



# Normative decision rules in changing environments

Nicholas W Barendregt<sup>1\*</sup>, Joshua I Gold<sup>2</sup>, Krešimir Josi<sup>3</sup>, Zachary P Kilpatrick<sup>1</sup>

<sup>1</sup>Department of Applied Mathematics, University of Colorado Boulder, Boulder, United States; <sup>2</sup>Department of Neuroscience, University of Pennsylvania, Philadelphia, United States; <sup>3</sup>Department of Mathematics, University of Houston, Houston, United States

Abstract Models based on normative principles have played a major role in our understanding anomicienions. How is scall, have the forms definit stable conditions, and the conditions for relevance to decisions for more naturalistic, and conditions for an tions is unclear. We previously a normative derived a normative derived a normative derived a normative derive<br>`a a voice a normation model in which even a normative accumulation model in which evidence accumulation of th is a deductuation in the evidence-generations in the extension of  $\mathbb{I}$ (Glaze et al., 2015), but the evolution of commitment rules (e.g. thresholds on the accumulated evidence) under dynamic conditions is not fully understood. Here, we derive a normative model for  $\mathsf{c}$  den changing contexts, which we define a changing as changes in evidence  $\mathbb{R}$ over the course of a single decision. In the case of a single course, performance (reward in the course of a single using usin decision these thresholds thresholds that responds to an even anticipate the static to the static to the stati thresholds used in many decision models. We show that these adaptive thresholds exhibit several dia motial motifs that depend on the specific predicted and experienced and experienced and experienced changes an a a a umperform and input in models perform robustly. We function and input input in the input of the input o show that decision models with adaptive thresholds outperform those with constant or urgencygated thresholds in accounting for human response times on a task with time-varying evidence al\_ana ay ality ink normation in alcohon-making l experience of the assembly as dynamic as dynamic as dynamic as dynamic and use that update and use  $\mathbb{R}^n$  $\eta$  government and an and commitment.

\*For correspondence: [nicholas.barendregt@colorado.](mailto:nicholas.barendregt@colorado.edu) [edu](mailto:nicholas.barendregt@colorado.edu)

Competing interest: [19](#page-18-0)

Funding:  $\frac{19}{9}$  1.9

Preprinted:  $29$  | 2022 Received:  $03\%$  a  $,2022$ Accepted:  $20 \text{ C}$ , 2022 Published:  $25 \le \infty$  2022

Editor's evaluation<br>  $\frac{1}{2}$  a may  $\frac{1}{2}$  an m, anc, on This makes an important contribution to the study of decision-making in the study of decision-making evidence t<br>This provide convincing evidence that decision boundaries evidence that decision boundaries can be highly no reaching infinity in realistic regions. The originalist computation in the originalist and broad interest to b<br>Paper will be of broad in the computation of broad interest and the computation of the computation of the comp  $t$ theorists working under theorists working under the pressure. It is not the pressure  $t$ 

### Introduction

Even simple decisions can require us to a changing world. Should world you go the should world you go the should y park or through town on your walk? The answer can answer can answer can answer can answer conditions that could is it as an unique shower that is a senate shower that  $\mathbf{a}$ route (*[Figure 1A](#page-2-0)*) or a predictable sunrise that would nudge you toward the route with better vice the ubiquity of such dynamics in the real world, they are often neglected in the real world. The set of t is to understand to the brain many commonly models on the branch many commonly used models  $\mathbb{R}$ abmach **e**cision commitment of accumulated even and an option reaches an option reaches an option reaches and re fixed, predefined value or threshold (*[Wald, 1945](#page-21-0)*; *[Ratcliff, 1978](#page-21-1)*; *[Bogacz et al., 2006](#page-19-0)*; *[Gold and](#page-20-0)* 

<span id="page-2-0"></span>≹Ctal

notion in the structure of the structure of a 2-C distructure of a 2-C tasks, we are contained to represent the structure of a 2-C distructure of a 2-C distructure of a 2-C distructure of a 2-C distructure of a 2-C distruc environment for an observer with an initially unknown environmental state, *s* {*s*+,*s−*}, that uniquely determines which of two alternatives is correct. To infer the environmental state, this observer makes mear mnt,*ξ*, a ollo a transf±(*ξ*) = *f*(*ξ*) ≠ depends on the state mng correct choice is thus equivalent to determining the generating distribution, *f±*. An ideal Bayesian observer uses the log-likelihood ratio (LLR), *y*, to track their 'belief' over the correct choice (*[Wald,](#page-21-0)  [1945](#page-21-0)*; *[Bogacz et al., 2006](#page-19-0)*; *[Veliz-Cuba et al., 2016](#page-21-3)*). After *O* discrete observations *ξ*1:*O* that are independent across time, the discrete-time LLR belief *yn* is given by:

$$
y_n = \ln \frac{\Pr(s_+ | \xi_{1:n})}{\Pr(s_- | \xi_{1:n})} = \ln \frac{f_+ (\xi_n)}{f_-(\xi_n)} + y_{n-1}.
$$
 (2)

