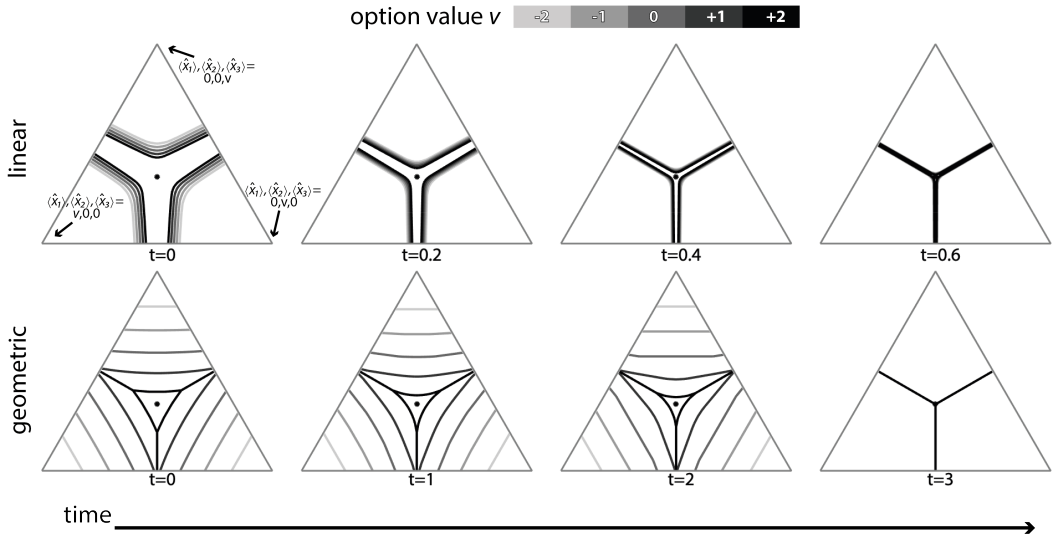


**Figure 2.** Linear time costs lead to weakly magnitude-sensitive optimal policies (top row), while geometric discounting of reward leads to strongly magnitude-sensitive optimal policies (bottom row). In the linear time cost (Bayes Risk) case nonlinear subjective utility changes complex time and value-dependent decision boundaries in estimate space into a simple mostly magnitude-insensitive 'best-vs-next' strategy (top row; see ***Tajima et al.*** (***2019***), Fig. 6C). For geometric discounting of rewards over time, optimal decision boundaries are strongly magnitude-sensitive and interpolate between simple 'best-vs-average' and 'best-vs-next' strategies (see ***Tajima et al.*** (***2019***), Fig. 6). Triangles are low dimensional projections of the 3-dimensional evidence estimate space onto a plane moving along the equal value line, at value $v$ (***Tajima et al., 2019***). Dynamic programming parameters were: prior mean $\bar{x}_{p,i} = 1.5$ and variance $\sigma^2_{p,i} = 5$, waiting time $t_w = 1$, temporal costs $c = 0$, $\gamma = 0.2$, and utility function parameters $m = 4$, $s = 0.25$ (for the linear time cost) and $m = 4$, $s = 3.5$ (for the geometric time cost).

## Multi-Alternative Decisions: Optimal Policies are Weakly Magnitude-Sensitive for Nonlinear Subjective Utility Under Bayes Risk-Optimisation

Under Bayes Risk-optimisation it is known that, for binary decisions, optimal policies are magnitude-insensitive when subjective utility is linear, whereas they are magnitude-sensitive when subjective utility is nonlinear (***Tajima et al., 2016***, ***2019***).

For ternary decisions, however, even with nonlinear subjective utility, policies exhibit very weak magnitude-sensitivity early in decisions, becoming magnitude-insensitive as decisions progress (Fig. 2, row 'linear'). Sensitivity analysis shows that magnitude-insensitivity is a general pattern. An informal understanding of this can be arrived at by appreciating that sigmoidal functions have two extremes of parameterisation; in one extreme they are almost linear (low $s$ in Eq. ()), hence will be mostly magnitude insensitive due to the known result (***Tajima et al., 2016***). At the other extreme, the function becomes step-like (high $s$ in Eq. ()); in this case options are either good or bad, and the optimal policy rapidly becomes 'choose the best' (Fig. 3), since under such a scenario sampling is of minimal benefit as early information quickly indicates whether an option is good or bad, and choosing the first option that appears to be good is optimal.

