Journal of Mathematical Psychology 88 (2019) 32–47

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Optimizing sequential decisions in the drift–diffusion model

Khanh P. Nguyen [a], Kre²imir Josi¢ [a,b,c,1], Zachary P. Kilpatrick [d,e,*,1]

[a] Department of Mathematics, University of Houston, Houston TX 77204, USA
[b] Department of Biology and Biochemistry, University of Houston, Houston TX 77204, USA
[c] Department of BioSciences, Rice University, Houston TX 77005, USA
[d] Department of Applied Mathematics, University of Colorado, Boulder, CO 80309, USA
[e] Department of Physiology and Biophysics, University of Colorado School of Medicine, Aurora, CO 80045, USA

highlights

- A normative model for evidence accumulation across correlated 2AFC trials.
- Initial beliefs of ideal observers depend on prior responses and sequence correlation.
- Reward rate is maximized by deliberating longer on early trials.
- Model accounts for experimentally observed repetition biases in 2AFC tasks.

article info

Article history:
Received 11 June 2018
Received in revised form 18 September 2018

Keywords:
Decision-making
Drift–diffusion model
Reward rate
Sequential correlations

abstract

To make decisions organisms often accumulate information across multiple timescales. However, most experimental and modeling studies of decision-making focus on sequences of independent trials. On the other hand, natural environments are characterized by long temporal correlations, and evidence used to make a present choice is often relevant to future decisions. To understand decision-making under these conditions we analyze how a model ideal observer accumulates evidence to freely make choices across a sequence of correlated trials. We use principles of probabilistic inference to show that an ideal observer incorporates information obtained on one trial as an initial bias on the next. This bias decreases the time, but not the accuracy of the next decision. Furthermore, in finite sequences of trials the rate of reward is maximized when the observer deliberates longer for early decisions, but responds more quickly towards the end of the sequence. Our model also explains experimentally observed patterns in decision times and choices, thus providing a mathematically principled foundation for evidence-accumulation models of sequential decisions.

1. Introduction

Organismal behavior is often driven by decisions that are the result of evidence accumulated to determine the best among available options (Brody & Hanks, 2016; Gold & Shadlen, 2007). For instance, honeybee swarms use a democratic process in which each bee's opinion is communicated to the group to decide which nectar source to forage (Seeley, Camazine, & Sneyd, 1991). Competitive animals evaluate their opponents' attributes to decide whether to fight or flee (Stevenson & Rillich, 2012), and humans decide which stocks to buy or sell, based on individual research and social information (Moat et al., 2013). Importantly, the observations of these agents are frequently uncertain (Brunton, Botvinick, and Brody 2013; Hsu, Bhatt, Adolphs, Tranel, and

* Corresponding author at: Department of Applied Mathematics, University of Colorado, Boulder, CO 80309, USA.
E-mail addresses: kpnguyen@math.uh.edu (K.P. Nguyen), josic@math.uh.edu (K. Josi¢), zpkilpat@colorado.edu (Z.P. Kilpatrick).
[1] Equal authorship.

& Newsome, 2001). The response trends and underlying neural activity are well described by the drift diffusion model (DDM), which associates a subject's belief with a particle drifting and diffusing between two boundaries, with decisions determined by the first boundary the particle crosses (Bogacz et al., 2006; Ratcliff & McKoon, 2008; Stone, 1960).

The DDM is popular because (a) it can be derived as the continuum limit of the statistically-optimal sequential probability ratio test (Bogacz et al., 2006; Wald & Wolfowitz, 1948); (b) it is an analytically tractable Wiener diffusion process whose summary statistics can be computed explicitly (Bogacz et al., 2006; Ratcliff & Smith, 2004); and (c) it can be fit remarkably well to behavioral responses and neural activity in 2AFC tasks with independent trials (Gold & Shadlen, 2002, 2007) (also see Latimer, Yates, Meister, Huk, and Pillow (2015)).