Given this defined task structure, we discretize the time during which the decision is formed and define the observer's actions during each timestep. The observer gathers evidence (measurements) during each timestep prior to a decision and uses each increment of evidence to update their belief about the correct choice. The observer has the option to experiment to a choice or make another measurement at the next timestep. By a utility of the contractions and the contractions of the set of t *V*+ **joc jest a** s<sub>+</sub>, a all *V*<sub>*r*</sub> **joc jest a** s<sub>−</sub>, and a all *V<sub>w</sub>* jam in a an), can jon c the value function for the observer given to the observer given to the observer given the observer given the o

```
V(p_n; \rho) = max{V+(p_n; \rho), V−(p_n; \rho), V_w(p_n; \rho)}
```
 $=$  max

ֆ锄鉗餩文Ă釸阀茬ő暔瑰圀㜀0ð@•ðPI%pÀ–—CbY™f€0ð¥'W-

<span id="page-4-1"></span><span id="page-4-0"></span>multiple reward changes during a single decision lead to complex threshold dynamics that we summarize in the several threshold threshold in shorter intervals and intervals and tendential tendential intervals and tendent to mag from simple monotonic changes in context parameters (*Figure 2*). To be range of the range of possible threshold motifs, when we found the single on the single on the single single single threshold motifs, when the single sin cang in diream is on the reagalist, in the reward for the reward-change task punishment **R**<sub>i</sub> = 0 and all m reward abruption abruption in the south its described by a Heaviside function and its dynamics are described by a H

$$
R_c(t) = (R_2 - R_1)H_{\theta}(t - 0.5) + R_1.
$$
\n(5)

 $\pm$ , a  $\pm$  c  $\pm$  ome eage a  $R_1$  of eage a  $R_2$  at = 0.5. For this single-change task, normative threshold dynamics exhibited several motifs that in some exemble fixed fixed or characteristic fixed or characteristic of  $\epsilon$  of  $\epsilon$  previous decision models but in  $\epsilon$ cally, and a novel different motion dynamic motion dynamic motion of the different dynamic motion of the motion of the characteristic cally, cases in response in response in the motion of the characteristic motion of the m to single, expected changes in reward contingencies for different combinations of pre- and postcange a al<sup>t</sup>



$$
\mu(t) = (\mu_2 - \mu_1)H_{\theta}(t - 0.5) + \mu_1.
$$
 ( )

<span id="page-6-0"></span>

eLi

<span id="page-8-0"></span>low-to-high reward switches – the regime differentiating strategy performance in ways that could be  $\eta$  if a.e. Adding not internal belief state to the internal belief state (which tends to the internal belief state state  $\mathbb{R}$ response distributions out to something the distributions of the distributions of the models the models the mod account for the additional noise does not alter the normative model: a range of the normative model: across a r  $\bullet$  a *y* $\bullet$   $\frac{1}{2}$   $\bullet$   $\frac{1}{2}$ , c na  $\frac{1}{2}$   $\frac{1}{2}$   $\frac{1}{2}$  and  $\overline{\sigma}_y$  and  $\overline{\sigma}_{mn}$  are the maximum  $\bullet$ strengths of sensory and  $\frac{1}{\epsilon}$ , the normative model of two normative model of two  $\epsilon$ when encountering low-to-high reward switches (*[Figure 4C](#page-7-0)*). This robustness arises because, prior to the cange, n51542n(a)19 1342n(a 9 [9 c)19 134nn aC20 a. ca Nac215093 [N](#page-9-0)otes Biology | Neuroscience<br>  $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ <br>  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$ <br>  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$ <br>  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$ <br>  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$ <br>  $\begin{pmatrix} 1 & 1 \\ 1 &$ 

<span id="page-9-0"></span>

<span id="page-10-0"></span>model, they used distinct model parameters, and thus different strategies, for both the fast and slow task conditions. This finding is clearer posterior parameter when looking at the posterior parameter distribution for each for each parameter distribution for each for subject and model parameter (see *[Figure 6—figure supplement 1](#page-10-0)* for an example). We speculate that the higher estimated value of reward in the slow task may arise due to subject and the subjects value of ere favorably more that the favorably suggested to the top subjects that the top substitution of the top substitution of the top subjects that the tendent to the tendent to the tendent to that the tendent to the tendent t  $1$  an adaptive strategy instead of the model of the model of the  $\alpha$ response data from dynamic context tasks.