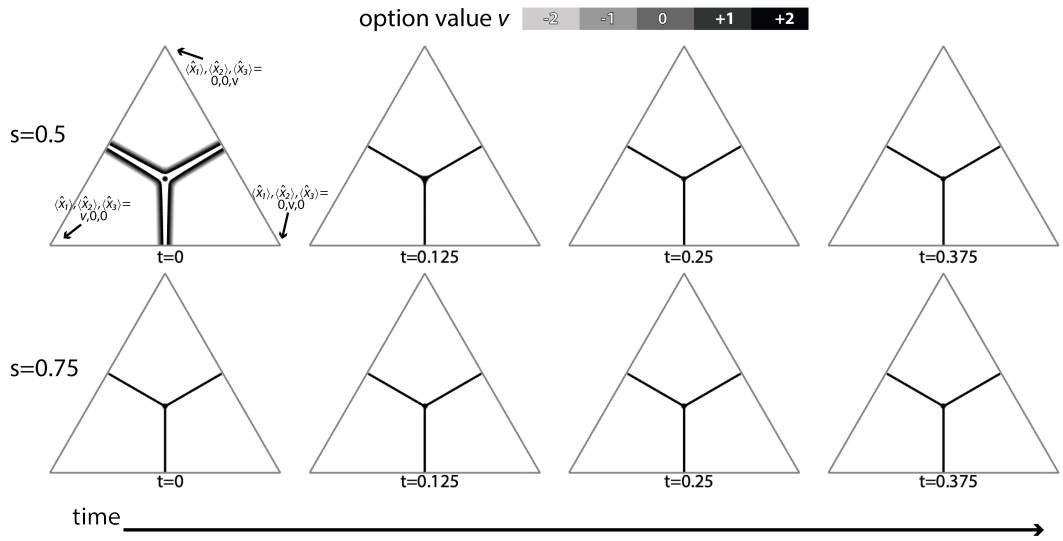


**Figure 3.** Optimal policies for linear time cost (Bayes Risk) rapidly transition from approximately linear subjective utility, and hence weakly magnitude-sensitive, decision boundaries in estimate space (Fig. 2, top row for $s = 0.25$; present figure, top row for $s = 0.5$, to more step-like subjective utility where immediate 'choose the best' decision-boundaries are necessarily magnitude-insensitive (bottom row for $s = 0.75$, and higher values of $s$). Triangles are low dimensional projections of the 3-dimensional evidence estimate space onto a plane moving along the equal value line, at value $v$ (***Tajima et al., 2019***). Dynamic programming parameters were: prior mean $\bar{x}_{p,i} = 1.5$ and variance $\sigma^2_{p,i} = 5$, and utility function parameters $m = 4$, $s \in \{0.5, 0.75\}$.

### Multi-Alternative Decisions: Optimal Policies Become Magnitude-Sensitive Under Geometric Discounting

As previously shown (***Marshall, 2019***; ***Steverson et al., 2019***), assuming geometric temporal discounting, the optimal policy for binary decisions is magnitude-sensitive. In ternary decisions, geometric discounting has the same effect; regardless of utility function linearity, the optimal policy is magnitude-sensitive (Fig. 2, row 'geometric').

## Numerical Simulations

Since noise-processing is fundamental in determining reaction times, we confirmed the results on magnitude sensitivity from our optimal policy analysis via numerical simulation of Bayes-optimal evidence-accumulating agents using those policies (see Materials and Methods). These numerical simulations confirmed the qualitative results from the optimal policy analysis; reaction times for ternary decisions under linear time costing are only weakly magnitude sensitive even for nonlinear subjective utility functions, while under geometric time costing reaction times become strongly magnitude sensitive for most utility functions examined.

### Multi-Alternative Decisions: Simulated Reaction Times are Weakly Magnitude-Sensitive for Nonlinear Subjective Utility Under Bayes Risk-Optimisation

Across all nonlinear subjective utility functions considered, linear time costing resulted in weakly magnitude-sensitive simulated reaction times (Fig. 4). This agrees with the weak magnitude-

sensitivity observed in the optimal policies derived above (Fig. 2). Note, however, that this contrasts with the binary decision case in which optimal policies, and hence reaction times, become magnitude sensitive under linear time cost when subjective utility is nonlinear (***Tajima et al., 2016***). An informal justification for this is given above in analysis the optimal decision boundaries computed via dynamic programming.

