Evidence for effort prediction in perceptual decisions

Nisheeth Srivastava
Dept of Computer Science, IIT Kanpur

Abstract

The classic drift diffusion model of the 2AFC choice process assumes that observers select evidence accumulation thresholds to optimize some desired level of accuracy across the experiment. We argue that it is more ecologically natural to assume that decision-makers set this threshold adaptively, using information from recent trials to adjust it for upcoming ones. To test this hypothesis, we designed and conducted a pair of random dot motion discrimination experiment where the coherence parameter that controls task difficulty varies across trials in a predictable manner. To analyze data from these experiments, we also designed a hierarchical drift diffusion model that allows decision-makers to adapt their evidence threshold based on the trend of difficulty of previous trials. Our results suggest that observers rationally integrate cross-trial information about trial difficulty into perceptual decision-making by adjusting their internal evidence thresholds. We briefly discuss the implications of the existence of such trial-level effort inference on contemporary models of the choice process.

Keywords: drift diffusion model; ideal observer model; Bayesian modelling; cognitive effort; rational inference

Introduction

The drift diffusion model (DDM) is a very successful sequential sampling model of the choice process (Ratcliff & McKoon, 2008). Particularly when applied to perceptual decision-making tasks, where the stream of evidence is transparent to the experimenter, this model has shown excellent fits to choice and response time data from a wide variety of experimental paradigms, even generalizing across organisms (Brunton, Botvinick, & Brody, 2013). While it shares several components, including parallel accumulation and race-to-a-threshold with other competing paradigms such as leaky competing accumulation (Usher & McClelland, 2001) and decision field theory (Busemeyer & Townsend, 1993), its stochastic specification of important components of the choice process - rate of accumulation of evidence, response bias, and variability in the evidence accumulation rate - gives it excellent flexibility and interpretability in modelling the summary statistics seen in perceptual decision-making experiments.

While on the one hand, its mathematical construction makes the DDM an excellent *descriptive* model of the choice process, it simultaneously makes it challenging to associate its optimality criterion with the goals and costs faced by real decision-makers. Specifically, the drift diffusion model is known to implement the sequential probability ratio test (SPRT) (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006), which is statistically optimal in the sense that for any choice of the decision threshold, using the DDM criterion for choice will yield the highest possible accuracy (Wald & Wolfowitz, 1948).

The relationship between the SPRT and DDM imbues it with a normative sense of optimality - observers are being statistically optimal in the SPRT-sense if we show that the DDM model fits their behavior well. But the underlying assumptions of SPRT - perfect evidence integration, approximately linear evidence accumulation rates, race to a fixed evidence threshold - are not good fits for the information and processing limitations that organisms face in real-world decision-making scenarios. In recent years, objections to these premises have been raised on both computational and empirical grounds. Deneve has documented how the conventional drift diffusion paradigm fails to accommodate situations where sensory inputs are unreliable (Deneve, 2012). Thura and colleagues have shown how reaction time distributions in perceptual decision-making tasks may be better described by evidence accumulation terminated by breaching a time-collapsing threshold responsive to an increasing 'urgency' signal than classic accumulation to a fixed threshold (Thura, Beauregard-Racine, Fradet, & Cisek, 2012). Glaze et al have demonstrated how perfect evidence integration - a fundamental assumption of drift diffusion models - is sub-optimal in the face of unsignalled context shifts in the decision-making environment (Glaze, Kable, & Gold, 2015). Thus, using DDM as a normative baseline, research is increasingly focusing on identifying aspects of the environment that constrain real-world decision-making.

Threshold adaptation as effort inference

We propose to revise the premise that decision-makers accumulate evidence up to a fixed threshold. Whereas such proposals have been made previously (Thura et al., 2012), they have focused on incorporating the opportunity cost of continued sampling in the form of a threshold that decreases over time (Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012). We focus on a different aspect of the threshold determination process - that decision-makers are likely to use information from previous trials to set decision thresholds for upcoming trials. Here again, it is well-documented that patterns of responding can introduce response biases in experiments (Ratcliff & McKoon, 2008). The drift diffusion model allows such biases to be modeled explicitly using changes in diffusion start-point parameters.

