Title
Modeling sampling duration in decisions from experience

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors
Srivastava, Nisheeth
Muller-Trede, Johannes
Schrater, Paul
et al.

Publication Date
2016

Peer reviewed

Modeling sampling duration in decisions from experience

Nisheeth Srivastava, Johannes Müller-Trede
UC San Diego, 9500 Gilman Drive
La Jolla, CA 92093 USA

Paul Schrater
University of Minnesota
Minneapolis, MN 55455 USA

Edward Vul
UC San Diego, 9500 Gilman Drive
La Jolla, CA 92093 USA

Abstract

Cognitive models of choice almost universally implicate sequential evidence accumulation as a fundamental element of the mechanism by which preferences are formed. When to stop evidence accumulation is an important question that such models do not currently try to answer. We present the first cognitive model that accurately predicts stopping decisions in individual economic decisions-from-experience trials, using an online learning model. Analysis of stopping decisions across three different datasets reveals three useful predictors of sampling duration - relative evidence strength, how long it takes participants to see all rewards, and a novel indicator of convergence of an underlying learning process, which we call predictive *volatility*. We quantify the relative strengths of these factors in predicting observers' stopping points, finding that predictive volatility consistently dominates relative evidence strength in stopping decisions.

Keywords: response time; decision-making; evidence accumulation; sequential sampling; decisions from experience

Introduction

In orienting decision research from analyzing static economic choices towards more dynamic decisions akin to those people face in everyday life, the decisions-from-experience (DFE) paradigm presents an important step forward (Hertwig, Barron, Weber, & Erev, 2004). The DFE paradigm may be thought of as a modification of the classic certainty equivalence method commonly used procedure to elicit utility functions for money (see von Winterfeldt & Edwards, 1993). Participants in typical certainty equivalence experiments choose between a risky option that pays H with a probability p and L with probability $1-p$ and a safe option that always pays M, where $H > M > L$. They learn about the two options' payouts and the associated probabilities from explicit descriptions.

Decisions from experience modify this protocol: Participants are not shown payouts or probabilities, and must learn about them experientially. Several interesting observations emerge from research on DFE. Subjects sample more variable options and options with higher stakes for longer, for example (Lejarraga, Hertwig, & Gonzalez, 2012). They also appear to underestimate the probability of rare events (Hertwig et al., 2004), though much less so with increasing experience (Zhang & Maloney, 2012).

In the particular variety of DFE we consider throughout this paper, typically known as the sampling paradigm, participants are permitted to sample each of two options without consequence as long as they like, before finally committing to a binding choice. This protocol is particularly interesting since it closely mirrors the flow of information in many everyday settings–learn from the environment *ad libitum*, then make a choice. Importantly, such choices are actually composed of two decisions: a latent decision to terminate learning, and an overt decision to choose the risky or the safe option, based on the learned information.

Efforts to model the overt decision about which option to choose have been relatively successful (Erev et al., 2010). Little attention has been paid, however, to modeling the earlier latent decision about how long to sample information (or when to stop learning). Research on DFE has used simple statistical approaches as place-holders, assuming an underlying probability distribution over sampling lengths and fitting this distribution to the empirical distribution of sampling lengths observed in the data (Gonzalez & Dutt, 2011). Recently, Markant et al. (2015) proposed a model that jointly predicts choices and sampling length distribution. However, since their model uses *known* lottery stakes, it cannot speak to the effects of the specific characteristics of the actual, realized samples that comprise individual learning experiences in the DFE paradigm.

In this paper, we investigated variables that, on theoretical grounds, are expected to predict sampling lengths in DFE, without assuming *a priori* knowledge of lottery stakes and probabilities. Our analyses of three different datasets revealed the influence of two important variables on sampling duration. One is the difference in the options' expected values during sampling. The second is a measure of outcome uncertainty we call predictive *volatility*, which tracks abrupt changes in the magnitude of prediction error an observer experiences while learning about a DFE decision. We further developed a computational model of sequential sampling for DFE that uses these predictors to make accurate sampling length predictions for individual DFE trials.

