This is a repository copy of *Magnitude-sensitive reaction times reveal non-linear time costs in multi-alternative decision-making*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/193043/

Version: Published Version

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

James A. R. Marshall[1,2]*, **Andreagiovanni Reina**[1,3], **Célia Hay**[4], **Audrey Dussutour**[4], **Angelo Pirrone**[5]

1 Department of Computer Science, University of Sheffield, Sheffield, United Kingdom, **2** Opteran Technologies, Sheffield, United Kingdom, **3** IRIDIA, Université Libre de Bruxelles, Brussels, Belgium, **4** Research Centre for Animal Cognition (CRCA), Centre for Integrative Biology (CBI), Toulouse University, Toulouse, France, **5** Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, London, United Kingdom

* james.marshall@sheffield.ac.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 multiplicative 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, in both humans and slime moulds. 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.

Author summary

Analysis of decisions based on option value (*e.g.* which pile of coins would you like?) suggests that the optimal rules correspond to simple mechanisms also known to be optimal for perceptual decisions (*e.g.* which light is brighter?) But, crucially, these analyses assume that the cost of time is linear—when the more usual assumption is made that time discounts multiplicatively (*e.g.* 'a bird in the hand is worth two in the bush (and so two in the hand are worth four in the bush)') then optimal decision-making looks quite different—in particular, the theory predicts that decision-making should be sensitive to the absolute magnitude of the opportunities, such as coin pile sizes, under consideration, in a way that the optimal perceptual mechanisms are not. As well as the theory, we present novel experimental evidence from human decision-making experiments, and foraging slime mould, of precisely such magnitude-sensitivity. This is a rare example of theory in behaviour

making a falsifiable prediction that is confirmed in two, highly divergent, species, one with a brain and one without.

Introduction

While the normative, optimal policy, approach to understanding decision-making is now well established for perceptual decisions (*e.g.* [1]), it has only recently been applied to value-based decisions [2–4]; 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.* [3–7]). Recently researchers have analysed multi-alternative value-based decision-making [4], building on earlier work in optimal decision policies for binary value-based choices [3]. Through sophisticated analysis based on the standard tool for solving such decision problems, stochastic dynamic programming [8, 9], the authors also present neurally-plausible decision mechanisms that may implement or approximate the optimal decision policies [3, 4]. 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 [3], and nonlinear time-varying thresholds that interpolate between best-vs-average and best-vs-next in the multi-alternative case [4].

Interestingly, the theoretically optimal policy for the binary decision case [3] is inconsistent with empirical observations of magnitude-sensitive reaction-times ([5, 10–14], but see [15]), 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 [3]. Furthermore, even single-trial dynamics have been observed to lead to magnitude sensitive reaction times [16]. While some descriptive models of decision-making can capture aspects of magnitude-sensitivity in the binary case [10, 17, 18], those models have not yet been extended to the multi-alternative case [5], hence their performance in the multi-alternatives case is unknown. Here we do not consider which extension of descriptive models could account for magnitude-sensitivity with multiple alternatives. Instead, we aim to fill two important gaps in the literature by establishing whether multi-alternative decision-making is magnitude-sensitive and what optimal policy could give rise to magnitude-sensitive multi-alternative decisions.

Previous analyses made an assumption that appears widespread in psychology and neuroscience, that decision-makers should optimise their Bayes Risk from such decisions [3, 4]; this is equivalent to maximising the expected value of decisions in which there is a linear cost for the time spent deciding [1, 6]. 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.* [1, 6]). 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 γ^2, and so on (see Materials and methods). This is a standard assumption in behavioural ecology [8, 9], 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.* [19], although this frequently suggests hyperbolic rather than geometric discounting, a fact that in itself merits an explanation based on optimality theory [20]).

Rederiving optimal policies to account for geometric [21] or general multiplicative [12] costs of time qualitatively changes them in the binary decision case, introducing magnitude-sensitive reaction times [12, 21]. However, disentangling these from nonlinear subjective utility is challenging, and cannot be excluded as an explanation for previous results [10–15, 22, 23].

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, both in human subjects undertaking a psychophysical task, and unicellular organisms engaged in foraging. 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

As we were testing theory developed to explain decision-making by animals with brains, we conducted psychophysical experiments with human subjects. However, we also conducted foraging experiments with a unicellular slime mould; testing theory across multiple species and behavioural tasks increases confidence when multiple agreements with theory are observed [11], and slime moulds have become a model system, with multiple experiments seeking to reproduce behavioural predictions from neuroscience and psychology [24–26].

Multi-alternative decisions in human psychophysical trials are magnitude-sensitive

Here we provide strong empirical evidence for magnitude sensitivity with multiple alternatives in humans, using an experimental paradigm similar to the one used to show magnitude sensitivity for two-alternative decision making [10, 11]. Details of the experiment (methods, participants, etc.) are reported in Materials and Methods. Participants had to choose which of three above-threshold grey patches was brighter in an online experiment (Fig 7A). 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 [11, 16, 23].

As previously done for binary decisions [11, 16], here we focused our analyses exclusively on equal alternatives. For the analyses, we did not censor any datapoints.

As shown in Fig 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 in Python) were made faster.

