Choice anticipation as gated accumulation of sensory expectations

Brandon Caie[1], Dominik Endres[3], Aarlenne Khan[2], and Gunnar Blohm[1]

[1]Queen's University
[2]Universite de Montreal
[3]Philipps-University Marburg

Abstract

Expectations are combined with sensory information when making choices. Some models of the choice process have conceptualized expectations as trial by trial updates to baseline evidence in an accumulator framework. These models have been successful in explaining the influence of choice history across trials on reaction times and choice probabilities, however they do not account for variability in the delay interval within trials. Here, we derive a gated accumulator that models the onset of evidence accumulation as a combination of delayed sensory information and an expectation of sensory timing that changes within trials. To test how the delay intervals interact with trial by trial expectations, we designed a free choice saccade task where participants directed eye movements, as fast as possible, to either of two targets that appeared with variable delays and asynchronies. Despite being instructed to not to anticipate, participants responded prior to target onset on a subset of trials. We reasoned that anticipatory responses may reflect a trade-off between gating the onset of evidence accumulation prior to target onset and releasing this gate as target appearance became more likely. Using a choice history analysis, we found that anticipatory responses were more likely following repeated choices, despite task randomization, suggesting that the balance between anticipatory and sensory responses was driven by an expectation of sensory timing. By fitting the gated accumulator model to reaction time data, we determined that within-trial fluctuations in baseline evidence occurring prior to the onset of the target could explain the joint increase of anticipatory responses and faster sensory-guided responses as the delay period increased. Thus, we conclude that an expectation of sensory timing can shape how the accumulation of choice evidence is gated, allowing us to balance the costs of anticipation with lowering the amount of evidence required to trigger movement.

Introduction

We infer patterns, even when there are none. Sometimes the causal relationships we infer do in fact exist, so we can reduce the cost of continuously processing new information by relying on expectations formed from past experience (Griffiths and Tenenbaum 2006; Gershman, Horvitz, and Tenenbaum 2015; Richards and Frankland 2017). However, causal relationships are often inferred even when events are independent (Jarvik 1951; LaPlace 1951; Nickerson 2002). When the statistics of tasks are randomized, we often fail to take into account the global structure of an environment, instead incorporating recent history into our decision-making process. While different variants of this cognitive bias are well known, (Nash 1950; Bakan 1960; Gilovich, Vallone, and Tversky 1985; Bar-Hillel and Wagenaar 1991; Mookherjee and Sopher 1994; Abrahamyan et al. 2016), less is known about how these biases interact with sensory processing as we are planning and executing movements.

Reaction times between a stimulus and a movement are thought to be determined by the time it takes to integrate sensory information from a baseline to a threshold (Stone 1960; Busemeyer and Townsend 1993; Dzhafarov 1993; Carpenter and Williams 1995; Usher and McClelland 2001; Ratcliff and Smith 2004; Bogacz et al. 2006; Wong and Wang 2006; Standage, Wang, and Blohm 2014). Shifts in the baseline level of choice evidence has traditionally accounted for the influence of choice bias on the shape of reaction time distributions (Carpenter and Williams 1995; Bogacz et al. 2006; Ratcliff and McKoon 2008; Mulder et al. 2012; Glaze, Kable, and Gold 2015), although a direct and unique link has been questioned (Urai et al. 2019; Simen 2012; Fan, Gold, and Ding 2020). Reinforcement-learning models of reaction times have since been applied to update the baseline level of evidence in a trial-trial fashion, explaining sequential dependencies in reaction times (Bitzer et al. 2014; Kim, Kabir, and Gold 2017), thus casting baseline evidence as an expectation that is updated on a trial by trial basis according to the outcome of a choice.

Predicting the timing of future sensory information is informed by prior experience (Jazayeri and Shadlen 2010; Taatgen and Rijn 2011; Acerbi, Wolpert, and Vijayakumar 2012). One timing problem the brain faces when accumulating choice evidence is the delay between the onset of sensory information and the relay of a neural response (Schmolesky et al. 1998; Lee et al. 2016; Lee, Darlington, and Lisberger 2020). Gated accumulator models have been proposed to explain the filtering of a constant stream of visual information from neural integrators, thus preventing constant accumulation of a continuous stream of information in decision networks (Purcell et al. 2012; Schall et al. 2011; Basso and Wurtz 2002; Lo and Wang 2006; Brown, Bullock, and Grossberg 2004). Thus, the model is well-posed to ask questions about how expectations of sensory timing interact with sensory signals prior to evidence accumulation. However, gated accumulation has previously only considered 'sensory' signals as noisy representations of the visual stimulus of a given trial,.