Predictors of sampling duration

While largely neglected in the context of DFE, sampling durations have played a key role in research on perceptual decision tasks. Several of the leading theories in this domain utilize drift-diffusion models (Forstmann, Ratcliff, & Wagenmakers, 2016). In such models, the principal variable of interest tends to be the relative evidence strength in favor of either outcome (Bitzer, Park, Blankenburg, & Kiebel, 2014), measured in log posterior odds. The larger the log odds favoring a particular option, the more decisive–and quicker–the evidence accumulation in its favor. In DFE, relative evidence strength corresponds to the difference between the imputed value of the two options. Theories of perceptual decisions applied to DFE thus suggest that smaller differences in the options' imputed values should be associated with longer sampling durations.

Formally, we track the expected value difference (EVD) between the two gambles, measured at every sample for each

DFE trial,

$$EVD = pH + (1-p)L - M, \quad (1)$$

where $p = \frac{|H|}{|H|+|L|}$, and $|\cdot|$ is the number of times an outcome has occurred in the sequence up to the time at which the measurement is taken. In the following, we refer to this quantity as the EVD predictor.

Information theory, we suggest, points to a second, complementary factor that might influence sampling durations. From an information-theoretic perspective, the DFE observers' goal is to efficiently learn the reward rate of the gamble(s) they are sampling. The decision to terminate information gathering may then be seen as an agent's rational response to a learning procedure that has saturated. In practice, observers, while trying to learn useful models of their environment, have access to the prediction errors in such models. Unexpected increases in prediction error magnitude (as illustrated in Figure 1) signal the presence of unlearned environmental dynamics, stimulating rational observers to sample more.

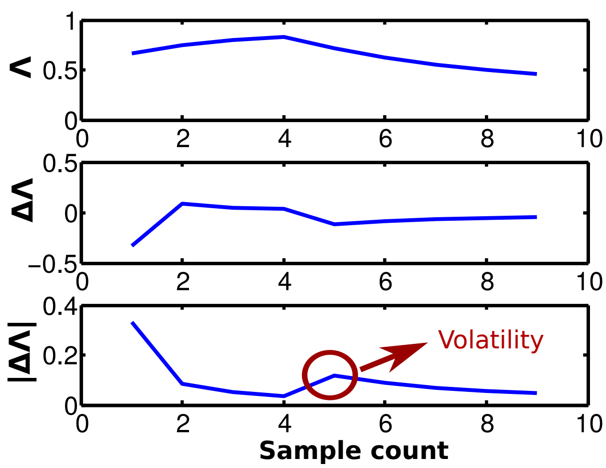

Figure 1: An intuitive view of predictive volatility. When the prediction error in a sequential learning process increases abruptly, it is reasonable to infer that the process has yet to converge. This indicator of the need to keep learning is what we call predictive volatility.

To understand the learning process, we can track the evolution of the learned parameter in a simple parametric observer model over individual sampling sequences. If a learning procedure is efficient, the prediction error is expected to show an asymptotic gradual decrease, reflecting that estimation is increasingly precise as more data are sampled. Critically, we assume that human observers are intuitive statisticians in this particular sense – they are implicitly aware that when learning is efficient, prediction error gradually declines. But in individual learning sequences, this decline is not always monotonic. We call deviations from the prediction error's expected trajectory episodes of *predictive volatility*. Observers who are sensitive to the expected trajectory of prediction error may rationally treat such episodes of volatility as evidence that learning has not yet completed, and respond by sampling for longer.

Formally, we model learning in DFE as a statistically efficient observer sequentially updating estimates of the mean parameter Λ of a Poisson distribution tracking the frequency with which the high outcome of the risky option occurs in the sampling sequence. With every new sample, the parameter estimate shifts by some quantity $\Delta\Lambda$. In absolute values, this quantity is mathematically–and, we argue, psychologically–expected to decay over time, so that $|\Delta\Lambda_t| < |\Delta\Lambda_{t-1}|$. Deviations from this trend constitute predictive *volatility*. Such deviations suggest to an observer that some aspects of the task may remain unlearned, and therefore justify continued sampling. In the case of human observers, the deviations must be sufficiently large to be noticeable, which suggests that volatility is perhaps best thought of as a binary variable which is either present or absent. This led us to formally operationalize volatility as

$$v(t) = 1 \text{ iff } |\Delta\Lambda_t| > \kappa \times |\Delta\Lambda_{t-1}|, \text{ and 0 otherwise.} \quad (2)$$