To assess if reaction times decreased as a function of the mean brightness of the equal alternatives, we used a linear mixed model in R. The model was fitted by specifying as fixed effect (explanatory variable) the brightness of equal alternatives as a continuous predictor. The
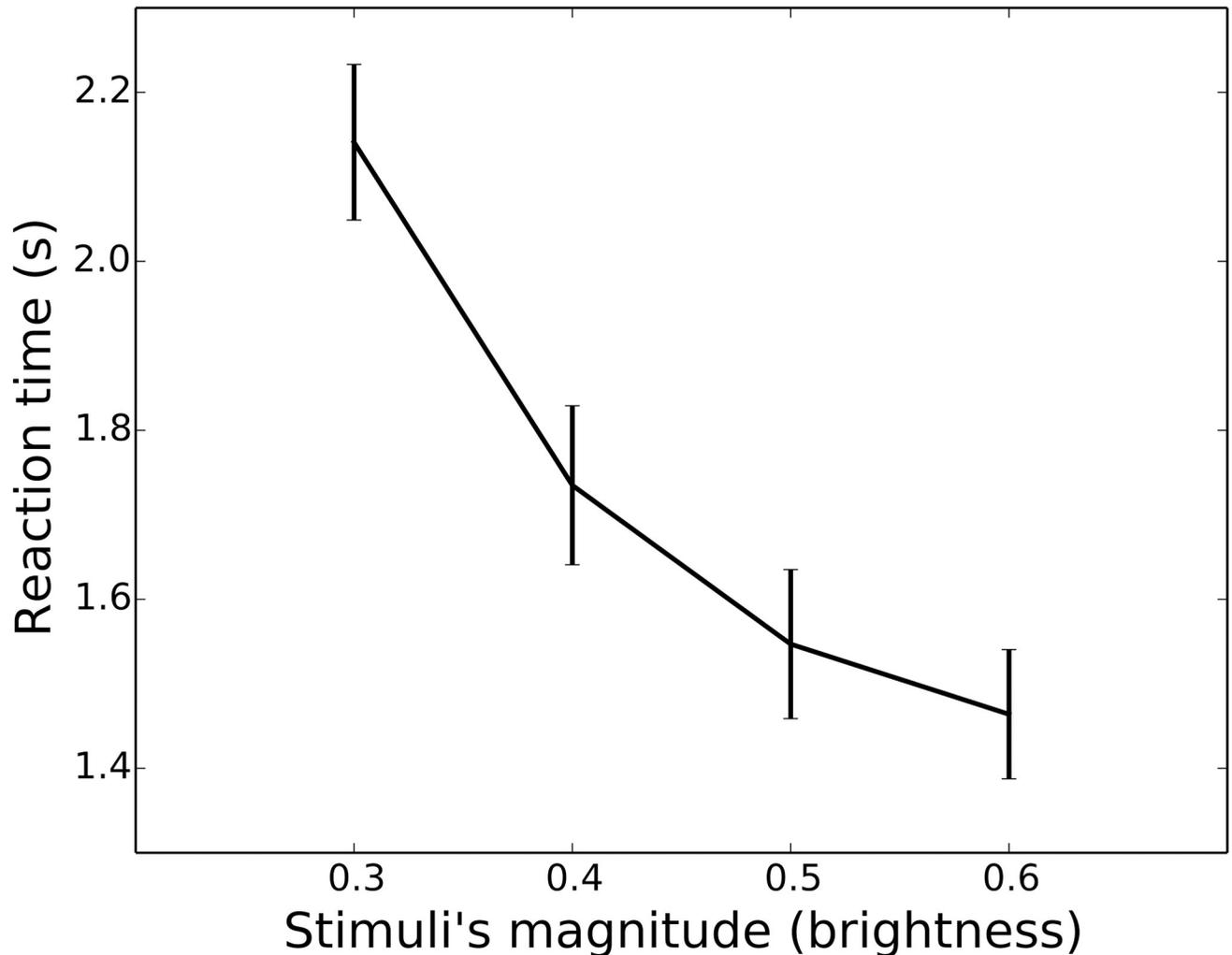
Fig 1. Empirical results from the behavioural online experiment. Decreasing reaction times as a function of the magnitude of the equal alternatives. X-axis presents mean brightness of equal alternatives (0.3, 0.4, 0.5, 0.6), on a scale of brightness from 0 to 1 in PsychoPy. Y-axis presents mean reaction times, in seconds. Bars show 95% confidence intervals. Participants experienced equal alternative conditions, interleaved with unequal alternative trials in pseudo-randomised order. Participants that performed the whole experiment experienced each equal alternative presentation ten times.

https://doi.org/10.1371/journal.pcbi.1010523.g001

participant ID was also added to the model as a factor for random effects. Reaction times significantly decreased as a function of the mean brightness of the alternatives (b = −1.95, p < .001, CI −2.14–1.75). Further details for the mixed-effect regression are presented in the supplementary information (S1 Table).

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 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 [10, 11]) should minimise effects due to randomly-distributed confounds.

Multi-alternative decisions in foraging trials by unicellular organisms are magnitude-sensitive