However, the classical DDM does not describe many aspects of decision-making in natural environments. For instance, the DDM is typically used to model a series of independent trials where evidence accumulated during one trial is not informative about the correct choice on other trials (

Fig. 1. Sequence of 2AFC trials with states H^n determined by a two-state Markov process with switching probability ϵ \D P(H^n D H^{n-1}). For instance, $H^{1:4}$ D ($H_C; H_C; H_C$) could stand for the drift direction of each drift diffusion process $y^n(t)$ corresponding to the decision variable of an ideal observer making noisy measurements, $\xi^n(t)$. The decision threshold (

an observer free to choose their response time can trade speed for accuracy in making a decision. This is typically modeled in the DDM by defining a decision threshold, θ_1; and assuming that at the first time, T_1, at which $|y^1(T_1)| \geq \theta_1$, the evidence accumulation process terminates, and the observer chooses $H = \text{sign}(y^1(T_1))$ D 1.

The probability, c_1, of making a correct choice in the free response paradigm can be obtained by analyzing the Kolmogorov backward equation (KBE) corresponding to Eq. (1) (Gardiner, 2009), as demonstrated in Appendix B. Given the initial condition $y^1(0) = y_0^1 = \frac{1}{2}\ln[P(H^1 = H_C)/P(H^1 = H)]$, threshold θ_1, and state $H^1 = H_C$, the probability of an exit through either boundary θ_1 is

$$\pi_{\theta_1}(y_0^1) = \frac{e^{(y_0^1+\theta_1)/D} - e^{-(y_0^1+\theta_1)/D}}{e^{2\theta_1/D} - e^{-2\theta_1/D}},$$

$$\pi_{-\theta_1}(y_0^1) = \frac{e^{(y_0^1+\theta_1)/D} - e^{-(y_0^1+\theta_1)/D}}{e^{2\theta_1/D} - e^{-2\theta_1/D}},$$

As the decision, $d_1 = -1$, determines the probability of each state at the end of trial 1, individual observations, $\xi^1_{1:s}$, are not needed to define the belief of the ideal observer at the outset of trial 2. The ideal observer uses a sequence of observations, $\xi^2_{1:s}$, and their decision on the previous trial, d_1, to arrive at the probability ratio

$$R^2_s = \frac{P(H^2 = H_C|\xi^2_{1:s}, d_1)}{P(H^2 = H_-|\xi^2_{1:s}, d_1)} = \frac{P(\xi^2_{1:s}|H^2 = H_C)P(H^2 = H_C|d_1)}{P(\xi^2_{1:s}|H^2 = H_-)P(H^2 = H_-|d_1)}.$$

Taking the logarithm, and applying conditional independence of the measurements $\xi^2_{1:s}$, we have that

$$L^2_s = \ln R^2_s = \sum_{j=1}^{s} \ln \frac{f_C(\xi^2_j)}{f_-(\xi^2_j)} + \ln \frac{P(H^2 = H_C|d_1)}{P(H^2 = H_-|d_1)},$$

indicating that $L^2_0 = \ln\left[P(H^2 = H_C|d_1)/P(H^2 = H_-|d_1)\right]$. Taking the temporal continuum limit as in Appendix A, we find

$$dy^2 = g^2 dt + \sqrt{2D}dW, \tag{11}$$

with the Wiener process, W, and variance defined as in Eq. (1). The drift $g^2 = \pm 1$ is determined by $H^2 = H_\pm$.

Furthermore, the initial belief is biased in the direction of the previous decision, as the observer knows that states are correlated across trials. The initial condition for Eq. (11) is therefore

$$y^2(0|d_1 = -1) = D \ln \frac{P(H^2 = H_C|d_1 = -1)}{P(H^2 = H_-|d_1 = -1)} = -D \ln \frac{(1-\epsilon)e^{\theta_1/D} + \epsilon}{\epsilon e^{\theta_1/D} + (1-\epsilon)} = -y^2_0, \tag{12}$$

where we have used Eqs. (4), (9), and (10). Note that $y^2_0 \to 0$ in the limit $\epsilon \to 0.5$, so no information is carried forward when states are uncorrelated across trials. In the limit $\epsilon \to 0$, we find $y^2_0 \to \theta_1$, so the ending value of the decision variable $y^1(T_1) = \theta_1$ in trial 1 is carried forward to trial 2, since there is no change in environmental state from trial 1 to 2. For $\epsilon \in (0, 1/2)$, the information gathered on the previous trial provides partial information about the next state, and we have $y^2_0 \in (0, \theta_1)$.