The constant $\kappa > 1$ determines the smallest noticeable deviation. All our analyses use $\kappa = 2$; substantially larger values would degrade the information present in this signal (since such large fluctuations are statistically infeasible in DFE reward rate estimation, where the set of possible outcomes is very limited); substantially smaller values would add noise to the signal, in the form of volatility false positives. Across an entire sampling sequence, the cumulative effect of such episodes is measured by the trial volatility load

$$V = \sum_t v(t), \quad (3)$$

and is referred to, in our following analysis, as the *volatility* predictor.

Finally, observers who know (i.e., have learned) that the standard DFE paradigm pits a safe option against a risky option (see above) may want to see all three reward outcomes at least once before terminating sampling. Depending on how skewed the risky option's odds are, this can take a relatively long time. The number of samples it takes to see all three reward outcomes at least once thus also contains valuable information about the length of the sampling sequence. It enters our predictive model in the form of a *counting* predictor, counting the number of samples it took a participant to see all three options at least once.

Results

We present two sets of results. We first demonstrate, using a proportional hazards regression analysis, that both the magnitude of difference in expected value and the amount of volatility seen in the sampling sequence influence sampling durations as predicted by our theory. Model selection reveals that volatility plays a more influential role in this process.

These results, however, are calculated using *post hoc* predictors that an actual observer would not have access to in real-time. Our second set of results uses sequential counterparts to these predictors to develop a sequential model that simulates the trajectory of the stopping probability of any DFE trial, sensitive to the influence of both differences in expected value and episodes of volatility experienced in real-time.

Data

We procured data from two sources: the decisions-by-sampling condition from the Technion Prediction Tournament, which involved two sets (an estimation and a competition set) of 40 participants each solving 30 such problems, and a sample of 37 participants solving 19 different DFE problems we collected in the DFE condition of a different experiment (Experiment 2 in Lejarraga & Müller-Trede, 2016). We refer to the Technion estimation dataset as **TE**, the Technion competition dataset as **TC**, and our own sample as **LM**.

Experimental protocols were largely identical across the datasets.[1] Participants could sample both options in each lottery pair as often as they liked, and subsequently committed to one final draw that would correspond to their actual payout. All participants were compensated via a random incentive scheme, and earned real money corresponding to their payout in one randomly selected choice problem. We note that participants in LM revisited each choice problem in a group setting between individual trials (for details, see Lejarraga & Müller-Trede, 2016), whereas participants in TE and TC did not.

Volatility matters more

What can the expected value difference between options tell us about the decision to stop sampling? Recall that small EVDs may trigger further sampling, whereas large EVDs supply a reason to choose one option over the other (and thus to terminate sampling). The average EVD predictor, defined as $\frac{1}{T}\sum_t^T |EVD_t|$ for sequences of duration T, should then be negatively correlated with sampling lengths. If observers use the weight of economic evidence to decide when to terminate information search, greater average magnitudes of the expected value difference should correspond to earlier sampling termination and vice versa.

We tested this hypothesis by running a Cox proportional hazards regression, assessing the direction and magnitude of effect the average expected value difference measured during a sampling sequence has on the hazard rate across all trials per subject. The top panels in Figure 2(A) show that, as expected, EVD consistently increases hazard rates across participants in all three datasets.

We ran a similar proportional hazard regression using volatility load as a predictor of sampling sequence lengths. As the bottom row in Figure 2(A) demonstrates, volatility load consistently retards hazard rates in participants across all three datasets tested. Panel B in Figure 2 compares the regression coefficients obtained when we use both predictors–EVD and volatility–normalized and combining all participants in each dataset. In all three cases, the normalized predictors are uncorrelated ($r < 0.05, p > 0.25$), so the regression coefficients (β-weights) are informative about the relative importance of the two predictors (Nathans, Oswald, & Nimon, 2012). Volatility consistently dominates EVD as a predictor across all three datasets.