Unicellular organisms have also been suggested to implement optimal decision rules [27], have been used to test evidence accumulation theory [24, 25], and are describable with dynamical models closely related to neural network models [28]. Here, using slime moulds of the species *Physarum polycephalum*, we also observed strong empirical evidence for magnitude sensitivity with three alternative foraging sources. Details of the experiment are reported in Materials and Methods. Slime moulds were confronted with a choice offering three equal food sources (Fig 7B). We increased the magnitude of the options by increasing the quality of the food sources. As shown in Fig 2, the latency to reach one of the alternatives depended on the quality of the food sources; the higher the quality the faster the slime mould. This was confirmed by a linear mixed model similar to the one applied to the human data, in which reaction times significantly decreased as a function of food quality (b = −0.03, p < .001, CI −0.03—0.02; further details, S2 Table in the Supplementary Information).
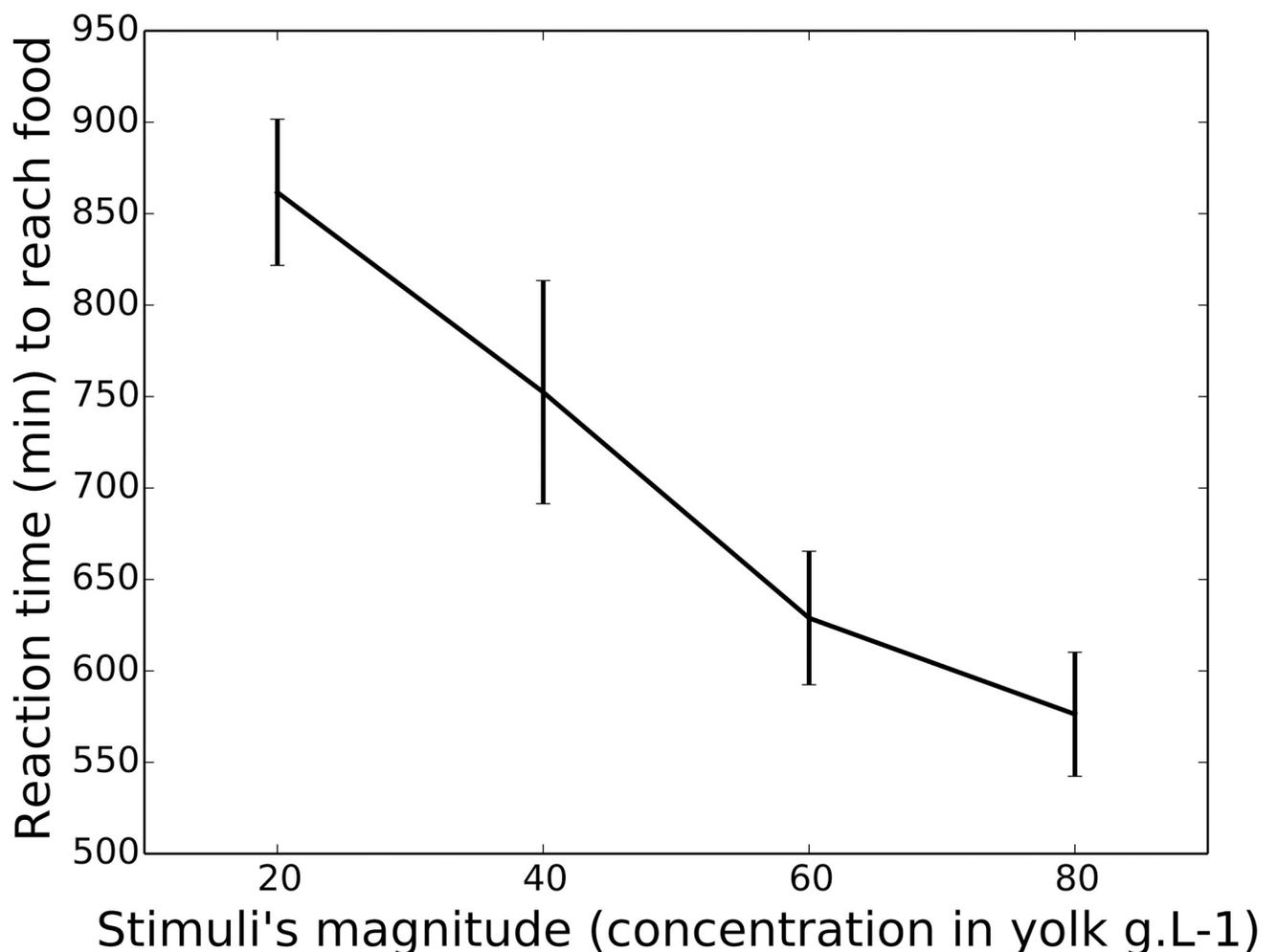
Fig 2. Empirical results from the slime mould experiment. Decreasing latencies to reach a food source as a function of the magnitude of the equal alternatives. X-axis presents the concentration in egg yolk of equal food sources (20, 40, 60, 80 g.L^{-1}). Y-axis presents mean latency to reach a food source, in minutes. Bars show 95% confidence intervals. 50 slime moulds were tested for each magnitude for a total of 200 slime moulds.

https://doi.org/10.1371/journal.pcbi.1010523.g002

Optimal policies

Tajima *et al.* used dynamic programming to derive optimal policies for decision makers receiving simultaneous evidence streams on the values of two [3] or three [4] different options, where the reward to the decision-maker is the value of the option selected, discounted by the time taken to make the selection. In the binary case, the optimal policy is approximated by a drift-diffusion model with collapsing boundaries [3]; that is, choice dynamics are largely dominated by value difference, and only under specific constraints is the optimal policy magnitude-sensitive [21]. In the three-alternatives case, the optimal policy is more complex and determined by nonlinear, time-dependent boundaries—while interested readers should refer to the original article [4] for details of the optimal policy, below we show that the optimal policy derived therein for three alternatives is only *weakly magnitude-sensitive*. 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 subjective 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 [12, 21] or nonlinear subjective utility with linear time costs [3]. 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 ([4] Fig 7) into either simple linear ones that collapse over time for lower-value option sets (Fig 3), or nonlinear boundaries that evolve over time similarly to the Bayes Risk-optimising case for higher-value option sets ([21]; Fig 3). 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 [4]; interestingly simply moving to nonlinear subjective utility with linear time costs simplifies the decision strategy to the 'best-vs-next' strategy (Fig 3; see [4], Fig 7).

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 [3, 4].

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 3, row 'linear'). Sensitivity analysis shows that magnitude-insensitivity is a general pattern (see next section). An informal understanding of this can be arrived at by appreciating that sigmoidal functions have two extremes of parameterisation (see S1 Fig); in one extreme they are almost linear, hence will be mostly magnitude insensitive due to the known result [3]. At the other extreme, the function becomes step-like; in this case options are either good or bad, and the optimal policy rapidly becomes 'choose the best' (Fig 4), 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.

Multi-alternative decisions: Optimal policies become magnitude-sensitive under geometric discounting

As previously shown [12, 21], assuming geometric temporal discounting, the optimal policy for binary decisions is magnitude-sensitive. In ternary decisions, geometric discounting has
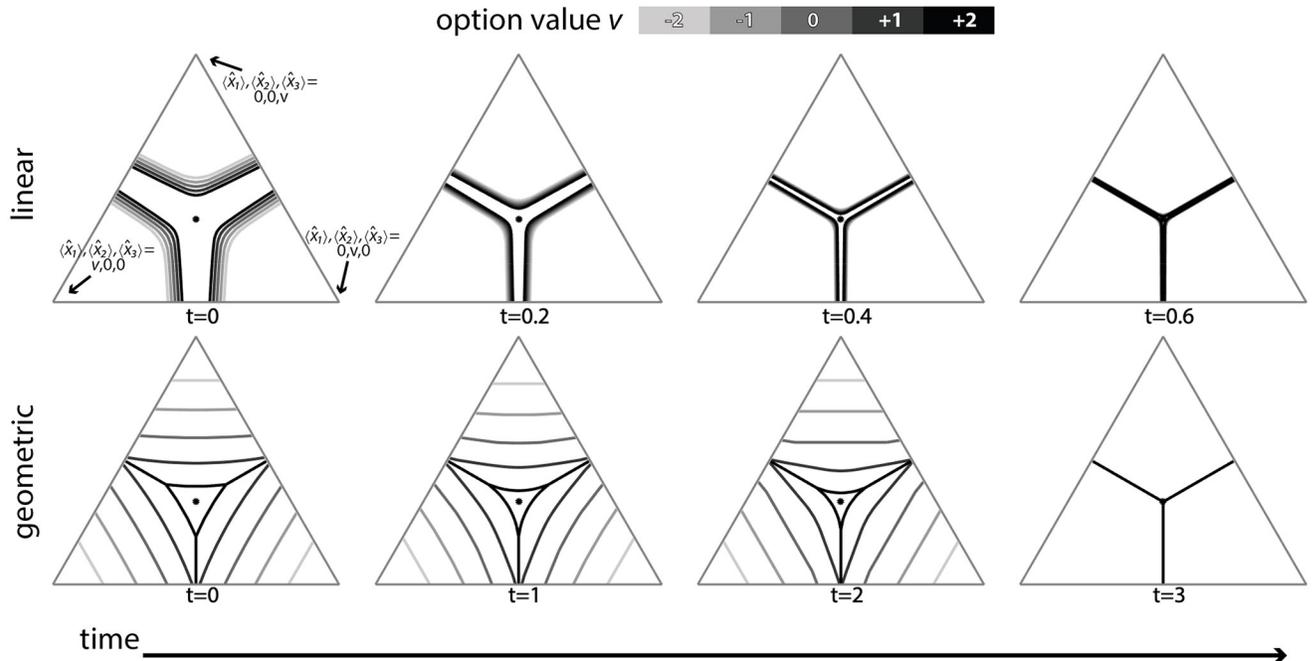
Fig 3. Optimal decision boundaries in the space of possible estimates of option values: Linear time costs lead to weakly magnitude-sensitive optimal policies (top row), while geometric discounting of reward leads to strongly magnitude-sensitive optimal policies (bottom row). In the linear time cost (Bayes 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 [4], 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 [4], 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 [4]. Dynamic programming parameters were: prior mean $\bar{x}_{p,i} = 1.5$ and variance $\sigma_{p,i}^2 = 5$, 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). Time steps chosen to illustrate boundary collapse.