We assume that the decision variable, $y^2(t)$, evolves according to Eq. (11), until it reaches a threshold $\pm\theta_2$, at which point they

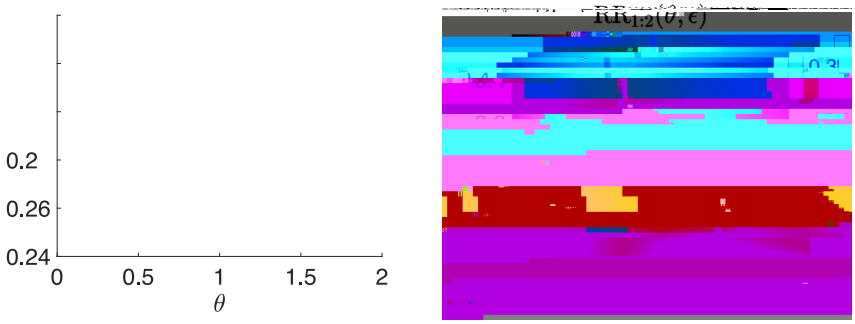

Fig. 2. The reward rate (RR) depends on the state change rate ϵ ∨D P(H^2 D H^1) for a constant threshold θ ∨D $\theta_{1\vee 2}$. Here and below stochastic simulations (dots) are compared to theoretical predictions (solid lines). A. The probability of a correct response in trials 1 and 2 ($c_{1\vee 2}$) only depends on the threshold θ, and increases with θ. B. Decision time DT_2 in trial 2 increases with the threshold θ and with the change rate, ϵ. Threshold crossings happen sooner on average for lower ϵ, since trial 2 is initiated closer to threshold. C. The reward rate (RR) as defined by Eq. (13) is unimodal. D. Colormap plot of RR as a function of θ and ϵ. The maximal RR increases and occurs at higher thresholds, θ; when state changes are less likely (lower ϵ): As ϵ decreases, the observer can afford to integrate evidence for a longer period of time in trial 1, since trial 2 will be much shorter (as shown in B). Here and in subsequent figures we used delay time T_D D 2 and noise amplitude D D 1.

threshold, and noise variance, D, as long as $\theta_2 > y_0^2$. However, when $\theta_2 < y_0^2$, the response in trial 2 is instantaneous (DT_2 D 0) and c_2 D (1 ϵ)c_1 C ϵ(1 c_1). Therefore, the probability of a correct response on trial 2 is defined piecewise

$$c_2 \text{ D } \begin{cases} \dfrac{1}{1 \text{ C e } ^{\theta_2=D}} & \vee\ \theta_2 > y_0^2; \\ \dfrac{(1 \quad \epsilon) \text{ C } \epsilon e^{\ \theta_1=D}}{1 \text{ C e } ^{\theta_1=D}} & \vee\ \theta_2 \quad y_0^2. \end{cases} \tag{15}$$

We will show in the next section that this result extends to an arbitrary number of trials, with an arbitrary sequence of thresholds, $\theta_{1\vee n}$.