[1] Participants in the LM experiment "chose" by alloting fractions of an allocation budget to either option; TE and TC used binary choices. Importantly for our purposes, the sampling procedure was the same in all three studies.

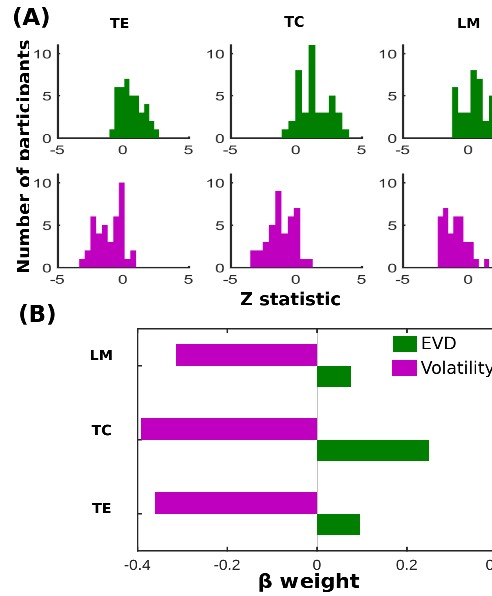

Figure 2: **(A)** Histograms of subject-wise Z-statistics obtained by Cox regression of predictors against sampling sequence lengths. Negative Z-statistics indicate that the predictor reduces the hazard rate, yielding longer sampling durations than baseline expectations. **(B)** Coefficients from Cox multiple regressions using normalized EVD and volatility predictors for all three datasets.

Table 1: Model selection using BIC. Lower values are better within individual datasets. ΔBIC from best model reported in brackets.

	TE	**TC**	**LM**
EVD	14312 (+165)	14643 (+177)	7090 (+63)
Volatility	14157(+10)	14529 (+63)	7031(+4)
Both	14147	14466	7027

A similar conclusion can be drawn by computing the Bayesian Information Criterion (BIC) for regressions using the two predictors indvidually, and then together. Not only is the full model preferable by BIC, corresponding ΔBIC values show that the volatility-alone model is considerably closer to the full model than the EVD-alone model, suggesting that it is a more powerful predictor. Together, these analyses demonstrate (i) that EVD and volatility are independent sources of information for predicting sampling duration and (ii) that volatility is a more informative predictor for sampling durations than EVD.

DFE sampling durations are *post hoc* predictable

How well can our account predict actual sampling durations in the DFE paradigm? A simple additive model combining these two predictors,

$$\text{duration} = \text{volatility} + \beta\, \text{EVD}$$

yields correlations $r = \{0.56, 0.53, 0.45\}$ with human sampling lengths for the TE, TC and LM datasets respectively, suggesting that these predictors can explain around 20-30% of the variance in sampling durations for human observers in DFE.[2] For lack of competitors–to the best of our knowledge, ours is the first model for predicting DFE sampling durations at the individual trial level–we cannot assess this performance comparatively.

To further improve predictive ability, we can add the counting predictor to the model. Its incorporation yields an augmented linear model

$$\text{duration} = \text{count} + \alpha\, \text{volatility} + \beta\, \text{EVD}$$

which improves the correlations with human data to $r = \{0.56, 0.69, 0.71\}$ for the TE, TC, and LM datasets respectively.[3]

Note that adding the counting predictor did not improve the data-model correlation for the TE dataset. This is because participants in this dataset strongly violated the expectation that participants would want see all three outcomes at least once. Of the 1200 total trials in this dataset (40 participants $\times$ 30 problems), as many as 742 trials (62%) were terminated without having seen all three outcomes, including 192 (16%) that were terminated after drawing just two samples altogether. For comparison, participants in the TC and LM datasets terminated 41% and 14% of all trials before seeing all three possible outcomes, respectively. We suspect that the higher sampling effort in the LM dataset may reflect additional intrinsic motivation participants in that study derived from the repeated social interactions between choice problems.