Here, we show that an expectation of sensory timing is combined with delayed visual information prior to the onset of evidence accumulation during saccadic choice. We developed a gated accumulator model in which within-trial fluctuations in baseline evidence are combined with delayed sensory information to trigger the onset of evidence accumulation. The model can thus describe multiple reaction time distributions conditioned on different delay times for a single parameter set, as well as anticipatory response distributions arising from baseline dynamics in absence of sensory input. To test how trial by trial changes to expectation

influence gated accumulation, we designed a free choice saccade task where participants had to balance withholding anticipatory responses with responding as fast as possible to one of two choice targets that were presented with variable asynchrony. A choice history analysis suggested that the outcome of the previous trial, despite randomization, influenced the balance between anticipatory and sensory-guided reaction times. By fitting the gated accumulator model to the choice history data, we show that changes to the baseline dynamics during the delay period were necessary to explain how repetitions changed the distribution of reaction times. This suggests that evidence accumulation is triggered by combining delayed sensory information with expectations from previous trials in order to control the timing of accumulation onset.

Gated accumulation of sensory expectations

The gating of information between different neural structures is an established mechanism of information transfer (Gisiger and Boukadoum 2011). In Purcell et al. 2012, evidence accumulation onset is proposed as a specific instance of a gating mechanism that controls the transformation of delayed visual information to choice evidence. It is commonly thought that such evidence is integrated or amplified towards a threshold, upon which a corresponding motor command is produced (*Fig 1A*, Schall 2019; Cisek, Puskas, and El-Murr 2009). Models of the accumulation process have thus served a joint role in both providing an algorithmic explanation of the variability observed in how long it takes to transform visual information into a motor command, as well as a mechanistic explanation of the ramping neural activity observed prior to a movement (Hanes and Schall 1996; Gold and Shadlen 2007). In recognizing the necessity for such a process to be under active inhibition, Purcell et al. 2012 extended the descriptive capacity of accumulator models to explain behavioural and neural variability arising from different set sizes in visual search (see also Schall et al. 2011).

Algorithmically, gated accumulation (*Fig 1B*) describes the transformation of sensory information into choice evidence as a two-stage process. The first stage of the model describes a sensory signal that is relayed to an evidence integrator, upon which the signal is accumulated towards a threshold. This replaces the notion of a baseline level of evidence plus a fixed sensory delay, common in classic accumulator frameworks, with what is essentially a sensory filtering stage. Although it is thought by many that predictions about prior experience are present in the earliest stages of visual processing, the sensory information in the gated accumulator proposed by Schall et al. 2011 only accounts for information contained in the stimulus present at the time. Thus, the delay time between trials and any information from past stimuli are not considered.

In this section, we derive a gated accumulator that receives a continuous input to evidence accumulation during the delay interval of a trial. We will then propose that gated accumulation provides a sufficient explanation of experimental findings in saccadic choice that are often treated independently: anticipatory responses, history effects, and urgency. To explain our approach, we will first decompose the model into two sections: a baseline dynamics section explaining the generative statistics of the model during the delay period between experimental trials, followed by an explanation of the accumulation of choice evidence to a threshold. Then, we will explain how our gating model combines these processes mathematically.

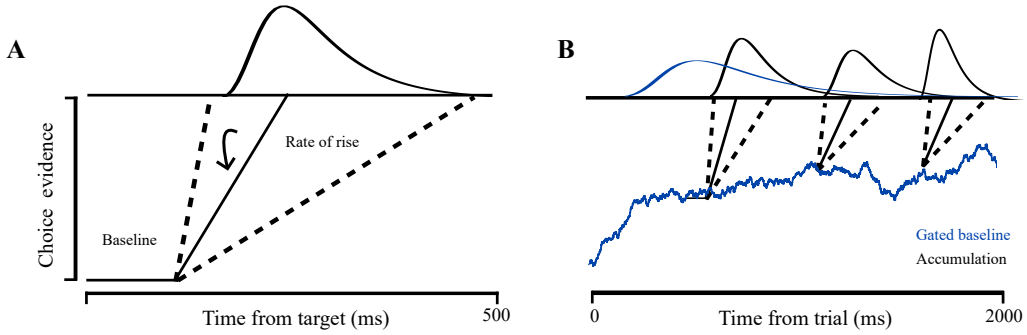

Figure 1: Gated accumulation of choice evidence. **A**: Schematic of a linear approach to threshold with ergodic rate (LATER) model. Sensory evidence is integrated upon a baseline level of choice evidence towards a movement-triggering threshold. The strength of sensory evidence typically controls the average rate of rise. Variability in the rate of rise yields variability in reaction times (top curve). Time in the model begins relative to the onset of a repeated stimulus **B**: Gated accumulation. Here, a baseline dynamic evolves from the beginning of a trial with variable target onset periods, yielding an anticipatory response function (blue) and separate integration processes for different delay intervals (black).