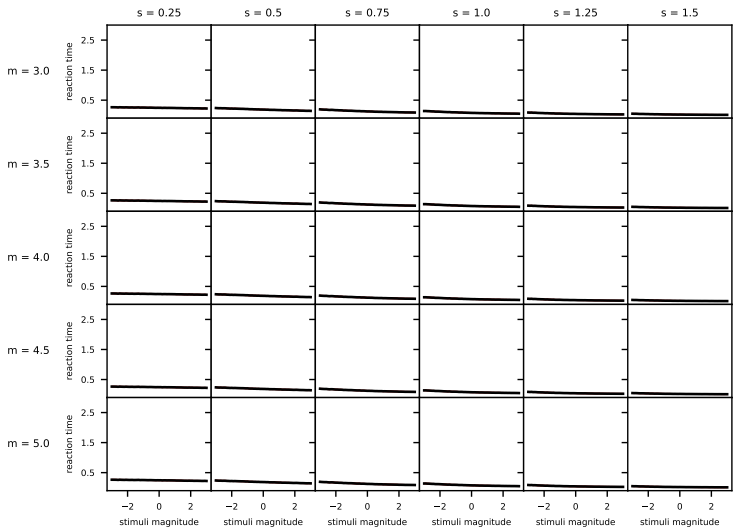


**Figure 4.** Linear time costs lead to weakly magnitude-sensitive simulated reaction times across a range of nonlinear subjective utility functions for equal value option sets. Simulation parameters were: prior mean $\bar{x}_{p,i} = 1.5$ and variance $\sigma^2_{p,i} = 5$, observation noise variance $\sigma^2_{a,i} = 2$, temporal cost $c = 0$, waiting time $t_w = 1$, and simulation timestep $dt = 5 \times 10^{-3}$. Lines are the mean reaction time for $10^4$ simulations, $95\%$ confidence intervals are shown as red shading (mostly invisible because smaller than the linewidth).

### Multi-Alternative Decisions: Simulated Reaction Times Become Magnitude-Sensitive Under Geometric Discounting

In contrast to linear time costing, across all nonlinear subjective utility functions considered, geometric time costing resulted in strongly magnitude sensitive simulated reaction times (Fig. 5), with longer reaction times for lower value equal-value option sets; this strategy was previously hypothesised to be optimal (***Pais et al., 2013***). The strong magnitude-sensitivity in the numerical simulations corresponds with the strong magnitude-sensitivity observed in the optimal policies derived above (Fig. 2).

## Discussion

In understanding the brain, which is a product of evolution, searching for optimal algorithms for typical decision problems can provide great insight. This normative approach can explain observed behaviours, and predict new behavioural patterns, based on evolutionary advantage. Yet the assumptions underlying such model analyses can prove crucial. Recently it has been asked what optimal decision algorithms look like for multi-alternative value-based choices, in which subjects