https://doi.org/10.1371/journal.pcbi.1010523.g003

the same effect; regardless of utility function linearity, the optimal policy is magnitude-sensitive (Fig 3, 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 5). This agrees with the weak magnitude-sensitivity observed in the optimal policies derived above (Fig 3). 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 [3]. An
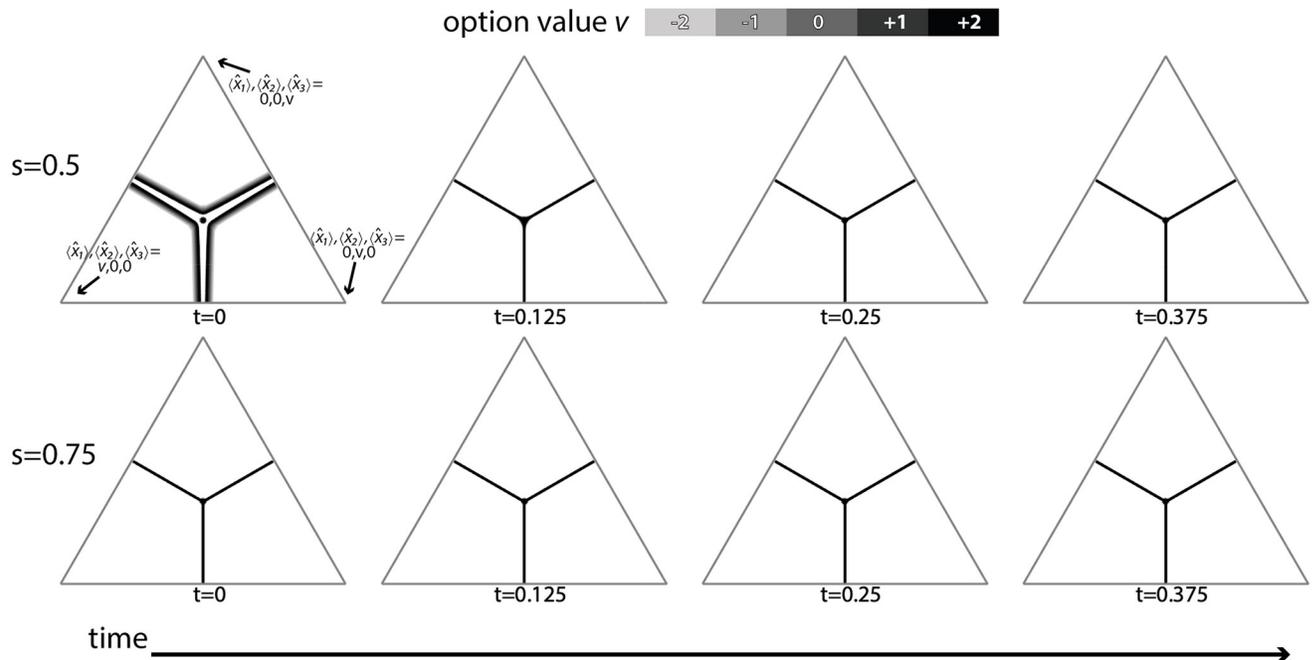
Fig 4. Optimal decision boundaries in the space of possible estimates of option values: Optimal policies for linear time cost (Bayes Risk) rapidly transition from approximately linear subjective utility, and hence weakly magnitude-sensitive, decision boundaries in estimate space (Fig 3, top row for *s* = 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* [4]. 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* ∈ {0.5, 0.75}. Time steps chosen to illustrate boundary collapse.

informal justification for this is given above in analysing the optimal decision boundaries computed via dynamic programming.

Multi-alternative decisions: Simulated reaction times can be 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 6 and S2 Fig), with longer reaction times for lower value equal-value option sets; this strategy was previously hypothesised to be optimal [29]. The strong magnitude-sensitivity in the numerical simulations corresponds with the strong magnitude-sensitivity observed in the optimal policies derived above (Fig 3). Note that the intervals of slope parameter *s* over which reaction times vary differ between linear (Fig 5) and geometric (Fig 6) time costing. The slope parameters used for the simulations interpolate between approximately linear and approximately piecewise utility functions, as shown in S1 Fig. We report the results for the same interval of slope parameter *s* as Fig 5 in S2 Fig and for linear utility function in S3 Fig, both of which also show strong magnitude-sensitivity.