Baseline dynamics

We first describe the baseline dynamics that evolve throughout the delay period of a single trial. In order to account for reaction times that occur throughout the trial, we assume that the baseline dynamics can either trigger anticipatory responses, or set a baseline evidence for accumulation if an anticipatory response is not triggered prior to the presence of new sensory input. We chose an Ornstein-Uhlenbeck (OU) process to model the baseline dynamics [1]. The OU process can describe either purely random baseline dynamics or a reversion to a steady state at different speeds, allowing us to model baseline dynamics as either carrying an informative or uninformative prior on the timing of future sensory information. The OU process is a stochastic differential equation that describes the time evolution of a random variable changing proportionally to the distance from an expected mean (*Fig 2*). This means that the OU process has a steady-state it will eventually reach, and the parameters of the process dictate the initial conditions, the speed, and the variance. Let $B_1, B_2, ...B_n$ be a set of processes which evolve as a time series during a trial of length $T_{0:Tmax=2000}$, where $b_i = b_1, b_2, ...b_n$ is a single process satisfying the stochastic differential equation

$$db_t = \lambda(\theta - b_t)dt + \sigma_B dW_t \tag{1}$$

where λ is the drift rate, θ is the steady-state baseline the process reverts to as time approaches infinity, and W is a Wiener process scaled by σ_B.

We further assume that if this system reaches the expected baseline θ prior to the onset of a choice target, then an anticipatory responses is triggered. Thus the rate λ dictates if and how close to θ the system

[1]For simplicity, we chose a dynamical system with analytically tractable properties that match the statistics of the saccadic choice paradigm we later study. In particular, the first-passage solution for a generic dynamical system is typically not known *a priori*, and often requires expensive simulation when analytical solutions are not known. However, it should be noted that this is purely a technical justification, and is not meant to argue for this particular dynamical system as being either necessary or sufficient to explain ground-truth dynamics. Nor would we claim that such a ground-truth dynamic, if it was even theoretically decodable, would generalize to any behaviour outside the task set. It is simply a sufficient explanation for the behavioural variability we do report, and may be supplanted by other technical justifications or experimental specifications

reaches prior to target onset, and thus if an anticipatory response is triggered. The anticipatory response distribution is thus defined by the first-passage problem where the threshold is equal to the long-run mean θ, which is described in Lipton and Kaushansky 2018 as the solution to the problem

$$T_{A,C} = \inf\{t \geq 0 : B_t = A | X_0 = C\} \tag{2}$$

with change of variables

$$\tilde{t} = \lambda t \tag{3}$$

$$\bar{C} = \frac{\sqrt{2\theta}}{\sigma_B}(C - \theta)$$

$$\bar{A} = \frac{\sqrt{2\theta}}{\sigma B}(A - \theta)$$

The anticipatory density function for the case where $\bar{A} = 0$ is known.

$$P_A(t) = \lambda\sqrt{\frac{2}{\pi}} \cdot |\bar{C}| \cdot \frac{\exp(-\theta t)}{(1 - \exp(-2\theta t))^{3/2}} \cdot \exp\left(-\frac{\bar{C}^2 \cdot \exp(-2\theta t)}{2 \cdot (1 - \exp(-2\theta t))}\right) \tag{4}$$

Because $\bar{A} = 0$ is equivalent to setting the threshold crossing A to equal θ, this equation becomes the first passage time solution for when the threshold $A = \theta$. Thus, this equation describes the probability density of how long it takes the OU process to reach the long-run mean θ, scaled by the rate parameter λ.