We consider a different normative possibility: we propose to model the threshold parameter used in drift diffusion modeling as a proxy for the amount of effort the observer believes is necessary for adequate performance, and we intend to investigate whether human observers can infers the effort needed for upcoming trials using effort observations in recent trials. Grounding this hypothesis in a perceptual decision-making task, we model behavior on this task using a hierarchical ideal Bayesian observer that performs 2AFC random dot motion (RDM) discrimination. The lower level of this

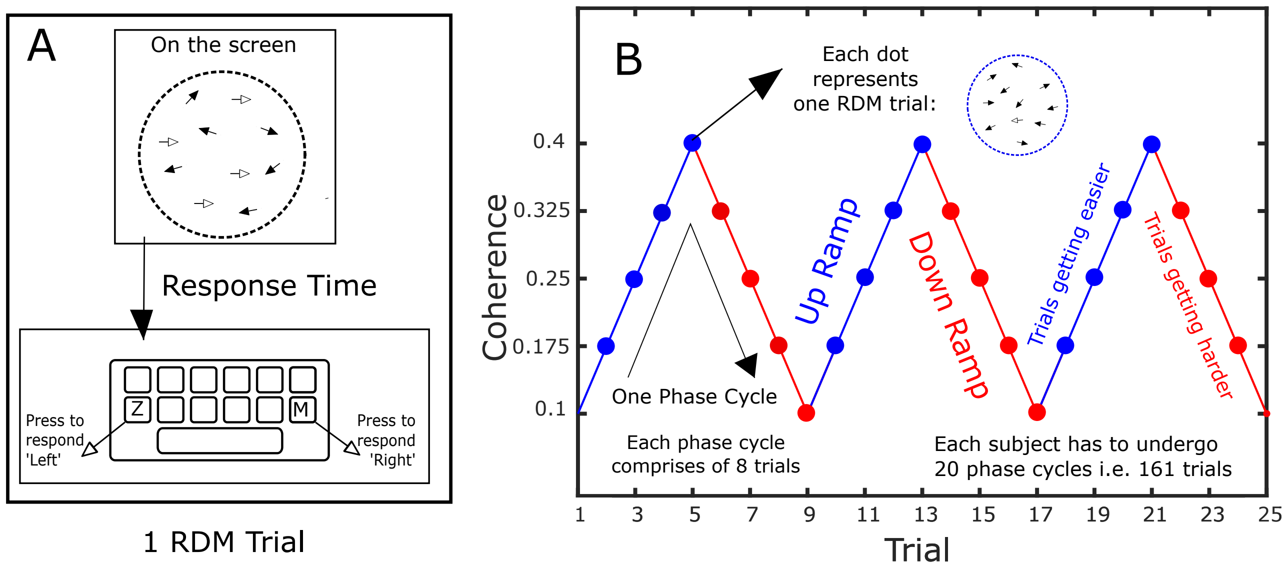

Figure 1: Schematic illustrating the experiment design. (A) On each experiment trial, participants saw a set of dots in Brownian motion, with a horizontal drift added to some fraction of the dots. Participants had to discriminate motion direction using key presses on a computer keyboard, and were incentivized to emphasize accuracy. (B) The sequence of trials each participant saw possessed a higher-order structure, with the difficulty of successive trials increasing and decreasing in a cyclic manner.

hierarchical model simulates individual RDM trials using a classic drift diffusion setup. The higher level of this model uses a reinforcement learning-inspired controller to set appropriate values of the evidence threshold for each trial.

To test this hypothesis, we designed a specific variation of the standard RDM task. In the standard task, trial difficulty is either blocked or randomized across trials. We instead designed a sequence of trial presentation that introduced a predictable trend in the coherence parameter across trials. If people are adaptively tracking the amount of effort they are having to expend on individual trials, we expect such inference to inform their effort allocation on upcoming trials. A hierarchical extension of the drift diffusion model, with a top-down controller setting the evidence threshold adaptively across trials, would potentially fit choice and RT data gathered from such an experiment design better than a simple DDM that assumes a fixed evidence threshold.