Discussion

In understanding behaviour, 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.
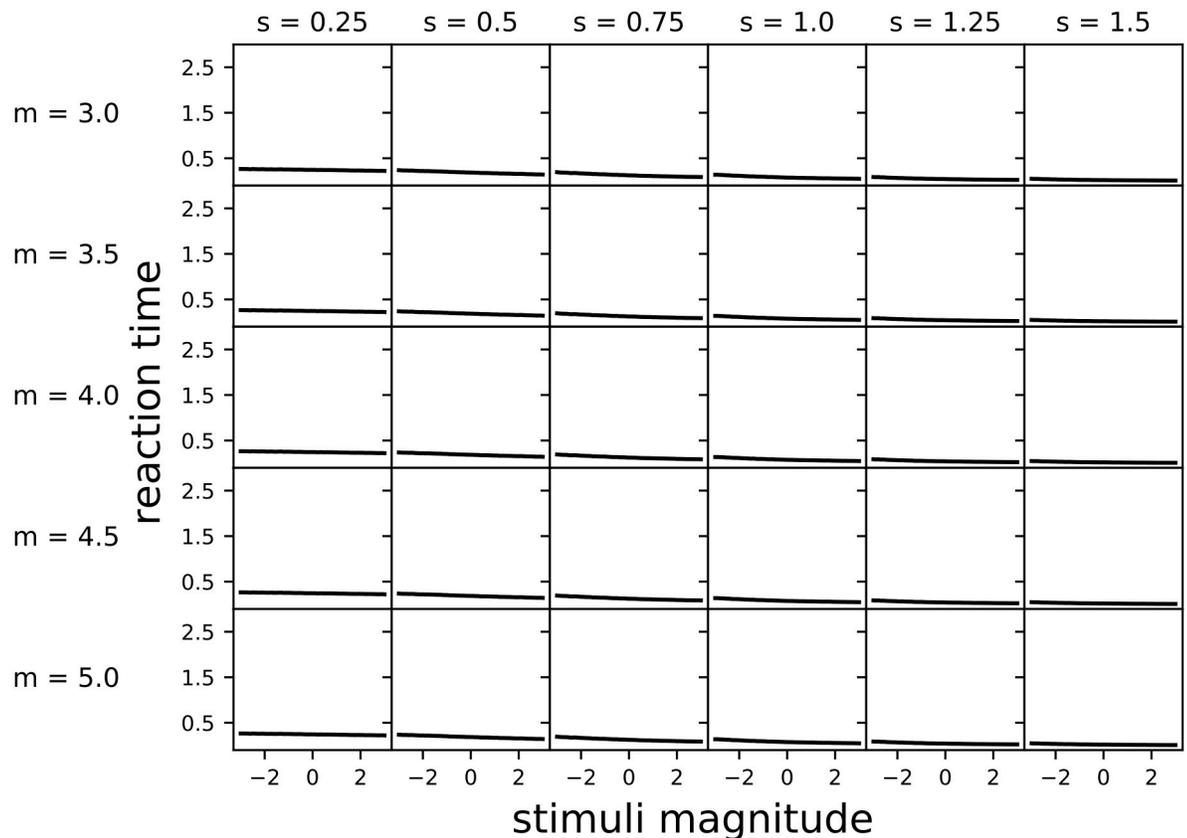
Fig 5. Linear time costs lead to weakly magnitude-sensitive simulated reaction times across a range of nonlinear subjective utility functions for equal value option sets. Simulation parameters were: prior mean $\bar{x}_{p,i} = 1.5$ and variance $\sigma_{p,i}^2 = 5$, observation noise variance $\sigma_{a,i}^2 = 2$, 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). Y-axis made consistent with Fig 6 for comparison. Non-decision time was implicitly zero.

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 are rewarded not by whether their decision was correct or not, but by the value to them of the selected option [4]. 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 deciding human subjects and foraging unicellular organisms do, however, exhibit marked magnitude sensitivity in ternary decisions, as previously shown for binary decisions [11, 26]. 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 [4], 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
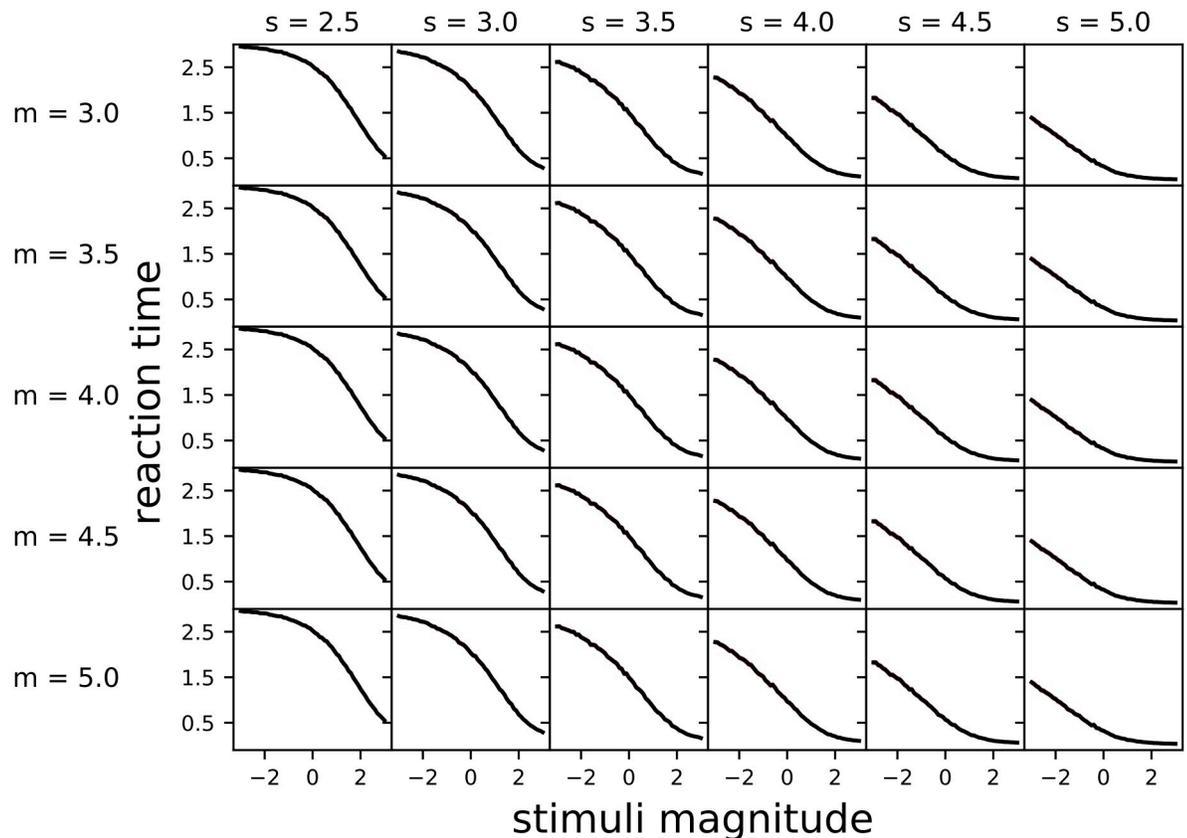
Fig 6. Geometric discounting of reward leads to strongly magnitude-sensitive simulated reaction times across a range of nonlinear subjective utility functions, with decisions postponed for low equal-value option sets. Simulation parameters were: prior mean $\bar{x}_{p,i} = 1.5$ and variance $\sigma^2_{p,i} = 5$, observation noise variance $\sigma^2_{a,i} = 2$, temporal cost $\gamma = 0.1$, and simulation timestep $dt = 5\times10^{-3}$. 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). Non-decision time was implicitly zero.