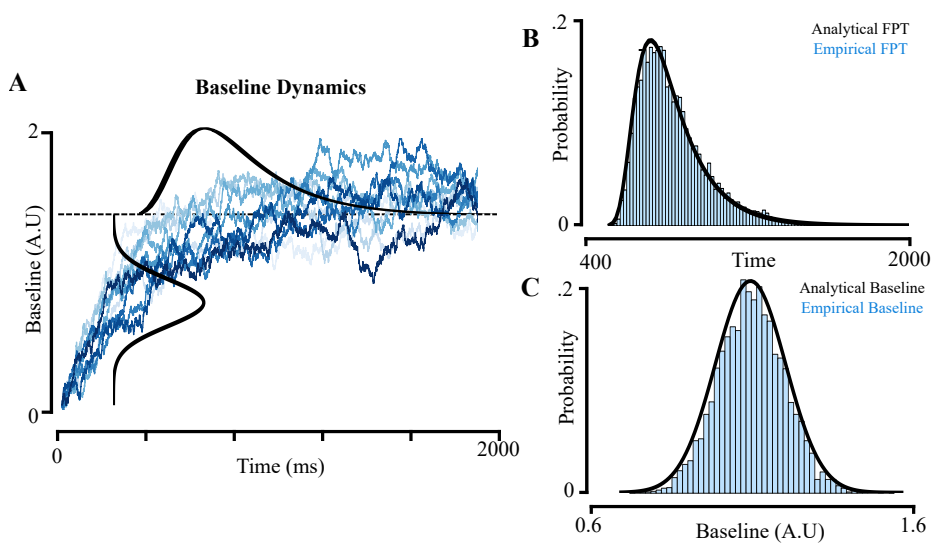

Figure 2: Baseline Dynamics. **A**: Sample paths of an Ornstein-Uhlenbeck process used to model the baseline dynamics. Each blue trace is one sample path drawn from the same parameter set: θ = 1, μ = 1.5, σB = 0.5, X0 = 0 **B**: The first passage time for this parameter (black: analytical solution, blue: empirical histogram for 1000 processes) set is plotted as the probability that the OU process has reached the μ value at time t. This is also seen in the heavy-tailed distribution above the threshold in **A**. **C**: probability distribution of baseline values at t = 250 ms. The distribution at any time follows a Gaussian distribution (colour scheme same as **B**).

A graphical depiction of the baseline dynamics of a sample OU process is depicted in *Fig. 2*. In panel

A, different instances of the same OU parameter set are plotted in the blue traces, with each trace settling to an average value according to the long-run mean parameter $\theta = 1.5$ at a rate according to λ with some variance σ_B. We can solve analytically for the first passage time for a given parameter set, yielding our anticipatory response function, as the time it takes to reach the steady state (*Eq. 4*), while also describing the distribution of baseline activity at each time point (*Fig 2C*).

Evidence accumulation

We next describe the integration of evidence to a threshold following the onset of sensory information. For simplicity, we chose to describe the evidence accumulation process that occurs after the onset of a choice target as minimally as possible, so as to focus on the effects of input during the delay period. Perhaps the simplest evidence accumulation model is the linear approach to threshold with ergodic rate (LATER) model. The LATER model captures reaction time as a deterministic signal with variability across trials using 3 free parameters. Any single trial is determined by

$$s = rt \tag{5}$$

where t is the response time, s is the threshold-baseline difference, and r is the rate of rise. The distribution of reaction times can be derived using process parameters μ_r, the mean rate of rise, s the distance from baseline to threshold, and σ_r, the variance of the rate of rise

$$P_L(t) = \frac{1}{t^2}\frac{1}{\sqrt{2\pi}\mu_r/s}\exp\left[(-\frac{1}{2(\frac{\sigma_r}{s})^2})(\frac{1}{t} - \frac{\mu_r}{s}^2))\right] \tag{6}$$

In this model, the baseline is a single value, and the variance required to explain a ***distribution*** of reaction times comes from trial-trial variance in the integration rate. Although the LATER model fits the majority of a reaction time distribution, it typically does not account for very-short latency reaction times (Noorani 2014) [2]. In order to treat the entire reaction time distribution under this framework, and crucially the responses close to the time of target onset, a derivation of the LATER model with variance to the baseline value was proposed (Nakahara, Nakamura, and Hikosaka 2006). This 'extended LATER' model has an analytical solution if the prior is Gaussian, $N(\mu_s\sigma_s)$

$$P_E(t) = \frac{t + r_1\alpha^2}{(t^2+\alpha^2)^{3/2}}\frac{1}{\sqrt{2\pi}\Sigma}exp\left(-\frac{1}{2\Sigma^2}(\frac{1}{t} - r_1)^2\frac{t^2}{t^2+\alpha^2}\right) \tag{7}$$

where $\alpha = \frac{\sigma_s}{\sigma_r}, \Sigma = \frac{\sigma_r}{\mu_s}, r_1 = \frac{\mu_r}{\mu_s}$.