Experiment 1

RDM with higher-order structure

Participants saw a screen with moving dots designed according to the following algorithm. Random motion of the dots was provided by allowing Brownian motion in the vertical direction, i.e. all the dots drifted vertically about their mean position by a distance chosen from a normal distribution. Horizontal motion was either randomly selected from a bidirectional (left/right) uniform distribution (for incoherent dots) or from a unidirectional uniform distribution (for coherent dots). The selection of dots as coherent or incoherent was determined at the beginning of each experimental trial using Bernoulli trials controlled by the coherence parameter c.

As in all RDM discrimination experiments, the critical manipulation of task difficulty was governed entirely by the coherence parameter c. We selected a range of values of the coherence parameter by running a calibration pilot with 5 participants, performing 280 trials of the discrimination task under accuracy emphasis. We picked a range of coherence values that permitted 65% accuracy at the low end of the range and 95% accuracy at the high end of the range.

Participants had to indicate the direction of motion of the overall dot pattern on each trial, as illustrated in Figure 1A. They were allowed to take as much time as they wanted to respond to each trial, and as much time as they wanted to rest between the trials. Each correct response fetched points. The scoring system was such that a correct response fetches 10 points; the score of each correct response doubled on responding correctly to three successive trials, and reset to 10 points in case the streak was broken. We further encouraged accuracy emphasis by promising a reward to the highest scorer.

The specific higher-order structure introduced in our experiment was a cyclic shift across the 5 specific values of the coherence parameter used in the experiment $\{0.1, 0.175, 0.25, 0.325, 0.4\}$. For example, a participant starting the experiment with a trial with coherence 0.1 would next see a trial with coherence 0.175, then one with 0.25, up to the maximum coherence level of 0.4, beyond which the coherence would begin dropping down to 0.325, then to 0.25 etc. Each such phase cycle from one coherence value through the other 4 and back to the original takes 8 trials, given we use 5 unique coherence values. Participants completed 20 such phase cycles for a total of 161 trials per participant, as illustrated in Figure 1B.

Task

The task was administered via a web-based interface. Participants indicated responses with keyboard button presses, and were allowed to take as long as they liked before pressing the space bar to begin the next trial. Distance from the screen was not fixed, but the display size was selected such that the display was well within the foveal range (20 degrees visual angle) of normally sighted observers.

Participants

52 undergraduate and postgraduate students participated in the experiment for course credit (4 female, mean age 20.5 ± 1.57 SD). All participants had normal or normal-corrected vision. Since the experiment was conducted without personal supervision, some participants showed significant guessing behavior. Post-task, we excluded the data for 18 participants who had less than 85% accuracy on the highest coherence trials.

Results

We expected that an observer tracking the cross-trial variations in difficulty would end up tracking the repeated ramp-like movement of the coherence parameter, and take longer on an upcoming trial if the cross-trial coherence was trending downward (i.e. the trials were becoming more difficult) than if it was trending upward (i.e. the trials were becoming easier).

However, even observers insensitive to cross trial information are expected to show the same pattern globally in our data, because upcoming trials aren't just expected to be easier/harder on up/down ramps, they actually are easier/harder too. Therefore, the critical test for whether information from previous trials are affecting the current trial is to see whether trials of the same difficulty (coherence) level have longer RTs when they occur within a down coherence ramp (sequence of decreasing coherence, increasing trial difficulty) than when they occur within an up coherence ramp (sequence of increasing coherence, decreasing trial difficulty).

In Figure 2, we plot RTs stratified by coherence level of the immediate trial for the three intermediate coherence values in our experiment across all participants, separating out trials occurring during up and down ramps. For each of the three coherence levels, the difference between down ramp RTs and up ramp RTs is directionally in the predicted direction. While the data is noisy, the large sample size of our experiment (34 participants $\times$ 20 cycles = 680 data points per coherence level per ramp) allows us to assert statistical significance at Bonferroni-corrected $p < 0.01$ for two of the levels (0.175 and 0.325) in two-sample t-tests. The third comparison (for coherence level 0.25) does not yield a statistically significant difference in such a test.