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 [30]; 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 [20]; 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 [31]; 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

Ethics statement. For this experiment, all procedures were approved by the University of Sheffield, Department of Computer Science Ethics Committee. Formal consent was not obtained as all participants were adults and all data were anonymised.

Participants. This experiment was conducted during the COVID-19 pandemic, from May 13th to June 1st 2020. Given that it was not possible to recruit participants for a laboratory experiment, we instead recruited them online using Pavlovia [32, 33], 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.

Notwithstanding these limitations, recent research has shown that web-based experiments yield reliable results, comparable to those obtained with lab-based experiments for reaction time tasks ([34] and references therein). Furthermore, our within-subjects study and analyses reduce the risk of conclusions driven by between-participant variability [35].

While in previous two-alternative experiments [10, 11] 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 aged 18 years or older. The experiment lasted about 5 minutes and participation was voluntary; participants did not receive any reward for their participation.

Participants were informed that the task involved 100 trials, and would take about 5 minutes to complete. After reading the instructions (see S1 Text), 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. As the experiment was conducted online and anonymously, verbal or written consent could not be provided by participants.

Experimental setup. Similarly to previous studies [10, 11], stimuli consisted of three homogeneous, round, white patches in a triangular arrangement on a grey background, as depicted in Fig 7. 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
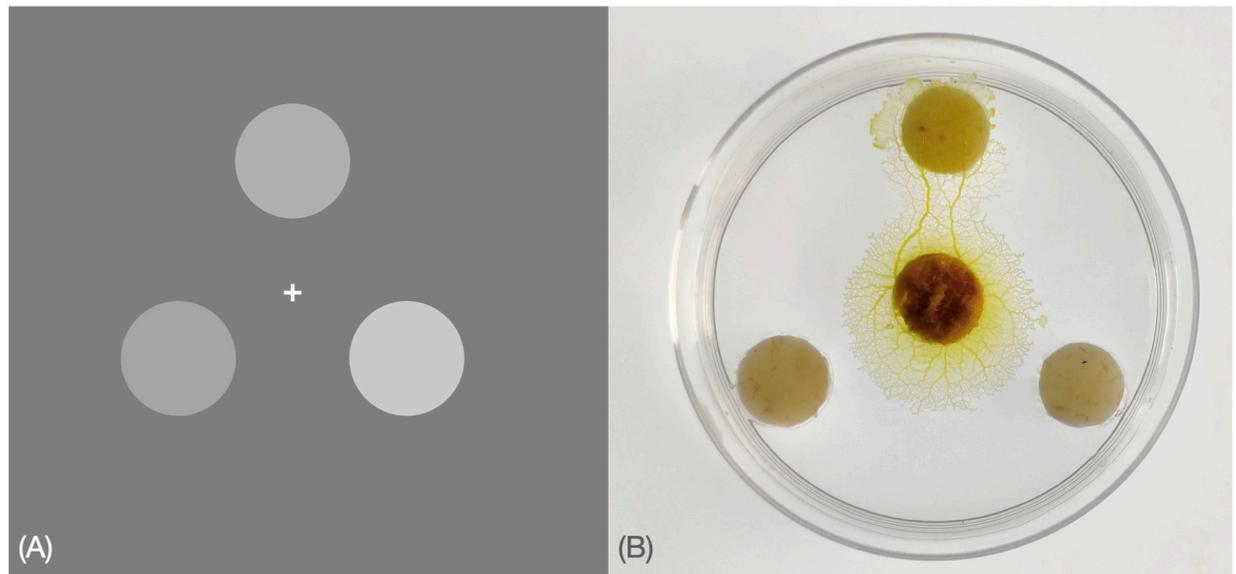
Fig 7. (A) Stimuli example for human psychophysical experiments: Participants were requested to decide as fast and accurately as possible which of the three stimuli was brighter; they were asked to maintain fixation on the cross at the centre of the screen and minimise distraction for the short duration of the experiment. Unknown to participants, conditions of interest were conditions for which the stimuli had equal mean brightness. (B) Photograph showing a slime mould that chose one food alternative among three equal ones. The slime mould was placed in the centre of a petri dish (60 mm ∅) filled with agar gel (10 g.L^{-1}) at a distance of 2 mm from each food alternative.

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 (see S1 Text); 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.

Slime mould experiment

Physarum polycephalum, also known as the acellular slime mould, is a giant polynucleated single cell organism that inhabits shady, cool, and moist areas. In the wild, *P. polycephalum* eats bacteria and dead organic matter. In the presence of chemical stimuli in the environment, slime moulds show directional movements (*i.e.* chemotaxis).

Slime moulds of strain LU352 kindly provided by Professor Dr Wolfgang Marwan (Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany) were used for the experiments. Slime moulds were initiated with a total of 10 sclerotia which are encysted resting stages. The sclerotia were soaked in water and placed in petri dishes (140 mm ∅) on agar gel (1%). Once revived, slime moulds start to explore the agar gel, usually 24h after

the reactivation of the sclerotia. The slime moulds were then reared for a month on a 1% agar medium with rolled oat flakes (Quaker Oats Company in Petri dishes (140 mm ∅). They were kept in the dark in a thermoregulated chamber at a temperature of 20 degrees Celsius and a humidity of 80%. The day before the experiment the slime moulds were transferred on a 10% oat medium (powdered oat in a 1% agar solution) in Petri dishes (diameter 140 mm). The experiments were carried out in a thermoregulated chamber and pictures were taken with a Canon 70D digital camera.