The extended LATER model allows us to calculate reaction time distributions for different ***distributions*** of baseline activity, rather than using the single baseline parameter in the LATER model. As we will see, the convenience of having a normally distributed baseline will allow us to plug in a stochastic process that

[2] In the study of the saccadic system, these short latency responses are often called 'express' saccades. The need to classify these responses distinctly arises from the bimodality observed in reaction times. It is a matter of debate whether these responses are functionally distinct, in the sense that a separate physical mechanism is required to explain their generation (Kingstone and Klein 1993). Our approach does not make a distinction – we will consider very short latency responses as arising from a mixture of anticipatory responses and integrated responses

evolves during the delay period, resulting in extended LATER processes with different baseline distributions, and corresponding reaction time distributions, at each delay time for a single model.

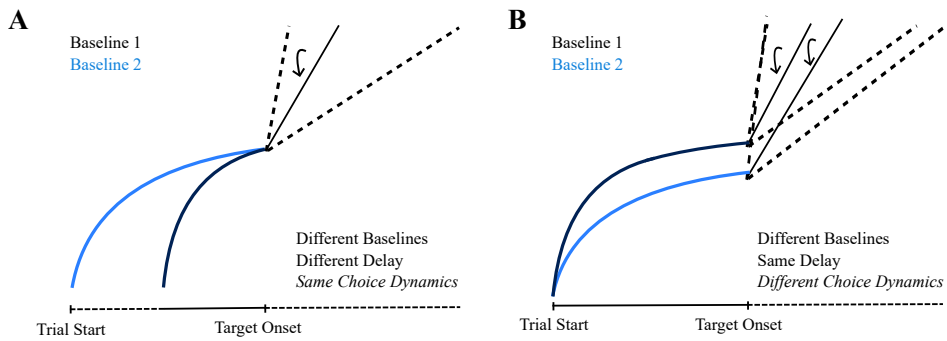

Figure 3: Inferring difference in baseline dynamics depends on starting point. **A**: two trials where baseline dynamics (high rate parameter: light blue, low rate parameter: dark blue) with the same delay time reach the different baseline levels, resulting in distinct accumulation processes. **B**: two baseline dynamics with different delay interval lengths, resulting in identical baseline evidence at target onset, and thus indistinguishable reaction time distributions.

Combining evidence accumulation with baseline dynamics via gating

We next need to describe how the two probability density functions we have previously described, the anticipatory response distribution P_A and the evidence accumulation distribution pT, are combined into a common process with a single probability density function relative to a given delay interval. We first treat the simplest case, where only a single target onset time is used. Thus, we need to compute the anticipatory first passage problem for the OU process, as well as the mean and variance of the process itself at the corresponding delay time, which we will use as the mean and variance to the baseline Gaussian distribution in the extended LATER model. We can thus compute the mean of the extended LATER baseline μ_r at time t as the expectation of the Ornstein-Uhlenbeck process, which is well-known

$$\mathbb{E}[x_t|x_0] = x_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}) \tag{8}$$

and the covariance

$$\mathbb{COV}[X_s, X_t] = \frac{\sigma^2}{2\theta}(e^{-\theta|t-s|} - e^{-\theta((s+t)}) \tag{9}$$

where $s = 0$ and $t = y$. Thus, while adding 3 free parameters from the OU process, we reduce two free parameters from the extended LATER process pT by allowing the OU process to control the baseline Gaussian at a given target onset time t. We can then define the gating process, G, as a combination of the anticipatory first passage distribution P_A, and the extended LATER solutions P_L for a distribution of delay intervals D over the interval $[a, b]$.

$$G = P_A + \int_a^b P_E(t_0 = D)dD$$

The key distinction between the way in which we have derived this model and other evidence accumulator models is that by explicitly accounting for the delay interval and anticipatory responses, the onset of

sensory information is no longer a fixed, experimentally controlled parameter that references $t = 0$, or $t = 0$ plus some fixed sensory delay parameter. Instead, the model output is the combination of the parameters of the model and the statistical distribution of delay times in the task. Under this view, we assume that every individual trial, with an associated delay time drawn from some experimental task distribution, has a theoretical reaction time distribution associated with a given parameter set. Thus, a distribution predicted by the model is the joint distribution of the process normalized over each observed delay time.