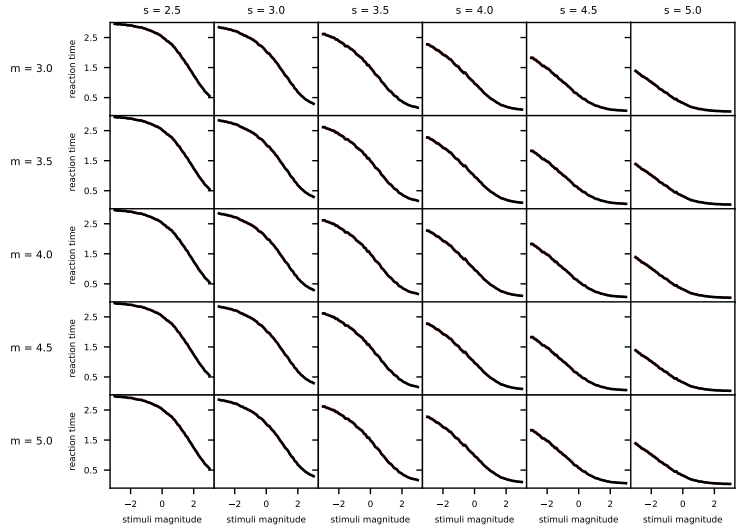


**Figure 5.** Geometric discounting of reward leads to strongly magnitude-sensitive simulated reaction times across a range of nonlinear subjective utility functions, with decisions postponed for low equal-value option sets. Simulation parameters were: prior mean $\bar{x}_{p,i} = 1.5$ and variance $\sigma^2_{p,i} = 5$, observation noise variance $\sigma^2_{a,i} = 2$, temporal cost $\gamma = 0.1$, and simulation timestep $dt = 5 \times 10^{-3}$. Lines are the mean reaction time for $10^4$ simulations, $95\%$ confidence intervals are shown as red shading (mostly invisible because smaller than the linewidth).

are rewarded not by whether their decision was correct or not, but by the value to them of the selected option (***Tajima et al., 2019***). The resulting algorithms correspond to earlier simple models for perceptual and value-based decision-making. These findings, however, rest on an assumption that time is a linear cost for subjects. Here we have shown that subjects do, however, exhibit marked magnitude sensitivity in ternary decisions, as previously shown for binary decisions. We have also shown that optimality theory that discounts future rewards multiplicatively based on time is the foremost explanation for such observations of magnitude-sensitivity; nonlinear subjective utility alone is not sufficient to give rise to strongly magnitude-sensitive decision times when time is treated as a linear cost.

## Behavioural Predictions

The Bayes Risk optimal policy is approximated by a neural model that is consistent with observations of economic irrationality (***Tajima et al., 2019***), hence it will be important to see if a revised neural model based on the revised optimal policy still shows such agreement. For example, while in the binary case magnitude-sensitive reaction times can be explained both by nonlinear subjective utility functions, and by multiplicative discounting rather than Bayes Risk, in the multi-alternative case our analysis suggests that the same phenomenon is explained *primarily* by multiplicative discounting of future rewards and not by nonlinear utility.

### Optimality Criteria

Practitioners of behavioural ecology have established principles to deal with empirically-observed deviations from the predictions of optimality theory (***Parker and Smith, 1990***); two of the most useful are to consider that the optimisation criterion has been misidentified, or the behaviour in question is not really adaptive. Tajima and colleagues employ an exemplary approach, attempting to combine the best of the approaches of normative and mechanistic modelling (***McNamara and Houston, 2009***); yet it bears remembering that subjects may not be trying optimally to solve the simple decision problem they are presented in the lab, but rather making use of mechanisms that evolved to solve the problem of living in their natural environment (***Fawcett et al., 2014***); indeed the experimental data presented here were produced by subjects who received no reward, yet nevertheless acted as if they were making an economic, value-based, decision rather than a purely perceptual, accuracy-based one.

## Materials and Methods

## Psychophysical Experiment

### Participants

This experiment was conducted during the COVID-19 pandemic, from May 13th – June 1st 2020.. Given that it was not possible to recruit participants for a laboratory experiment, we instead recruited them online using Pavlovia (***Peirce et al., 2019***; ***Peirce, 2007***), an online platform for psychophysical experiments implemented in PsychoPy.

Running a perceptual experiment online has a number of limitations: first, there is no way to ensure that participants are focused on the task and minimising distractions - to mitigate this we kept the task short and participants were instructed to concentrate on it; second, Pavlovia (as of March-May 2020) only allows stimuli to be drawn in units relative to window size (i.e., the window in which the experiment is displayed) or in pixels, hence their size and position relative to the fixation cross will vary across devices, depending on specific window sizes. However, even if the size and location of stimuli could be kept constant across participants, participants' distance from the screen cannot be controlled during an online experiment.