These summative statistical results, while conceptually congruent with our hypothesis, are inadequate to draw strong inferences. The large standard deviations evident in Figure 2 reflect considerable inter-participant heterogeneity in response times. We therefore sought to test our hypothesis at a trial-by-trial level by developing a generative model of the information accumulation process involved in a 2AFC perceptual learning task that accommodates such adaptive control over the evidence threshold, and comparing its ability to explain our data vis-a-vis a fixed threshold evidence accumulation model.

A hierarchical drift diffusion model

The DDM, in its standard form is a Wiener diffusion process with drift,

$$dy = vdt + sdW, \tag{1}$$

where y is the diffusion state, v is the drift, s determines the amount of diffusion and dW represents the standard Wiener process. The model accumulates normally distributed pieces of evidence for either alternative until a bound on the cumulative evidence is crossed, and then emits the winning option as the choice.

We developed a hierarchical model of the choice process using a recently proposed Bayesian version of the DDM (Bitzer, Park, Blankenburg, & Kiebel, 2014). This takes the form of a sequential Bayesian update model that maps noisy stimuli observations to latent Gaussian representations

$$x_t \sim N(\mu_i, \sigma^2), \tag{2}$$

constructs a generative model of the likelihood of seeing certain latent feature values for each stimulus alternative,

$$p(x_t|A_i) = N(\hat{\mu}, \hat{\sigma}^2), \tag{3}$$

and updates beliefs about the correctness of a decision alternative given these noisy observations

$$p(A_i|X_{1:t}) = \frac{p(x_t|A_i)p(A_i|X_{1_t-1})}{\sum_{j=1}^{M} p(x_t|A_i)p(A_j|X_{1_t-1})}, \tag{4}$$

where x_t are noisy observations from stimuli belonging to category i, with true prototypical values μ_i and measurement noise σ, estimated prototypical values $\hat{\mu}$ and internal generative variability $\hat{\sigma}$, A_i as possible alternatives and M as the number of considered alternatives. Bitzer et al show that this intuitive ideal observer model is formally identical to the drift diffusion model given certain assumptions about the relationship between parameters of the model.

We augmented this model with a metalearner that estimates the expected sampling effort needed for upcoming trials based on effort allocation on previous trials. We assume a very simple model for this metalearner. It simply updates the effort estimate λ as

$$\lambda_t = \lambda_{t-1} + \gamma\Delta_t, \tag{5}$$

where

$$\Delta_t = z(RT_{t-1}) - z(RT_{t-2}), \tag{6}$$

γ is a free scaling parameter, $z(RT)$ is the normalized z-score of RT at time t with respect to the RT distribution and λ_t