Slime moulds were presented with a choice between three equal food sources in an arena consisting of 60 mm diameter Petri dish filled with plain 1% agar. We punched three holes (10mm ∅) in the arena and filled them with a food source (10mm ∅). We used four different food patches varying in quality: 2% w/v powdered oat mixed with either 2, 4, 6 or 8% w/v egg yolk. Once the food sources were set in each hole, we placed a slime mould (10mm ∅) in the centre of the arena 2 mm away from each food source. We replicated the experiment 50 times for each food quality. For each replicate, we measured the time taken by the slime mould to reach either one of the three food sources.

To assess the difference in the latency to reach the food as a function of the food quality, we used a linear mixed model (function "lmer", package "lme4") in R (RStudio Version 1.2.1335). The models were fitted by specifying the fixed effects (explanatory variables) and the concentration in yolk (continuous predictor). The sclerotia identity was also added to the model as a random factor. We transformed the dependent variable using the "bestNormalize" function (package "bestNormalize"). The outcome of the model is presented in S2 Table.

Optimal policies

Tajima *et al.* set the general problem of a decision-maker that must integrate evidence simultaneously on the value of different competing options, then reach a decision where the reward for that decision is the true value of the option chosen, discounted by the time taken to reach that decision [4]. Optimal policy computations were performed in Matlab (Matlab 2020b), and were adapted from the dynamic program of [4]. Optimal policies were computed for Bayes Risk and geometric discounting, for linear utility ($r := x$), and for non-linear utility functions having the form

$$r := m\left(\frac{2}{1 + e^{-sx}} - 1\right), \tag{1}$$

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, δt is the time interval to the next decision point, c is the linear cost per unit time, ρ 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 (δt) [4]. For the results presented here we set $c = 0$, $t_w = 1$ and found the optimal $\rho > 0$ using the methods of [4]; note, however, that since the prior was not varied this reward-rate optimisation could not induce magnitude-sensitive reaction times in itself.

For the geometric discounting case the Bellman equation becomes

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

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 $\mathbf{X_i} \sim \mathcal{N}(\bar{x}_i \, dt, \sigma_{a,i}^2 \, dt)$, where each sample $x_{\tau,i}$ is a piece of momentary evidence at a small timestep of length dt and with sequential index τ, $\bar{x}_i$ is the true raw value (before any nonlinear utility transformation) of option i and $\sigma_{a,i}^2$ is the variance in accumulation of evidence for i [3]. Before observing any evidence, the decision maker has prior mean and variance, $\bar{x}_{p,i}$ and $\sigma_{p,i}^2$ 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

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

Due to the computational and memory cost of determining the optimal policies outlined in the preceding section, model fitting to empirical data was not practical. Although time units in the simulations are in arbitrary units they were tuned by hand so that, if interpreted as being measured in seconds, the intervals of the linear and geometric time cost cases overlapped with the empirically measured reaction times for the human psychophysics experiment, and with each other.

Supporting information

S1 Text. Task instructions for the human psychophysical experiments.
(PDF)

S1 Table. Mixed-effect regression for reaction times as a function of the brightness of the equal alternatives in the human study. Participant ID was included as a random factor. The regression was performed using R (RStudio Version 1.2.1335; function 'lmer', package 'lme4'). Given the typical skewness of reaction times, the dependent variable was transformed (*i.e.* normalized) using the 'bestNormalize' function in R. As the brightness of equal alternatives increased, reaction times significantly decreased.
(PDF)

S2 Table. Mixed-effect regression for reaction times as a function of food quality in the slime moulds study. Sclerotia identity was included as a random factor. The regression was performed using R (RStudio Version 1.2.1335; function 'lmer', package 'lme4'). Given the typical skewness of reaction times, the dependent variable was transformed (*i.e.* normalized) using the 'bestNormalize' function in R. As the food quality of equal alternatives increased, reaction times significantly decreased.
(PDF)

S1 Fig. Shape of the utility functions for different values of s and m of the logistic function of [Eq (1)](#) in the main text. The top panel shows the values used in [Fig 5](#) and SF1; the bottom panel shows the values used in [Fig 6](#).
(PDF)

S2 Fig. Geometric discounting of reward leads to 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 $\gamma = 0.9$. Non-decision time was implicitly zero. and simulation timestep $dt = 5 \times 10^{-3}$. Lines are the mean reaction time for 104 simulations, 95% confidence intervals are shown as red shading (mostly invisible because smaller than the linewidth).
(PDF)

S3 Fig. Geometric discounting of reward leads to magnitude-sensitive simulated reaction times also for linear subjective utility function. 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 $\gamma = 0.4$. Non-decision time was implicitly zero. and simulation timestep $dt = 5 \times 10^{-3}$. Lines are the mean reaction time for 104 simulations, 95% confidence intervals are shown as red shading (mostly invisible because smaller than the linewidth).
(PDF)

Author Contributions

Conceptualization: James A. R. Marshall, Angelo Pirrone.

Data curation: Angelo Pirrone.

References

1. Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychological Review. 2006; 113(4):700. https://doi.org/10.1037/0033-295X.113.4.700 PMID: 17014301

2. Fudenberg D, Strack P, Strzalecki T. Speed, accuracy, and the optimal timing of choices. American Economic Review. 2018; 108(12):3651–84. https://doi.org/10.1257/aer.20150742

3. Tajima S, Drugowitsch J, Pouget A. Optimal policy for value-based decision-making. Nature Communications. 2016; 7:12400. https://doi.org/10.1038/ncomms12400 PMID: 27535638

4. Tajima S, Drugowitsch J, Patel N, Pouget A. Optimal policy for multi-alternative decisions. Nature Neuroscience. 2019; 22:1503–1511. https://doi.org/10.1038/s41593-019-0453-9 PMID: 31384015

5. Pirrone A, Reina A, Stafford T, Marshall JAR, Gobet F. Magnitude-sensitivity: rethinking decision-making. Trends in Cognitive Sciences. 2022; 26(1):66–80. https://doi.org/10.1016/j.tics.2021.10.006 PMID: 34750080

6. Pirrone A, Stafford T, Marshall JAR. When natural selection should optimize speed-accuracy trade-offs. Frontiers in Neuroscience. 2014; 8:73. https://doi.org/10.3389/fnins.2014.00073 PMID: 24782703