In the other two datasets, adding the counting predictor boosts the data-model correlation to ≈ 0.7, so the augmented model explains around 50% of the variance in those data. The large improvement in predictive ability somewhat overstates the predictor's true explanatory value, however. To see why, note that the sample count at which all three options have been seen once, by definition, cannot exceed the overall sampling sequence length. The counting predictor is thus upper-bounded by the independent variable it is used to predict. The resulting absence of counting predictor values greater than the actual sampling length inflates its correlation with the independent variable. Hence, while the counting predictor *prima facie* adds substantial predictive value to our model, it does so for reasons that need not be theoretically insightful.

[2] Best fit $\beta = 0, -0.1, -0.4$.

[3] For best fit values of $\alpha = \{3.8, 5.9, 2.6\}$ and $\beta = \{-0.2, -0.4, -0.9\}$ respectively.

Real-time stopping point prediction

While we show that sampling lengths in DFE are substantially predictable *post hoc* using objectively observable predictors, not all these predictors are available to decision-makers at the time of making their decisions. Neither the average EVD magnitude nor the cumulative volatility load across the complete sequence is available to an observer who is currently sampling. In the following analyses, we used elements of our predictors that are available to observers in real-time, and test how well they predict eventual stopping decisions.

To do so, we use computational models that perform the same sampling task as the observer, stepping through each trial sample by sample, predicting sequence lengths indirectly by estimating stopping probabilities λ_t at each sample. To make this analysis feasible, we make the simplifying assumption that there are no individual differences across participants within datasets. Doing so yields multiple data points for each DFE problem, instead of just one per problem-by-participant pair. This allows us to construct a stopping point distribution for each problem.

We then use these stopping point distributions to fit a basic piece-wise constant hazard model that assumes the stopping probability increases linearly with the sampling count t, i.e.,

$$\lambda_{t+1} - \lambda_t = \delta, \quad (4)$$

and fit $\{\lambda_0, \delta\}$ for each unique problem using a grid search, maximizing the statistical indistinguishability (measured using a two-sided T test; typical values $p > 0.95$) between the empirical stopping point distributions and the model's predicted stopping point distribution, averaged across multiple simulations (N=1000). We kept these $\{\lambda_0, \delta\}$ values fixed in the subsequent models we describe below, fitting only the additional parameters.

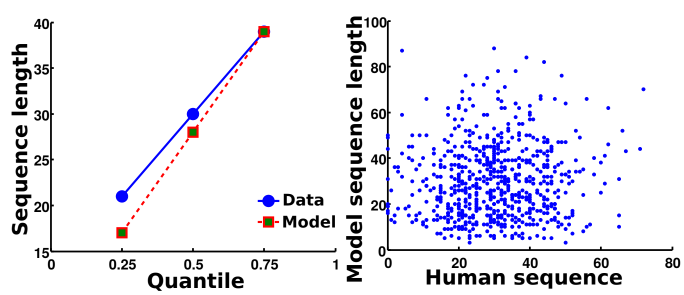

Figure 3: While simple statistical models of sampling lengths can reproduce population-level statistics *(left)*, they *(right)* fail to predict sampling sequence lengths for individual trials. Results shown for LM dataset.

This simple model theoretically and empirically resembles previous sampling length models proposed in the literature, which assess model fit by testing whether it produces the same statistical distribution of sampling lengths as the underlying data (Gonzalez & Dutt, 2011; Markant, Pleskac, Diederich, Pachur, & Hertwig, 2015). Figure 3 illustrates that our baseline model closely approximates the empirical distribution of sampling lengths, much like existing models (compare left panel with Figure 1 in Markant et al., 2015). It is a poor predictor of sampling lengths at the individual trial level ($r = 0.03$, right panel), however, which illustrates a basic limitation of this modeling approach.

Table 2: Best fit correlations of sampling durations predicted by sequential models with human data in all three datasets.