While in previous two-alternative experiments (***Teodorescu et al., 2016***; ***Pirrone et al., 2018a***) a limited number of participants (N < 10) performed a large number of trials, for our online experiment we aimed at a large number of participants performing a limited number of trials. This strategy is beneficial for online studies since the large number of participants helps ensure that variation in participants' motivation or viewing arrangements is averaged out. We therefore recruited 117 participants via external advertisement on Twitter, and internal email lists at the University of Sheffield (mean age = 40.4, SD = 11.2774, range 23 -77; 79 females, 37 males, 1 did not indicate their gender). We requested participants to follow the link to the experiment only if

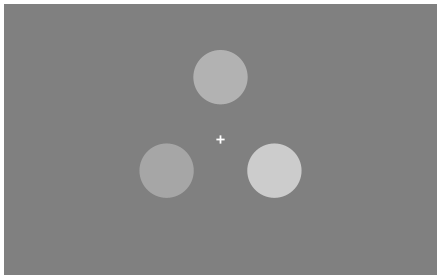

**Figure 6.** Stimuli example. Participants were requested to decide as fast and accurately as possible which of the three stimuli was brighter; they were asked to maintain fixation on the cross at the centre of the screen and minimise distraction for the short duration of the experiment. Unknown to participants, conditions of interest were conditions for which the stimuli had equal mean brightness.

aged 18 years or older. The experiment lasted about 5 minutes and participation was voluntary; participants did not receive any reward for their participation.

After reading the instructions, participants were informed that by continuing they were confirming that they understood the nature of the experiment and consented to participate. Participants were also informed that they could leave the experiment at any time by closing the browser. For this experiment, all procedures were approved by the University of Sheffield, Department of Computer Science Ethics Committee.

## Experimental setup

Similarly to previous studies (***Teodorescu et al., 2016***; ***Pirrone et al., 2018a***), stimuli consisted of three homogeneous, round, white patches in a triangular arrangement on a grey background, as depicted in Figure 6. Throughout the task participants were presented with a central fixation cross that they were requested to fixate on.

On a scale from 0 to 1 in PsychoPy, the patches could have a brightness of 0.3, 0.4, 0.5 or 0.6. There were $4^3$ = 64 possible trial combinations, of which 4 were equal alternatives (i.e., alternatives having a brightness of [0.3,0.3,0.3], [0.4,0.4,0.4], [0.5,0.5,0.5] or [0.6,0.6,0.6]). We selected 10 equal trial repeats, and only one trial repeat for all possible unequal alternatives, for a total of 100 trials per participant. On each frame, a Gaussian random variable with mean 0 and standard deviation of 0.25 $\times$ (mean brightness of the alternative) was added separately to the brightness level of each patch; the signal-to-noise ratio was thus kept constant across equal alternatives of different magnitude. The order of presentation of the alternatives was pseudo-randomised across participants and there was no systematic link between patch position and best option.

Three grey patches were presented simultaneously on the screen and subjects were asked to decide which of the three was brighter by pressing ‘left’, ‘right’ or ‘up’ on a keyboard using their second, fourth or third right-hand fingers, via a line-drawn diagram of a hand over a keyboard presented before the experiment began; specific instructions for left-handed participants were not provided, and we did not record handedness. The inter-trial interval, during which participants were presented with only the fixation cross, was selected at random between 0.5 seconds, 1 second

or 1.5 seconds for each trial. Subjects were instructed to be as fast and accurate as possible and to maintain their fixation on the cross at the centre of the screen throughout the experiment. Before the experiment they were presented with 6 training trials (unequal alternatives) to familiarise themselves with the task. Participants were not provided with any feedback after each trial and were not informed about the presence of the equal-alternatives conditions.

## Optimal policies

Optimal policy computations were performed in Matlab (***MathWorks, 2020***), and were adapted from the dynamic program of ***Tajima et al.*** (***2019***). Optimal policies were computed for Bayes Risk and geometric discounting, for linear utility ($r := x$), and for non-linear utility functions having the form

$$r := m\left(\frac{1}{1+e^{-sx}} - 1\right) \tag{1}$$

where $m$ and $s$ are shape parameters for the logistic function determining the interval of utilities and steepness of the slope respectively, and $x$ is the raw input value. We systematically varied the $m$ and $s$ parameters to test magnitude-sensitivity under Bayes Risk optimisation and geometric discounting, under a range of utility function shapes ranging from almost linear, to almost stepwise. Note that a sigmoid curve includes an interval in which subjective utility is an accelerating function of input value when the latter is negative, and an interval in which it is a decelerating function when the latter is positive; thus testing for magnitude-sensitivity over the full interval of raw input values tests a variety of utility function shapes over sub-intervals.