In the case $\theta_2 > y_0^2$, the average time until a decision in trial 2, is (see Appendix D)

DT_2 D θ_2

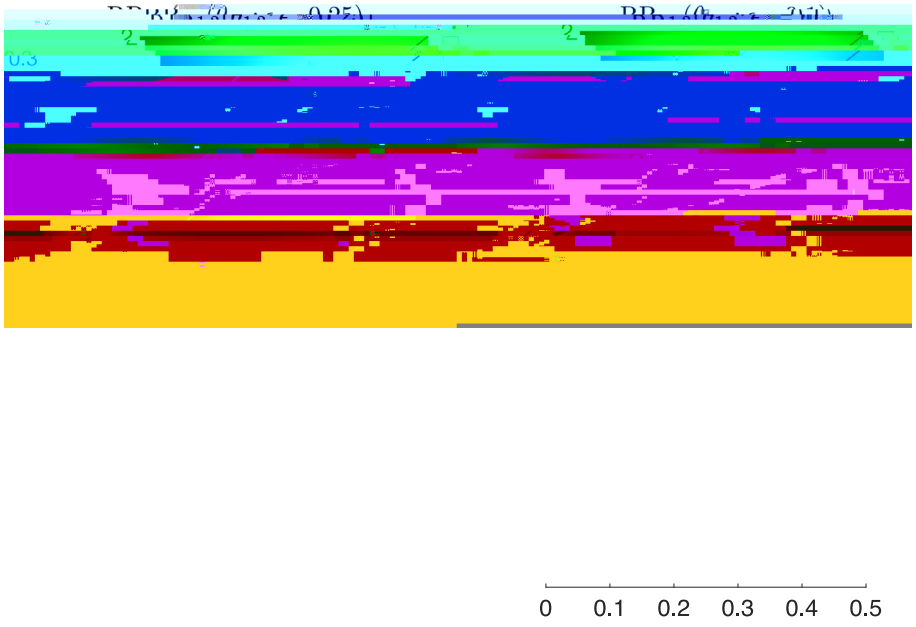

Fig. 3. A,B. Reward rate (RR), given by Eq. (17), depends on the change rate ϵ and thresholds $\theta_{1\backslash 2}$. For $\epsilon = 0.25$ (panel A), the pair $\theta_{1\backslash 2}^{\max}$ that maximizes RR (dot) satisfies $\theta_1^{\max} > \theta_2^{\max}$. Compare to the pair $\theta_{1\backslash 2}^{\max}$ (circle) for $\theta = \theta_{1\backslash 2}$ as in Fig. 2D. The boundary $\theta_2 = y_0^2(\theta_1$

Fig. 4. Reward rate ($RR_{1\backslash n}$) defined by Eq. (21) increases with n, and as ϵ is decreased, as shown for $\epsilon = 0.25$ (panel A); $\epsilon = 0.1$ (panel B); and $\epsilon = 0$ (panel C). The optimal decision threshold, θ^{max}; increases with n. The limiting reward rates $RR_{1\backslash n}(\theta \mid \epsilon) \to RR_1(\theta \mid \epsilon)$ as $n \to \infty$ defined by Eq. (22) are given by the topmost light curves.

a correct decision, c_j; is given by the more general, iterative version of Eq. (18). Therefore, while analytical results can be obtained for constant thresholds, numerical methods are needed to find the sequence of optimal dynamic thresholds.

The expression for the mean decision time, DT_j; on trial j is determined by marginalizing over whether or not the initial state y_0^j aligns with the true state H^j, yielding

$$DT_j = \begin{cases} \theta_j \dfrac{1 - e^{-\theta_j/D}}{1 + e^{-\theta_j/D}} - (1 - 2\epsilon)(2c_{j-1} - 1)y_0^j & \forall\ \theta_j > y_0^j; \\ 0 & \forall\ \theta_j \leq y_0^j. \end{cases}$$

Thus, Eq. (18), the probabilit0

Fig. 5. Optimal sets of thresholds, $\theta^{\max}_{1\backslash n}$; for n trials with correlated states. Top row: A. For n D 3, the set of optimal thresholds $\theta^{\max}_{1\backslash 3}$ obeys the ordering $\theta^{\max}_1$ $\theta^{\max}_2$ $\theta^{\max}_3$ for all values of ϵ, converging as ϵ ! 0:5 and as ϵ !