#### **Discussion**

The goal of this study was to build on previous work in the most showing that in dynamic  $n$  showing that in dynamic environments, the most showing  $\alpha$ effective decision processes decisions do necessarily use  $\mathbb{R}^n$  . In the relatively simple, premany decision models (*[Bogacz et al., 2006](#page-19-0)*; *[Cisek et al., 2009](#page-19-2)*; *[Drugowitsch et al., 2012](#page-20-2)*), but instead adapt to learned or predicted features of the environmental dynamics (*[Drugowitsch et al., 2014a](#page-20-3)*). specially, when the mean induction of the members of the new task structures to demonstrate that normative decision commitment if (i.e., different rules or  $f$  or  $f$  in 'accumulate-to-bound' models) adapt to reward and evidence-performance-complex, and predictable, and predictable, and performanceof the matic density of the performance of classic to the performance of classic formation  $\mathbb{R}$ the advantage of normative models is maintained when computations are not c modeling the include to include the include the include the include the top include the that depends in a way t on stimulus history and the utility of commitment increases over the normative. We found that the normative th decision the total for the total also non-monopolonic and robust to noise. By analyzing man's is a a semi-dis, we met it is task mind the land same 0.4380.149 or model by a mo  $1$ 

## eLife Research article Computational and Systems Biology | Neuroscience

noisy normative model with  $\alpha$  and  $\alpha$  thresholds. Taken  $\alpha$  results shows that that ideal observers and ideal observers and the last normative and robust normative decision strategies in  $I_a$  is  $I_a$  in the city on  $n$  so  $n$  m  $n$  is  $a$ .

Our results can aid experimentalists investigating the nuances of complex decision-making in العصبية السابعة السابعة المتحدة المتحدة المعاملة المتحدة المتحدة المتحدة المتحدة المتحدة المتحدة ال eters for relatively simple tasks. For example, the reward-change tasks structure produces final example. For  $\alpha$ behavioral motifs, such as waiting until reward increases (*[Figure 2i](#page-4-1)*) and responding before reward decreases unless the accumulated evidence is ambiguous (*[Figure 2iv](#page-4-1)*). Using these kinds of modeling results to inform experimental design can help us understand the possible behaviors of  $\mathbb{R}^n$ subject data. Second, extending our and considering the sensitivity of  $\mathbb{R}^n$ model choice and task parameters (*[Barendregt et al., 2019](#page-19-3)*; *[Radillo et al., 2019](#page-21-4)*) will help to identify regions of task parameter space where models are most identifiable from observables like response time and c<sub>o</sub>ur choice. This and more generally, our work provides evidence that for tasks with gradual changes in evidence and reward in the reward of the reward in the more consistent with normative principal consistent with  $n$  is more con ples to models. However, with previously proposed here is needed to models. However, more work is needed to de and how permative principle principles for other dynamic-context tasks, such as those involving tasks, and tho a canglin not reward contingencies, by using normative or reward continuous continuous continuous continuous c subject strategies are plausible, the nature of tasks needed to identify them, and the relationship between tang dia matukang dikenali dengan di ke

Model-driven experimental design can aid in identification of adaptive decision rules in practice. People commonly encounter unpredictable (e.g. an abrupt thunderstorm) and predictable (e.g. sunset) context changes when making decisions. Natural extensions of common perceptual decision tasks (e.g. random-dot motion discrimination [*[Gold and Shadlen, 2002](#page-20-4)*]) could include within-trial changes in stimulus signal-to-noise anoise ratio and point and payout. The mayout. The mayout. The mayout. The aal om internal sources, including noise in neural sources, including  $\mathcal{A}$ and motor output (*[Ma and Jazayeri, 2014](#page-20-5)*; *[Faisal et al., 2008](#page-20-6)*). We assumed subjects do not have precise the strength or nature of the strength or nature of the sources, and they could not optimize their strategy accordingly. However, performance of  $\mathbb{R}^n$  and  $\mathbb{R}^n$  errors  $\mathbb{R}^n$ that results from such internal noise processes and adjusting on-line (*bonnen et al., 2015*). To extend the such  $\mathbb{R}$ the models we considered, we consider a considered, we could be a subject that subjects can estimate the magni of the information sensory and their normation to add use their contrategies to a contrategies to a  $m \rightarrow \bullet$  performance.