The Bellman equation used in the dynamic programming analysis for the Bayes Risk-optimisation case was

$$V(t, \hat{\boldsymbol{x}}(t)) = \max\left\{\max_i\{r_i(t, \hat{\boldsymbol{x}}(t))\} - \rho t_w, \langle V(t+\delta t, \hat{\boldsymbol{x}}(t+\delta t))\rangle - (c+\rho)\delta t\right\}, \tag{2}$$

where $V(t, \hat{\boldsymbol{x}}(t))$ is the value of the state estimates vector $\hat{\boldsymbol{x}}(t)$ at time $t$, $r_i(t, \hat{\boldsymbol{x}}(t))$ is similarly the expected reward from choosing the $i$-th reward, $\delta t$ is the time interval to the next decision point, $c$ is the linear cost per unit time, $\rho$ is the reward rate per unit time based on optimal decision-making over a sequence of trials, $t_w$ is the inter-trial waiting time, and $\langle \ldots \rangle$ is expectation over the next time interval ($\delta t$) (***Tajima et al., 2019***). For the results presented here we set $c = 0$, $t_w = 1$ and found the optimal $\rho > 0$ using the methods of ***Tajima et al.*** (***2019***); note, however, that since the prior was not varied this reward-rate optimisationcould not induce magnitude-sensitive reaction times in itself.

For the geometric discounting case the Bellman equation becomes

$$V(t, \hat{\boldsymbol{x}}(t)) = \max\left\{\max_i\{r_i(t, \hat{\boldsymbol{x}}(t))\}, \langle V(t+\delta t, \hat{\boldsymbol{x}}(t+\delta t))\rangle^{\gamma}\right\}, \tag{3}$$

where $0 < \gamma < 1$ is a discount factor for rewards received in future timesteps; this discount factor is per-unit-time, hence to discount a reward $\delta t < 1$ timesteps in the future the appropriate factor is

$\gamma^{\delta t/1} = \gamma^{\delta t}$.

## Stochastic simulations

Since noise processing is important in determining reaction times, we derived optimal decision policies as above, then tested them through numerical analysis of stochastic models. To test for magnitude-sensitive decision-making we examined the case of $n = 3$ equal-quality alternatives, in which we varied the magnitude of the (equal) stimuli values. Through these stochastic simulations, we tested the impact of the different temporal discount methods—linear or geometric—and of different utility functions—linear or nonlinear—on the decision speed (reaction time RT). The stochastic models simulate the sequential accumulation of evidence for the three alternatives $i \in \{1, 2, 3\}$ that are used to compute the expected rewards $\langle \hat{x}_i(t) \rangle$. The three evidence estimates can be represented, with a little abuse of notation, as a vector $\hat{\boldsymbol{x}}(t)$ denoting a point in a cube that represents the estimate space. In that cube, we also include the decision boundaries computed as above, indicating the separation of the estimate space into decision regions; in one region continuing to sample is expected to maximise utility, and in the other region taking a decision for the leading option is expected to be the best action.

In our simulations, each time step of length $dt$ the decision-maker accumulates three pieces of evidence, one for each option. Evidence for an option $i$ is sampled from the normal distribution $\mathbf{X_i} \sim \mathcal{N}(\bar{x}_i\, dt, \sigma^2_{a,i}\, dt)$, where each sample $x_{\tau,i}$ is a piece of momentary evidence at a small timestep of length $dt$ and with sequential index $\tau$, $\bar{x}_i$ is the true raw value (before any nonlinear utility transformation) of option $i$ and $\sigma^2_{a,i}$ is the variance in accumulation of evidence for $i$ (***Tajima et al., 2016***). Before observing any evidence, the decision maker has prior mean and variance, $\bar{x}_{p,i}$ and $\sigma^2_{p,i}$ for the distribution of $\mathbf{X_i}$. Each new piece of accumulated evidence is used by the decision maker to update the posterior expected reward as

$$\langle \hat{x}_i(t) \rangle = \frac{\bar{x}_{p,i}\, \sigma^2_{a,i} + \sum_{\tau=1}^{t} dx_{\tau,i}}{\sigma^2_{a,i} + \sigma^2_{p,i}\, t} \,. \tag{4}$$

## Code and Data Availability

Optimisation codes for the results presented here can be downloaded from https://github.com/DiODeProject/MultiAlternativeDecisions/tree/stochastic-sim

Experimental code and data for the results presented here can be downloaded from https://osf.io/8jumk/

## Acknowledgements

We thank Satohiro Tajima for sharing the code for the binary decision model. Dr Tajima was an exceptionally promising scientist who is sadly missed. We thank Jan Drugowitsch and Alex Pouget for discussions of their results and our own, and Thomas Bose, Nathan Lepora and Sophie Baker

for comments on an earlier draft.