Fig. 7. Repetition bias trends in the adaptive DDM. A. The psychometric functions in trial j conditioned on d_{j-1} shows that when $d_{j-1} = +1$, the fraction of $d_j = +1$ responses increases (top curve). When $d_{j-1} = -1$, the fraction of $d_j = +1$ responses decreases (bottom curve). The weighted average of these two curves gives the unconditioned psychometric function (middle curve). Coherence g/D is defined for drift $g = 1$, noise diffusion coefficient D, and assumed change rate $\epsilon_{\text{assumed}} = 0.25$. Inset: When $\epsilon_{\text{true}} = 0.5$ is the true change rate, the psychometric function is shallower (gray curve). B. Correct probability c_{XY}

(marginal), and second (sequential) order effects (Jones et al., 2013; Kim et al., 2017), and predicts that decision times should decrease for repeated trials and increase for alternating trials when an observer assumes a correlated environment. However, these studies did not formalize the link between the decision threshold and subsequent initial condition, or study the problem of optimizing the reward rate. In addition, modeling studies have used generalized versions of the DDM, to account for trends observed in experimental data (Fründ et al., 2014; Goldfarb et al., 2012), but have not examined how to optimize a cost function, like the reward rate. Here we have linked the normative model of evidence-accumulation in correlated environments to a decision strategy that optimally trades off speed and accuracy by maximizing an overall reward rate.

The fact that repetitions decrease response time has long been known (Falmagne, Cohen, & Dwivedi, 1975; Remington, 1969). Whether this effect is driven by sensory after-effects of the stimuli, or the evidence and resulting responses is still debated (Pashler & Baylis, 1991). However, there is experimental evidence that past outcomes bias the initial belief on a subsequent trial in human observers (Fründ et al., 2014; Jones & Sieck, 2003). In contrast to our model, experimental studies also show that responses depend on more than just one preceding trial (reviewed in Luce (1986), Chapter 6.6). Such trends could be due to continuous adaptation of the evidence-discounting process to the statistics of the environment (Jones et al., 2013; Yu & Cohen, 2008). The experimentally observed increases in response time, and reductions in erroneous responses following a correct response (Laming, 1979), are consistent with a dynamical threshold model. Normative models of the type we suggest also cannot explain negative recency effects, *i.e* the observed tendency to alternate choices (Jarvik, 1951).

To account for information obtained on one trial observers could adjust accumulation speed (Diederich & Busemeyer, 2006; Ratcliff, 1985; Urai, de Gee, & Donner, 2018), threshold (Bogacz et al., 2006; Diederich & Busemeyer, 2006; Goldfarb et al., 2012), or their initial belief on the subsequent trial (Bogacz et al., 2006; Braun et al., 2018). We have shown that an ideal observer adjusts their initial belief and decision threshold, but not the accumulation speed in a sequence of dependent, statistically-identical trials. Adjusting the initial belief increases reward rate by decreasing response time, but not the probability of a correct response. Changes in initial bias have a larger effect than adjustments in threshold, and there is evidence that human subjects do adjust their initial beliefs across trials (Braun et al., 2018; Fründ et al., 2014; Olianezhad et al., 2016). On the other hand, the effect of a dynamic threshold across trials diminishes as the number of trials increases: $RR^{\max}_{1/n}$ is nearly the same for constant thresholds θ as for when allowing dynamic thresholds $\theta_{1/n}$. Given the cognitive cost of trial-to-trial threshold adjustments, it is thus unclear whether human subjects would implement such a strategy (Balci et al., 2011).

We do note, however, that the fact that the optimal variable-threshold model predicts a decrease in decision times may lend some perspective to the finding that subjects tend to increase the speed of decisions as an experiment goes on Elliott, Hansen, Mendoza, and Tremblay (2004) and Rabbitt and Vyas (1970). Past studies have suggested that subjects successively decrease their decision threshold until they make an error, and then increase the threshold on the next trial. This theory relies on the presence

rate can then lead to bounds that collapse or increase depending on the assumed prior distribution over the evidence reliability (Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012). One possible extension of our model would be to consider cases in which evidence quality can switch from one trial to the next, potentially causing the noise amplitude and threshold to adapt continuously within a trial in response to new evidence.