Models	**TE**	**TC**	**LM**
Baseline	-0.04	0.02	0.03
Baseline + Vol	0.20	0.11	0.19
Baseline + EVD	0.15	0.11	0.06
Baseline + EVD + Vol	0.26	0.18	0.21
Baseline + EVD + Vol + Counting	**0.44**	**0.38**	**0.41**

Next, we added a volatility predictor to the model. In particular, we assumed that the baseline stopping probability would increase by Δ every time the observer encounters volatility in the sampling sequence,

$$\lambda_{t+1} - \lambda_t = \delta + v(t)\Delta. \quad (5)$$

Here, volatility refers to single episodes of volatility within a sampling sequence as defined in Equation 2, not the cumulative quantity ("load") measured across the entire sequence. If volatility retards the termination probability as predicted, negative values of Δ will yield greater correlations of the model's sample sequences with human data. As Figure 4 illustrates, observer models that reduce stopping probability when encountering volatility ($\Delta < 0$) indeed provide the best fit to the data in all three datasets. Our hypothesis about the influence of volatility thus finds clear support in the data.

We ran a similar analysis to measure the sample-by-sample impact of the EVD predictor. We assumed that incoming signals of greater relative evidence strength would affect the stopping probability following a logistic relationship, with larger values having a disproportionally larger effect. Thus,

$$\log \frac{\lambda_t}{1-\lambda_t} = \log \frac{\lambda'_t}{1-\lambda'_t} + k \log |d_t + 1|, \quad (6)$$

where λ'_t is obtained from Equation 4, k is fitted to the data to maximize the model-data correlation, and d_t is the EVD calculated using the sequence up to the t^{th} sample. Table 2 provides a summary of the results.

To combine the influence of volatility and EVD, we revisited Equation 6 with λ' calculated via Equation 5, and with $\{\Delta, k\}$ as free parameters. The best overall model fit yielded weakly positive correlations across the three datasets.[4]

[4]Best fit parameter values for all three datasets: $\Delta = \{-0.20, -0.15, -0.125\}, k = \{0.02, 0.02, 0.02\}$.

Finally, when we incorporated the counting predictor into our model in the form of a real-time decision threshold–if all options seen at sample t, terminate with probability λ_t, otherwise, terminate with probability λ_0–the correlations improved substantially, to $\{0.44, 0.38, 0.41\}$, for the same parameter values as in the previous model. Unlike in the aggregate analysis, the effect of the counting predictor in sample-by-sample data does not necessarily suffer from the problem of artifactual inflation of correlation. These numbers thus present a fair picture of the predictability of sampling durations using only real-time information. Empirical estimates of the test-retest reliability of participants' sampling lengths would be required to assess how well our model's performance matches the best possible performance.

Discussion

This paper develops theory and algorithms to predict sampling durations in economic decisions from experience in which observers freely sample options before committing to a binding choice. We argue and then empirically demonstrate that a combination of evidence strength, predictive volatility, and simply tracking how long it takes participants to observe the entire reward structure of the particular DFE problem they are solving goes a considerable way in explaining individual decisions to stop sampling. Previous attempts to modeling information search in DFE succesfully reproduced distributions of sampling duration but were effectively random in trial-level predictions. Our account matches these previous attempts in distribution-level performance and surpasses them in making reasonably accurate trial level predictions. Finally, since it yields direct stopping point predictions, our model could easily be combined with choice models that respect the epistemic limitations of sampling-based DFE such as primed sampling or natural means (Erev et al., 2010) to make joint predictions of choice and sampling duration.

Of the three predictors we examine, one–the number of samples required to observe all possible outcomes–is specific to DFE. Its predictive power is substantial, which is interesting because it suggests that participants in DFE experiments both discern an aspect of the task structure not explicitly described to them (i.e., that each choice is between a safe option and a risky option with exactly two possible outcomes), and adaptively react to it. The other two predictors–evidence strength and volatility–are substantially more general and may thus be used to predict sampling duration in other experimental modalities. For example, Juni et al. have demonstrated that observers sample for longer when encountering noisier stimuli in a visuomotor estimation task (Juni, Gureckis, & Maloney, 2011). In this modality, greater stimulus noise corresponds directly to lower evidence strength, and as in our case, lower evidence strength is associated with longer sampling.