## Declaration of Conflicting Interests

The authors declare that they have no conflicting interests.

## References

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

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and 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.

Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E., Trimmer, P. C., and McNamara, J. M. (2014). The evolution of decision rules in complex environments. *Trends in Cognitive Sciences*, 18(3):153–161.

Fudenberg, D., Strack, P., and Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108(12):3651–84.

Houston, A. I. and McNamara, J. M. (1999). *Models of adaptive behaviour: an approach based on state*. Cambridge University Press.

JASP Team (2020). JASP (Version 0.12.2)[Computer software].

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, 13(10):1292–1298.

Mangel, M. and Clark, C. W. (1988). *Dynamic modeling in behavioral ecology*. Princeton University Press.

Marshall, J. A. R. (2019). Comment on 'optimal policy for multi-alternative decisions'. *bioRxiv*.

MathWorks (2020). Matlab r2020b.

McNamara, J. M. and Houston, A. I. (2009). Integrating function and mechanism. *Trends in Ecology & Evolution*, 24(12):670–675.

Pais, D., Hogan, P. M., Schlegel, T., Franks, N. R., Leonard, N. E., and Marshall, J. A. R. (2013). A mechanism for value-sensitive decision-making. *PloS one*, 8(9):e73216.

Parker, G. A. and Smith, J. M. (1990). Optimality theory in evolutionary biology. *Nature*, 348(6296):27.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., and Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior research methods*, 51(1):195–203.

Peirce, J. W. (2007). Psychopy—psychophysics software in python. *Journal of neuroscience methods*, 162(1-2):8–13.

Pirrone, A., Azab, H., Hayden, B. Y., Stafford, T., and Marshall, J. A. (2018a). Evidence for the speed–value trade-off: Human and monkey decision making is magnitude sensitive. *Decision*, 5(2):129.

Pirrone, A. and Gobet, F. (2021). Is attentional discounting in value-based decision making magnitude sensitive? *Journal of Cognitive Psychology*.

Pirrone, A., Stafford, T., and Marshall, J. A. R. (2014). When natural selection should optimize speed-accuracy trade-offs. *Frontiers in neuroscience*, 8:73.

Pirrone, A., Wen, W., and Li, S. (2018b). Single-trial dynamics explain magnitude sensitive decision making. *BMC neuroscience*, 19(1):1–10.

Sellitto, M., Ciaramelli, E., and di Pellegrino, G. (2010). Myopic discounting of future rewards after medial orbitofrontal damage in humans. *Journal of Neuroscience*, 30(49):16429–16436.

Smith, S. M. and Krajbich, I. (2019). Gaze amplifies value in decision making. *Psychological Science*, 30(1):116–128.

Steverson, K., Chung, H.-K., Zimmermann, J., Louie, K., and Glimcher, P. (2019). Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. *Scientific Reports*, 9(1):1–14.

Tajima, S., Drugowitsch, J., Patel, N., and Pouget, A. (2019). Optimal policy for multi-alternative decisions. *Nature Neuroscience*, 22:1503–1511.

Tajima, S., Drugowitsch, J., and Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature Communications*, 7:12400.

Teodorescu, A. R., Moran, R., and Usher, M. (2016). Absolutely relative or relatively absolute: violations of value invariance in human decision making. *Psychonomic Bulletin & Review*, 23(1):22–38.

Turner, W., Feuerriegel, D., Andrejevic, M., Hester, R., and Bode, S. (2019). Perceptual change-of-mind decisions are sensitive to absolute evidence magnitude. *PsyArXiv*.

Zajkowski, W., Krzemiński, D., Barone, J., Evans, L., and Zhang, J. (2019). Reward certainty and preference bias selectively shape voluntary decisions. *bioRxiv*, page 832311.