Modeling approaches can also point to the neural computations that underlie the decision-making process. Cortical representations of previous trials have been identified (Akrami, Kopec, Diamond, & Brody, 2018; Nogueira et al., 2017), but it is unclear how the brain learns and combines latent properties of the environment (the change rate ϵ), with sensory information to make inferences. Recent work on repetition bias in working memory suggests short-term plasticity serves to encode priors in neural circuits, which bias neural activity patterns during memory retention periods (Kilpatrick, 2018; Papadimitriou, Ferdoash, & Snyder, 2015). Such synaptic mechanisms may also encode previously accumulated information in sequential decision-making tasks, even in cases where state sequences are not correlated. Indeed, evolution may imprint characteristics of natural environments onto neural circuits, shaping whether and how ϵ is learned in experiments and influencing trends in human response statistics (Fawcett et al., 2014; Glaze, Filipowicz, Kable, Balasubramanian, & Gold, 2018). Ecologically-adapted decision rules may be difficult to train away, especially if they do not adversely impact performance (Todd & Gigerenzer, 2007).

There is evidence in humans that outcomes on previous trials bias future decisions. This is the case even when stimuli provide no information about either choice on a trial, so there is technically no correct decision (Bode et al., 2012). Indeed the decision made on such trials can be decoded from electroencephalogram recordings prior to stimulus presentation, suggesting it is primarily determined by evidence on the previous trial. It would be interesting

$p(y^0; t|y; 0)$ denote the conditional density of state y^0 at time t, with corresponding Kolmogorov equation,

$$\partial_t p(y^0; t|y; 0) = -\partial_y p(y^0; t|y; 0) + D\,\partial_y^2 p(y^0; t|y; 0);$$

given initial condition $p(y^0; 0|y; 0) = \delta(y - y^0)$ and boun
condition $p(-\theta; t|y; 0) = 0$. The absorbing boundaries at $y =$
also endow Eq. (B.1) with the condition $p(y^0; t|-\theta; 0) =$
$y^0 \in (-\theta; \theta)$. The survival probability, that $y \in (-\theta; \theta)$ at time
defined:

$$G(y; t) = \int_{-\theta}^{\theta} p(y^0; t|y; 0)\,dy^0 = P(T > t);$$

where T is a random variable, corresponding to the time the s
y leaves the domain $y \in (-\theta; \theta)$. Integrating Eq. (B.1), we obta

$$\partial_t G(x; t) = -\partial_x G(x; t) + D\,\partial_x^2 G(x; t);$$

Boundary conditions for Eq. (B.1) imply $G(-\theta; t) = 0$.

The mean first passage time, $T(y|\theta)$, for the state to de
the domain $(-\theta; \theta)$ given initial condition y is then given
marginalizing t against the flux $-\partial_t G(y; t)$ of survival probability
of $(-\theta; \theta)$:

$$T(y|\theta) = \int_0^{\infty} t\,(-\partial_t G(y; t))\,dt = \int_0^{\infty} G(y; t)\,dt.$$

Noting $\int_0^{\infty} G(y; t)\,dt = \lim_{t\to\infty} G(y; t) - G(y; 0) = -1$
integrating Eq. (B.2), we find

$$\partial_y T(y|\theta) + D\,\partial_y^2 T(y|\theta) = -1$$

with boundary conditions $T(-\theta) = 0$, which can be solved
itly:

$$T(y|\theta) = \theta\,\frac{\left[e^{\theta/D} + e^{-\theta/D} - 2e^{-y/D}\right]}{e^{\theta/D} - e^{-\theta/D}} - y.$$

B.2. Probability of exit through each boundary

To derive the probability of exit from either boundary
when the decision variable begins at y, we integrate the p
ity current through the boundary of interest, $J(-\theta; t|y$
instance, the probability that the decision variable exits
after time t is

$$g_\theta(y; t) = \int_t^{\infty} J(\theta; t'|y; 0)\,dt'$$
$$= \int_t^{\infty} \left[p(\theta; t'|y; 0) - D\partial_y p(\theta; t'|y; 0)\right] dt'.$$