Our discovery of the significant influence of predictive *volatility* on sampling duration warrants further investigation. Predictive volatility is a highly accessible information signal

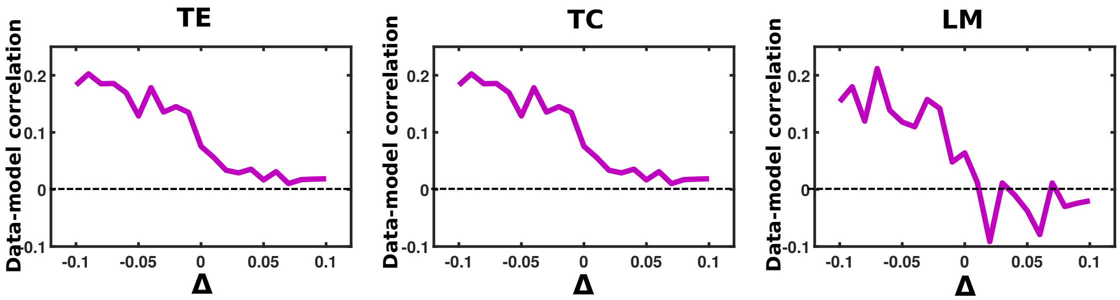

Figure 4: Model-data correlations for observer models fitted using different increments to stopping probability when encountering volatility.

that observers could draw on in a variety of real-world decisions from both experience and memory. To date, researchers have modeled response durations as arising from either evidence accumulation rising to a fixed threshold (Forstmann et al., 2016), or from a time-sensitive threshold collapsing to meet accumulating evidence (Thura, Beauregard-Racine, Fradet, & Cisek, 2012). Our results suggest that thresholds need not stay fixed or fall over time. Instead, they could rise and fall adaptively within trials in response to sequence-dependent predictive volatility. Volatility's representational generality, alongside our demonstration of its consistent and considerable impact on DFE stopping point decisions, invites further exploration in other experimental designs.

The role of predictive volatility in determining when to terminate sampling could also streamline the functional interpretation of cortico-striatal dopaminergic activity in decision-making (Schultz, Dayan, & Montague, 1997). Dopamine has been experimentally associated with encoding both reward and prediction error. The latter association appears to be more robust, however, in that it is congruent with a larger literature on the role of prediction error in multiple motor, cognitive and perceptual functions (Friston et al., 2012). Our account provides a rationale for why dopaminergic activity could be temporally correlated with the choice process without actually encoding reward: It may instead play a critical role in the latent decision to make a choice, and to terminate information search.

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*.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., … Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making*, *23*(1), 15–47.

Forstmann, B., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual review of psychology*, *67*, 641–666.

Friston, K., Shiner, T., FitzGerald, T., Galea, J., Adams, R., Sporns, O., et al. (2012). Dopamine, affordance and active inference. *PLoS Comput Biol*, *8*(1), e1002327.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review*, *118*(4), 523.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*(8), 534–539.

Juni, M. Z., Gureckis, T. M., & Maloney, L. T. (2011). Dont stop til you get enough: adaptive information sampling in a visuomotor estimation task. In *33rd Annual conference of the cognitive science society* (pp. 2854–2859).

Lejarraga, T., Hertwig, R., & Gonzalez, C. (2012). How choice ecology influences search in decisions from experience. *Cognition*, *124*(3), 334–342.

Lejarraga, T., & Müller-Trede, J. (2016). When experience meets description: How dyads integrate experiential and descriptive information in risky decisions. *Management Science (in press)*.

Markant, D., Pleskac, T. J., Diederich, A., Pachur, T., & Hertwig, R. (2015). Modeling choice and search in decisions from experience: A sequential sampling approach. In *37th annual conference of the cognitive science society.*

Nathans, L. L., Oswald, F. L., & Nimon, K. (2012). Interpreting multiple linear regression: A guidebook of variable importance. *Practical Assessment, Research & Evaluation*, *17*(9), 1–19.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.

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.

Von Winterfeldt, D., & Edwards, W. (1993). *Decision analysis and behavioral research.* Cambridge, MA (USA) Cambridge Univ. Press.

Zhang, H., & Maloney, L. T. (2012). Ubiquitous log odds: a common representation of probability and frequency distortion in perception, action, and cognition. *Frontiers in Neuroscience*, *6*.