However, we cannot model the observed distribution of single unique temporal events, so an alignment strategy must be chosen to find a common temporal reference point for different trials. *Fig 3* depicts a cartoon scenario where the choice of temporal alignment can obscure differences in the output of distinct systems. If two different baseline dynamics, depicted in dark and light blue respectfully, were aligned relative to the onset of a stimulus but had different delay times (*Fig 3A*), they may produce identical baselines at the time of target onset, and as such not be dissociable. However, if the same baseline dynamics were shifted such that the common starting time was the trial, and not the target, the same baseline dynamics would lead to an observable difference in reaction times (*Fig 3B*). This highlights that distinctions between underlying processes can be masked by the way in which reaction time data is aligned, while identical processes can result in quantitatively different behaviour if the temporal statistics of the task are appropriately varied.

Experimental evidence for gated accumulation of sensory expectations

To assess how predictions influenced the gating of evidence accumulation, we designed a free choice saccade task (*Methods: Free Choice Task*) where participants had to rapidly respond to the onset of sensory information while withholding anticipatory responses. Each trial begins with a central fixation period drawn from a uniform distribution between 750 and 1250 ms (*Fig 4A*). This period was followed by the appearance of two visual targets, one in each visual hemifield. The targets appeared at different times relative to each other, with the target onset asynchrony (TOA) randomized across 9 discrete values 0, +/- 16, 33, 66, 99ms. For our convention, positive asynchrony values will denote the right target appearing prior to the left, and negative asynchrony values will denote the left target appearing before the right. Participants were instructed to direct a saccade to either target as fast as possible, without anticipating. Although both delay and TOA were randomized, any given sample of trials from a uniform distribution may have biased statistics in the spatial/across-trial (target location) and temporal/within-trial (target onset and asynchrony) domain.

We depict the generative statistics of the TOA in *Fig 4C* as a random walk. Although the expected value of an ensemble of random walks is a constant value (solid black trace), an individual sequence (example trace highlighted in blue)) are likely to deviate from the ensemble. Thus, the variance of the ensemble grows as a function of the number of trials, as is depicted in the collection of grey background traces. Within each trial, there is an additional probabilistic element: the delay period. *Fig 4D* depicts the cumulative probability of the first target as a function of the length of time from trial onset.

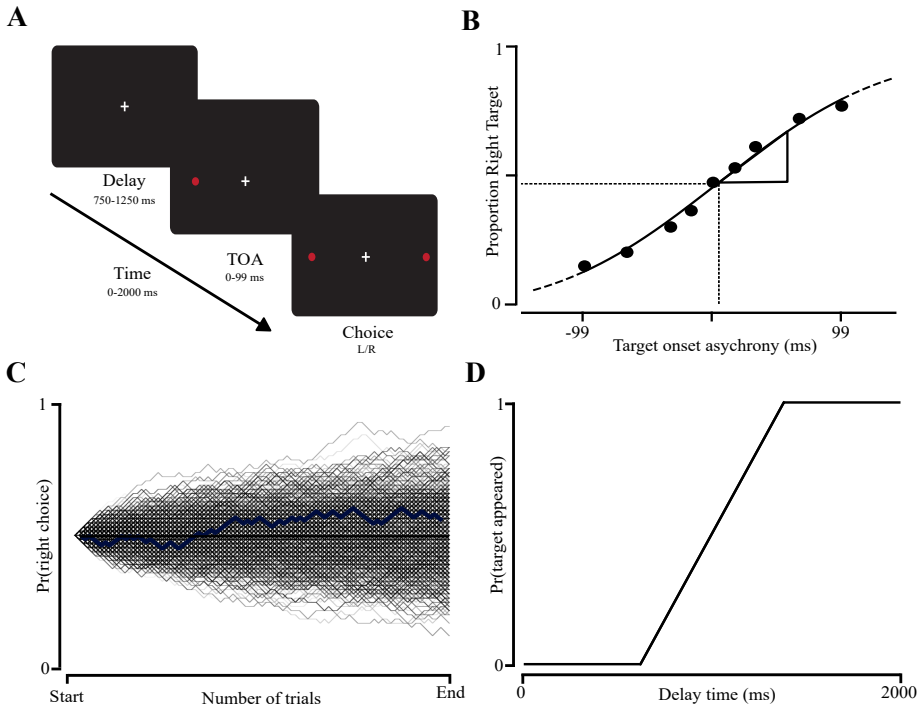

Figure 4: Free Choice Saccade Task. **A**: Participants directed saccades to either of two free choice targets as fast as possible, with a randomized delay time and target onset asynchrony (TOA).**B**: Group-average psychometric function, depicting the probability of right target (y-axis) as a function of TOA. **C**: Across-trial task statistics. Depicted is a random walk (left), showing the probability of right target selection, beginning at 0.5, across different hypothetical blocks. **D**: Within-trial task statistics. Depicted is the probability of the first target appearing as a function of the delay time of the trial.