Using the fact that $p(\theta; t|y; 0)$ satisfies Eq. (B.1), we find that g
satisfies

$$\partial_y g_\theta(y; t) + D\partial_x^2 g_\theta(y; t) = -\int_t^{\infty} \partial_{t'} J(\theta; t'|y; 0)\,dt'$$
$$= J(\theta; t|y; 0) = \partial_t g_\theta(y; t).$$

Taking $t \to 0$ and defining $\pi_\theta(y) = -g_\theta(y; 0)$, we see
$J(\theta; 0|y; 0)$ vanishes if $\theta \neq$ since $p(\theta; 0|y; 0) = \delta(y - \theta)$, so
right hand side goes to ze

$$\partial_y \pi_\theta(y) + D\partial_y^2 \pi_\theta(y) = 0;$$

where $\pi_\theta(\theta) = 1$ (θ, t) [illegible]

Using Eqs.

Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science*, *340*(6128), 95–98.

Chittka, L., Skorupski, P., & Raine, N. E. (2009). Speed–accuracy tradeoffs in animal decision making. *Trends in Ecology and Evolution*, *24*(7), 400–407.

Cho, R. Y., Nystrom, L. E., Brown, E. T., Jones, A. D., Braver, T. S., Holmes, P. J., et al. (2002). Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(4), 283–299.

Corcoran, A. J., & Conner, W. E. (2016). How moths escape bats: predicting outcomes of predator–prey interactions. *Journal of Fish Biology*, *219*(17), 2704–2715.

Dehaene, S., & Sigman, M. (2012). From a single decision to a multi-step algorithm. *Current Opinion in Neurobiology*, *22*(6), 937–945.

Deneve, S. (2012). Making decisions with unknown sensory reliability. *Frontiers in Neuroscience*, *6*, 75.

Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics*, *68*(2), 194–207.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, *32*(11), 3612–3628.

3628.

of accumulating6(Toward950 55a950 55.50tis19))0 555 051 rg0 55n lA., Science

hesis.)]TJ57,)]TJ0 0.50195 0.67451 rg 0 0.50195 0.67451 RG0 0.50195 0.67451 rg 0 0.50195 0.67451 RG [-250(194)]TJ0 0.50195 0.67451 rg 0 0.50195 0.7451 RG0 0.50

5 0.67451 RG 11.955 -9.066 Td [(optimally)-182(accumulate)-182(evidence)-ter(of)-74in p.50tis19) 73(p5 051 451 RG0 0.50195 0.67451 rg 0 0.50195 0.67451 RG/F56 6.6311

95 0.67451 rg 0 0.50195 0.67451 RG [-374(The)-374(cost)-373(of)-374(accumulatin29(I.5015(on)-325e.5(2Tr)]TJ07(judgments.5(2Trbiliing.)2Tr(on)-)2Trser(stimul80f674507((on)g 0

Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic Bulletin & Review*, *19*(1), 139–145.

Stevenson, P. A., & Rillich, J. (2012). The decision to fight or flee insights into underlying mechanism in crickets. *Frontiers in Neuroscience*, *6*, 118.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*(3), 251–260.

Todd, P. M., & Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science*, *16*(3), 167–171.

Urai, A. E., de Gee, J. W., & Donner, T. H. (2018). Choice history biases subsequent evidence accumulation. bioRxiv, 251595.

Usher, M., & McClelland, . J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, *108*(3), 550.

Veliz-Cuba, A., Kilpatrick, Z. P., & Josić, K. (2016). Stochastic models of evidence accumulation in changing environments. *SIAM Review*, *58*(2), 264–289.

Wald, A., & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, *32*, 6–339.

Yu, A. J., & Cohen, J. D. (2008). Sequential effects: Superstition or rational behavior?. *Advances in Neural Information Processing Systems*, *21*, 1873–1880.