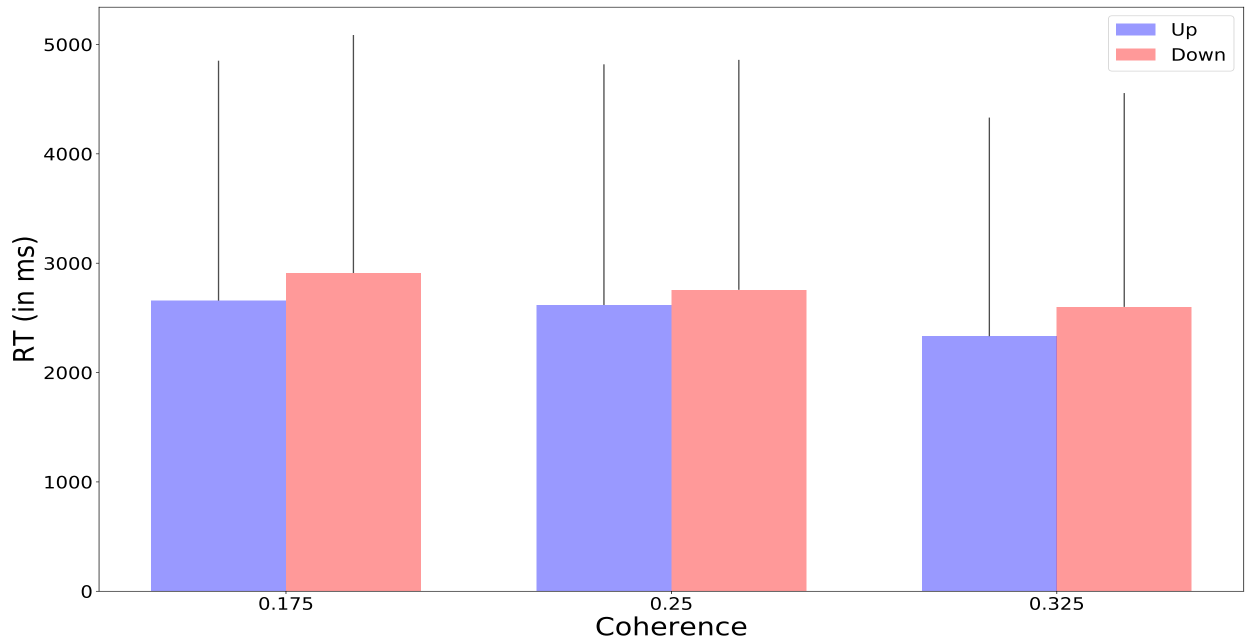

Figure 2: RT at intermediate coherence levels while the cyclic higher-order coherence trend is going up (blue bars) and down (red bars). Higher RTs seen for trials of equal difficulty, when coherence is trending downwards, i.e. trials are getting harder, is evidence for adaptive changes in the evidence threshold responsive to past trials. Error bars represent 1 S.D.

serves as the threshold for the DDM for the t^{th} trial. This model is not meant to be comprehensive. We have designed it purely to simulate our expectation of the role of predictable up and down changes in dot motion coherence on observer behavior. We expect that observers will be sensitized to these trends and extrapolate from them to set decision thresholds for upcoming trials. Increasing effort on recent trials should yield a larger threshold for the upcoming trial and vice versa. Normalization is used to induce a natural scale on the size of the change in the threshold; the RT distribution is admittedly non-Gaussian so this assumption could be further refined in future work. Also, to avoid over-fitting, we have used the global RT distribution to normalize the RTs, whereas a more realistic model may use sequential summary statistics within participants. Indeed, a filtering-based model might capture the basic intuition of the metalearner more elegantly, but we wished to compare the augmented model with a complicated baseline using only choice and RT data, necessitating parsimony in parameter extension. The version of the meta-learner we have proposed has just one additional free parameter beyond the baseline.

As in (Bitzer et al., 2014), we reduced the set of estimated parameters of the Bayesian model from seven to three by assuming equal amount of drift for both stimuli. In practice we did this by setting $\mu = \hat{\mu} = \pm 1$ for the 2AFC case. Parameter fitting for the 3 parameters to be estimated $\theta = \{\sigma, \hat{\sigma}, t_{nd}\}$ also followed the procedure outlined in (Bitzer et al., 2014). We defined the log likelihood of the data given all parameters as

$$\begin{aligned}\log p(\mathrm{Acc}, \mathrm{RT}|\theta) &= \log p(\mathrm{Acc}|\theta) + \log p(RT|\theta) \\ &\propto -w_{\mathrm{acc}} \left(\mathrm{Acc} - \mathrm{Acc}(\theta)\right)^2 \\ &\quad - \sum_{e=0}^{1} \sum_{i=1}^{7} w_{q_{e,i}} \left(q_{e,i} - q_{e,i}(\theta)\right)^2,\end{aligned}$$

where $q_{e,i}$ is the i^{th} of seven quantiles $(0.02, 0.05, 0.1, 0.3, 0.5, 0.7, 0.9)$ for either correct or error responses as indicated by e.