Choice probability is history-dependent

Psychometric analysis revealed that choices are probabilistic with respect to the magnitude of asynchrony (*Fig 4B*). Choice behaviour binned by asynchrony indicated a progressive increase in rightward choices when the right target appeared earlier. The binned choice data was fit to a sigmoidal curve with a psychometric fitting procedure using Psignifit (Schütt et al. 2016). Note in this analysis only choices occurring after the target onset plus a minimum delay of 70ms were included, as is often standard when excluding anticipatory trials (Schmolesky et al. 1998; Nijhawan 2008). Synchronous targets resulted in choices closest to chance with respect to left or right direction, while a target appearing well before another increased the likelihood of response towards it. This indicated that participants were aware of the task instructions, and that the range of TOAs used in the task were sufficient to evoke a change in choice direction probability..

In randomized tasks, trials are uncorrelated, so any notion of optimality should begin by treating trials as statistically independent. However, one may infer sequences that arise randomly to be a part of a pattern that generalizes to future trials. Although average choice probability in this task should yield an equal probability of selecting the left or the right target on each trial, that trial is preceded by and follows a particular history of repetitions and alternations that may carry local contextual effects. A binary choice tree for 5 choice in the past is schematized in *Fig 5A*. In the limit of contextual logic, each choice could

be considerd to follow a unique sequence, however this would limit our capacity to perform any statistical analysis. To this end, we compared sequences of choice repetitions to the same target direction, and sequences of choices alternating targets, as measures of choice sequences with a stereotyped local structure that may induce bias away from the true generative statistics of the task (Cho et al. 2002). *Fig 5B-C* shows a histogram of the number of repetitions and alternations in the present data-set. Because the data become increasingly limited with higher sequence number, further analyses were conducted up to a sequence length of 5.

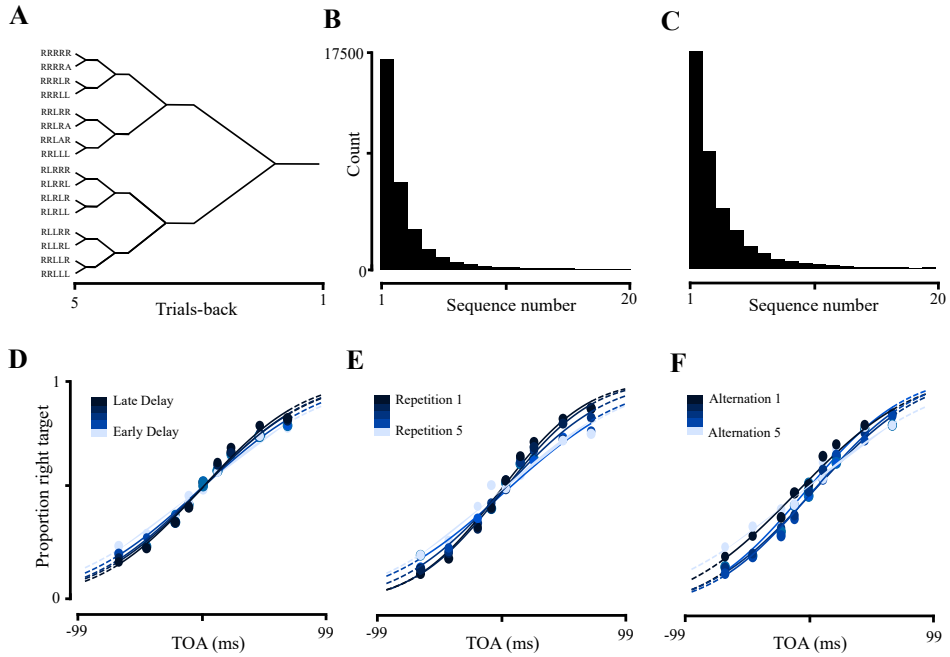

Figure 5: Choice history and delay time influence choice probability. **A**: Choice history tree depicting the possible alternatives of a binary choice up to 4 trials back. **B**: Histogram of the number of choice repetitions across all participants. For each sequence number, a single tally was recorded if the same choice direction had been chosen for that many trials previously. **C**: The same as B, but for the number of successive choice direction alternations. **D**: Group-average psychometric functions with respect to TOA, where the data has been split into 4 delay bins equally spaced between 750 and 1250 ms. The colour gradient denotes the delay period, low to high. **E**: Psychometric functions split into the number of previous choice repetitions. Here, each repetition sequence was given an independent repetition score, such that no trials in one repetition set are contained in another **F**: The same as **E**, for choice alternations.