7. Krajbich I, Armel C, Rangel A. Visual fixations and the computation and comparison of value in simple choice. Nature Neuroscience. 2010; 13(10):1292–1298. https://doi.org/10.1038/nn.2635 PMID: 20835253

8. Mangel M, Clark CW. Dynamic modeling in behavioral ecology. Princeton University Press; 1988.

9. Houston AI, McNamara JM. Models of adaptive behaviour: an approach based on state. Cambridge University Press; 1999.

10. Teodorescu AR, Moran R, Usher M. Absolutely relative or relatively absolute: violations of value invariance in human decision making. Psychonomic Bulletin & Review. 2016; 23(1):22–38. https://doi.org/10.3758/s13423-015-0858-8 PMID: 26022836

11. Pirrone A, Azab H, Hayden BY, Stafford T, Marshall JAR. Evidence for the speed–value trade-off: Human and monkey decision making is magnitude sensitive. Decision. 2018; 5(2):129. https://doi.org/10.1037/dec0000075 PMID: 29682592

12. Steverson K, Chung HK, Zimmermann J, Louie K, Glimcher P. Sensitivity of reaction time to the magnitude of rewards reveals the cost-structure of time. Scientific Reports. 2019; 9(1):1–14. https://doi.org/10.1038/s41598-019-56392-0

13. Zajkowski W, Krzemiński D, Barone J, Evans L, Zhang J. Reward certainty and preference bias selectively shape voluntary decisions. bioRxiv. 2019; p. 832311.

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

15. Bhui R. Testing Optimal Timing in Value-Linked Decision Making. Computational Brain & Behavior. 2019; 2(2):85–94. https://doi.org/10.1007/s42113-019-0025-9

16. Pirrone A, Wen W, Li S. Single-trial dynamics explain magnitude sensitive decision making. BMC Neuroscience. 2018; 19(1):1–10. https://doi.org/10.1186/s12868-018-0457-5 PMID: 30200889

17. Ratcliff R, Voskuilen C, Teodorescu A. Modeling 2-alternative forced-choice tasks: Accounting for both magnitude and difference effects. Cognitive Psychology. 2018; 103:1–22. https://doi.org/10.1016/j.cogpsych.2018.02.002 PMID: 29501775

18. Kirkpatrick RP, Turner BM, Sederberg PB. Equal evidence perceptual tasks suggest a key role for interactive competition in decision-making. Psychological Review. 2021; 128(6):1051. https://doi.org/10.1037/rev0000284 PMID: 34014711

19. Sellitto M, Ciaramelli E, di Pellegrino G. Myopic discounting of future rewards after medial orbitofrontal damage in humans. Journal of Neuroscience. 2010; 30(49):16429–16436. https://doi.org/10.1523/JNEUROSCI.2516-10.2010 PMID: 21147982

20. McNamara JM, Houston AI. Integrating function and mechanism. Trends in Ecology & Evolution. 2009; 24(12):670–675. https://doi.org/10.1016/j.tree.2009.05.011 PMID: 19683827

21. Marshall JAR. Comment on 'Optimal Policy for Multi-Alternative Decisions'. bioRxiv. 2019.

22. Smith SM, Krajbich I. Gaze amplifies value in decision making. Psychological Science. 2019; 30(1):116–128. https://doi.org/10.1177/0956797618810521 PMID: 30526339

23. Pirrone A, Gobet F. Is attentional discounting in value-based decision making magnitude sensitive? Journal of Cognitive Psychology. 2021;. https://doi.org/10.1080/20445911.2021.1890091

24. Latty T, Beekman M. Speed–accuracy trade-offs during foraging decisions in the acellular slime mould Physarum polycephalum. Proceedings of the Royal Society B: Biological Sciences. 2011; 278(1705):539–545. https://doi.org/10.1098/rspb.2010.1624 PMID: 20826487

25. Reid CR, MacDonald H, Mann RP, Marshall JAR, Latty T, Garnier S. Decision-making without a brain: how an amoeboid organism solves the two-armed bandit. Journal of The Royal Society Interface. 2016; 13(119):20160030. https://doi.org/10.1098/rsif.2016.0030 PMID: 27278359

26. Dussutour A, Ma Q, Sumpter D. Phenotypic variability predicts decision accuracy in unicellular organisms. Proceedings of the Royal Society B. 2019; 286(1896):20182825. https://doi.org/10.1098/rspb.2018.2825 PMID: 30963918

27. Perkins TJ, Swain PS. Strategies for cellular decision-making. Molecular Systems Biology. 2009; 5(1):326. https://doi.org/10.1038/msb.2009.83 PMID: 19920811

28. Zabzina N, Dussutour A, Mann RP, Sumpter DJ, Nicolis SC. Symmetry restoring bifurcation in collective decision-making. PLoS Computational Biology. 2014; 10(12):e1003960. https://doi.org/10.1371/journal.pcbi.1003960 PMID: 25521109

29. Pais D, Hogan PM, Schlegel T, Franks NR, Leonard NE, Marshall JAR. A mechanism for value-sensitive decision-making. PLoS one. 2013; 8(9):e73216. https://doi.org/10.1371/journal.pone.0073216 PMID: 24023835

30. Parker GA, Smith JM. Optimality theory in evolutionary biology. Nature. 1990; 348(6296):27. https://doi.org/10.1038/348027a0

31. Fawcett TW, Fallenstein B, Higginson AD, Houston AI, Mallpress DE, Trimmer PC, et al. The evolution of decision rules in complex environments. Trends in Cognitive Sciences. 2014; 18(3):153–161. https://doi.org/10.1016/j.tics.2013.12.012 PMID: 24467913

32. Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, et al. PsychoPy2: Experiments in behavior made easy. Behavior Research Methods. 2019; 51(1):195–203. https://doi.org/10.3758/s13428-018-01193-y PMID: 30734206

33. Peirce JW. PsychoPy—psychophysics software in Python. Journal of Neuroscience Methods. 2007; 162(1-2):8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017 PMID: 17254636

34. Bhui R. A statistical test for the optimality of deliberative time allocation. Psychonomic Bulletin & Review. 2019; 26(3):855–867. https://doi.org/10.3758/s13423-018-1555-1 PMID: 30684250

35. Brand A, Bradley MT. Assessing the effects of technical variance on the statistical outcomes of web experiments measuring response times. Social Science Computer Review. 2012; 30(3):350–357. https://doi.org/10.1177/0894439311415604