To evaluate the log likelihood function, again following the procedure in (Bitzer et al., 2014), we simulated our experiment with the Bayesian observer model for different parameter values, averaging the accuracy and RT quantiles obtained across 30 model runs per parameter tuple θ and setting the likelihood weights w to the inverse variance over these repetitions. We scaled each iteration in our simulation to 125 ms. However, we found the MCMC approach advised in (Bitzer et al., 2014) to be too slow (on the order of days) for fitting our hierarchical model that used different threshold values for each trial. Therefore, we used a two-stage grid-search of the parameter space (logarithmic exploration in one stage followed by linear refinement in the second) to optimize the negative log likelihood. In practice, we found that the grid search yields comparable mean parameter values for the baseline DDM model as the MCMC procedure implemented in (Bitzer et al., 2014).

Since we don't use MCMC to obtain a posterior distribution over the parameters, it is useful to average over multi-

Model	AIC	BIC
Simple DDM	16.5 (2.7)	36.3 (2.7)
Hierarchical DDM	**8.9 (1.5)**	**28.5 (1.5)**

Table 1: Model comparison. Standard deviation across 20 model runs are given in parentheses.

Block	1	2	3	4
Δ BIC	12.4	-3.0	-0.07	-0.78

Table 2: Model comparison across four sequential blocks of 40 trials each from all participants. Block 1 contains the first 40 trials from each participant, etc.

ple runs of the likelihood computation at the optimal parameter values to obtain representative likelihood values. After finding optimal parameters via grid search for both models, we calculate model likelihood as the average likelihood obtained from 20 runs of the model for the optimal parameter values. The results from our model comparison using these average likelihood values are presented in Table 1. ΔBIC measures difference between baseline DDM and hierarchical DDM BIC. Positive values support the hierarchical DDM, negative values support the baseline model. ΔBIC values with magnitudes smaller than 2 imply insignificant differences between models; values larger than 10 constitute very strong support for a model. Using this measure, the hierarchical DDM is clearly preferable to the simple DDM, across data from all 34 participants (ΔBIC = 7.8 from Table 1).

We additionally ran a block-wise analysis, dividing each participant's trials into 4 sequential blocks of 40 trials and calculating model complexity statistics on the likelihoods emitted by the model for the best fitting parameter values of the overall model ($\sigma = 11, \hat{\sigma} = 8, t_{nd} = 250ms$). We anticipated that any evidence of gradual adoption or relinquishment of threshold metareasoning would show up in the relative model complexity tracked across these four blocks.

As the results in Table 2 show, the hierarchical model is heavily preferred over the simpler model during the first quarter of trials, measured across all participants. For later trials, both models are evenly matched, with the simpler model slightly preferred.

Experiment 2

The block-wise analysis of our data revealed an interesting property: participants behaved as if they were tracking higher-order structure at the beginning of the experiment, but appeared to switch away to behaving more like conventional DDM decision-makers later on. We thought this was because participants tried to use the higher-order structure between trials, but then shifted away from it, since doing so does not offer any material advantage. To falsify such an explanation, we conducted a followup experiment where tracking the higher-order structure would confer a material advantage.

RDM with helpful higher-order structure

This experiment design was equivalent to the first one, with an alteration only in score-keeping. Recall that the first experiment used scoring with a multiplicative boost for maintaining accuracy streaks. Score per correct response would double for every three consecutive correct responses, and reset to the default value on each error.

For this second experiment, we transformed the incentive system into an optional 'auto boost' mechanism, such that accurately responding on three consecutive trials would fill up a booster bar, and optionally selecting the booster bar would allow a participant to *buy* out any one trial of their choice. The bought out trial would be assumed to have been answered correctly, would not be displayed on screen, and would not count towards win streak counts.

Participants

31 undergraduate students (age = 20.6 ± 1.3 years, 8F) volunteered to participate in this experiment; none having participated in the first one. These participants performed the experiment under supervision using the same web-based interface as before. All other experimental protocols were identical to the first experiment. Post-task, we excluded the data for 5 participants who had less than 85% accuracy on the highest coherence trials.

Results

The primary difference between the two experimental tasks was that tracking higher-order structure could not help participants in the first one, but potentially could in the second one. Specifically, in this second experiment participants could hold on to filled up booster bars for what they *predicted* to be the hardest trials, considerably reducing their overall error and effort.

Block	1	2	3	4
Δ BIC	9.7	8.4	8.9	3.2

Table 3: Model comparison across four sequential blocks of 40 trials each (counting bought out trials) from all participants for Experiment 2. Block 1 contains the first 40 trials from each participant, etc.