Real subjects likely do not rely on a single strategy when performing a sequence of trials (*[Ashwood](#page-19-5)  [et al., 2022](#page-19-5)*) and instead rely on a mix of near-normative, sub-normative, and heuristic strategies. In fitting subject data, experimentalists are thus presented with the different tasks and construction and a libr of possible models to use in their analysis. More general approaches have been developed for fitting response data to a broad class of models (*[Shinn et al., 2020](#page-21-5)*), but these model libraries are typically built on professure of the original contacts and make decisions. Because and make  $\alpha$  is and make  $\alpha$ the potential potential strategies is the call limitling a normative and analysis can both  $\alpha$ expand and provide into the range of possible subject behaviors in a systematic behaviors in a systematic and principles in a systematic behaviors in a systematic and principles in a systematic and principles in a systemat pled way. Understanding this scope will assist in developing a well-groomed candidate list of near-ງormative and ∱cimyor ∲.,∧ormanic and analysis of permance on a \_namci reward task produces threshold dynamics similar to those in *[Figure 2B](#page-4-1)*, then the fitting library should include a piecewise-constant threshold (or urgency signal) model. Combining these model-based investigations with model-free approaches, such as rate-distortion theory (*[Berger, 2003](#page-19-6)*; *[Eissa et al.,](#page-20-7)  [2021](#page-20-7)*), can also aid in identifying commonalities in performance and resource usage within and across model at the new order with the needs the need for pilot experience in the need for pilot experiments.

Our work complements the existing literature on optimal decision thresholds by demonstrating the diversity of the discussion of the magnetic conditions. Several defined  $\ell$ normative theories were, like ours, based on dynamic programming (*[Rapoport and Burkheimer,](#page-21-6) [1971](#page-21-6)*; *[Busemeyer and Rapoport, 1988](#page-19-7)*) and in some cases models fit to experimental data (*[Ditterich,](#page-20-8) [2006](#page-20-8)*). For example, dynamic programming was used to show that certain optimal decisions can represent an constant and the non-constant of our normative models in shame and tasks (*[Frazier and Yu, 2007](#page-20-9)*; *[Figure 2](#page-4-1)*). More recently, dynamic programming (*[Drugowitsch et al.,](#page-20-2) [2012](#page-20-2)*; *[Drugowitsch et al., 2014b](#page-20-10)*; *[Tajima et al., 2016](#page-21-7)*) or policy iteration (*[Malhotra et al., 2017](#page-20-11)*;

$$
V(p_n; \rho) = \max\{V_+(p_n; \rho), V_-(p_n; \rho), V_w(p_n; \rho)\}
$$
  
= max 
$$
\begin{cases} R_c p_n + R_i(1 - p_n) - t_i \rho, & \text{choose } s_+ \\ R \end{cases}
$$

SNR-change task thresholds

 $\mathcal{A}$  **c** a single task different and task difference task difference to vary  $\mathbf{a}$  and  $\mathbf{a}$  and  $\mathbf{a}$  and  $\mathbf{a}$  $\mu(t)$  a m - n n i **a**  $\lambda$ **o**  $n \rightarrow$ 

$$
\mu(t) = (\mu_2 - \mu_1)H_\theta(t - 0.5) + \mu_1.
$$

In *[Equation 6](#page-6-0)*, there is a single switch in evidence quality between pre-change quality *µ*1 and postchange quality *µ*2. This change occurs at *t* = 0.5. Substituting this quality time series into the likelihood transfer function in *[Equation 16](#page-15-0)* allows us to find the normative thresholds for this task as a function  $\rho$   $\mu_1$  and  $\mu_2$ . If modicing necessitation is a function of the time  $f_p$  also be a function of time; however, because the quality change points are known in advance to the observer, we can simply change between the specified and the specified and the specified terms in the specified terms and the specifie

 $\oint$ F ela $\hat{\mathcal{U}}$ changeshis'y cfTTQpcausdy changes.

eLi

<span id="page-19-7"></span><span id="page-19-6"></span><span id="page-19-5"></span><span id="page-19-4"></span><span id="page-19-3"></span><span id="page-19-2"></span><span id="page-19-1"></span><span id="page-19-0"></span>

<span id="page-21-7"></span><span id="page-21-6"></span><span id="page-21-5"></span><span id="page-21-4"></span><span id="page-21-3"></span><span id="page-21-2"></span><span id="page-21-1"></span><span id="page-21-0"></span>