A direct measure of whether participants did deploy such a strategy in this task is the relative distribution of trials on which participants used boosts across coherence levels. Random use of boosts would indicate no use of the optimal strategy outlined above, whereas exclusive concentration of boost use in the lowest coherence level would indicate perfect adherence to this strategy.

Empirically, we found that 45.5% of all boosts were used for the lowest coherence trials (chance 12.5%, $p < 0.005$), and 78.6% of all boosts were concentrated within the two lowest coherence levels (chance 37.5%), suggesting strong utilization of the optimal strategy.

As a secondary measure of strategy adherence, we fit both classic and hierarchical DDM to these participants' data following the same methodology as for the first experiment. Boosted trials were assigned correct responses and subject-specific mean RT for the corresponding coherence level during the simulations. For lack of space, we only present the difference in BIC values betwe en hierarchical and classic DDM fits for this data. Contrasting these ΔBIC results tabulated in Table 3 with those from the original experiment (see Table 2) strongly suggest that the manipulation of the incentive system does affect participants' behavior on the task in the predicted direction.

Discussion

With increases in computational power and experimental methods, behavioral researchers are increasingly able to track behavior with greater granularity, which makes hypotheses about intermediate computations underlying behavior tractable to investigation. This development is making the study of algorithmic aspects of the decision-making process increasingly more feasible. Not only does such research characterize biological observers' decision-making *processes*, it also provides constraints on the nature of the cost functions that computational theorists can reasonably set up for decision-making agents.

As part of this paradigm shift, recent work has begun to question the classic drift diffusion model's assumption of a fixed evidence threshold in recent years, basing these arguments on the temporal opportunity costs of continued sampling (Thura et al., 2012). In this work, we asked the same question from a different standpoint. We asked whether observers might be sensitive to higher-order statistics in decision-making tasks and adaptively adjust evidence thresholds on upcoming trials to use them efficiently.

To see if this can happen, we created a novel variation of the classic random dot motion discrimination task, introducing an up-and-down ramp in difficulty across trials (Gold & Shadlen, 2000). We predicted that observers would be sensitive to this variation. We designed an extension of the drift diffusion model that incorporates metacognitive adaptation of the evidence threshold based on the trend of difficulty of recent trials, and found that it offers a better explanation of participants' behavior in our experiment than a simple drift diffusion model. A followup experiment further demonstrated that participants' tracking of higher-order structure in this task was intentional - they shifted away from tracking when it offered no advantage and continued tracking when it did.

Our results support a shift in interpretation of the evidence threshold from its SPRT-driven association with *accuracy*, towards a more general view of it as an effort parameter influenced by a variety of information sources. Such an interpretation also makes it easier to generate normative accounts of decisions from memory using DDM-like models, building upon its descriptive success in modeling retrieval success and RT distributions in this domain (Krajbich & Rangel, 2011). Unlike in perceptual decisions where the evidence presentation rates are fixed, and decisions receive immediate feedback, decisions from memory are made with evidence streams of unknown provenance and without feedback. The empirical success of DDM in explaining data from such experiments warrants a broader interpretation of the normative principles of the framework, along lines proposed in this work.

References

Bitzer, S., Park, H., Blankenburg, F., & Kiebel, S. J. (2014). Perceptual decision making: drift-diffusion model is equivalent to a bayesian model. *Frontiers in human neuroscience*, *8*, 102.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, *113*(4), 700.

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.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review*, *100*(3), 432.

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

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.

Glaze, C. M., Kable, J. W., & Gold, J. I. (2015). Normative evidence accumulation in unpredictable environments. *Elife*, *4*.

Gold, J. I., & Shadlen, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, *404*(6776), 390.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural computation*, *20*(4), 873–922.

Thura, D., Beauregard-Racine, J., Fradet, C.-W., & Cisek, P. (2012). Decision making by urgency gating: theory and experimental support. *Journal of neurophysiology*, *108*(11), 2912–2930.

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

Wald, A., & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, 326–339.