Psychometric analysis (*Methods: Psychometric Fitting*) revealed that choice probability was influenced by the number of previous repetitions and alternations. In *Fig 5E*, psychometric fits to group-averaged repetition data are plotted as a function of repetition number (blue gradient), showing a progressive decrease in the slope of the sigmoid function. We interpret this as indicating that following choice repetitions, participants relied less on the TOA of the current trial, and more on information derived from past trials. To quantify this, we estimated parameters of the psychometric function using a Bayesian procedure (*Methods: Psychometric fitting*). We then bootstrapped estimates of the slope from the resulting probability distribution over the slopes, and performed a Kruskal-Wallis to ask if the parameter estimates came from the same distribution ($F = 2829.13, p = 0$). This suggests that the distribution of best-fitting psychometric slopes

was not constant across repetitions. We performed this same analysis for sequential alternations, finding that choice probability was also not constant ($F = 476.23, p = 9.18^{-102}$).

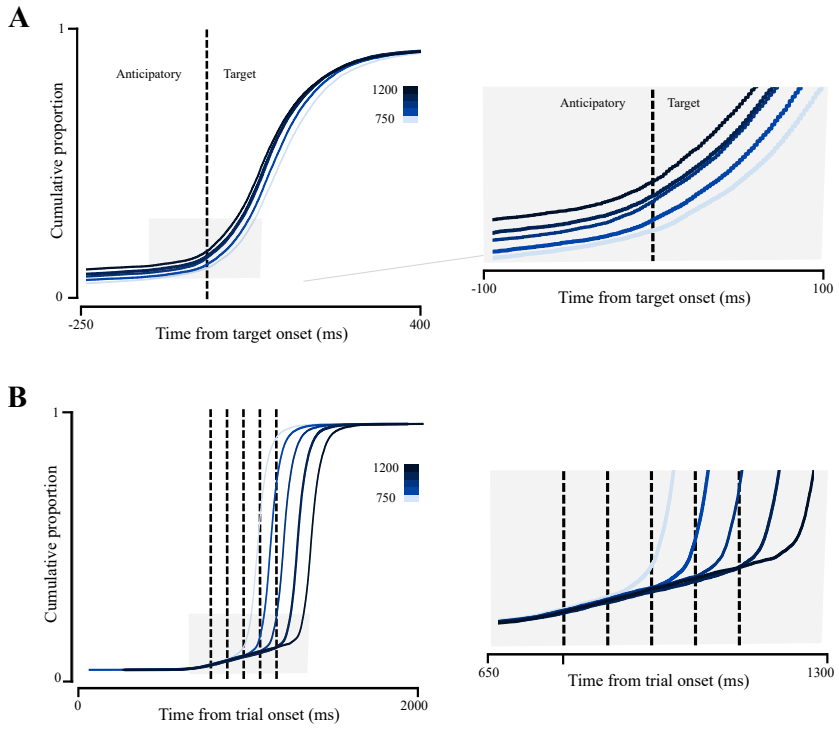

Figure 6: Anticipatory and sensory-guided reaction times are influenced by choice history. **A**: Cumulative probability of reaction times as a function of 10 delay bins, aligned to the onset of the target (left) and trial (right). The onset of the cumulative probability functions deviating from zero in target-aligned responses represent the proportion of anticipatory responses visible in the trial-aligned graph. **B**: Cumulative response functions separated by the number of previous repetitions of the same choice direction. **C**: The same, for alternations.

In addition to the TOA, we also controlled the length of the delay interval within-trials. We next assessed whether the length of the delay interval influenced choice probability. To do this, we performed the same binning procedure as previously outlined for choice sequence, but instead we binned the delay period into 5 equally spaced intervals between $750ms$ and $1250ms$. The resulting psychometric functions are shown in *Fig 5F*. Again, we performed a Kruskal-Wallis test, which supported the qualitative observation that the best-fitting psychometric functions were not constant when separating data by the length of the delay interval ($F = 384.77, p = 5.43^{-82}$). In combination, these analyses support the conclusion that both choice sequence and delay time influenced choice behaviour. However, choice probability only yields a discrete measure for each individual trial. In order to understand how choice history and delay time influenced underlying choice dynamics, we next analyzed reaction times.

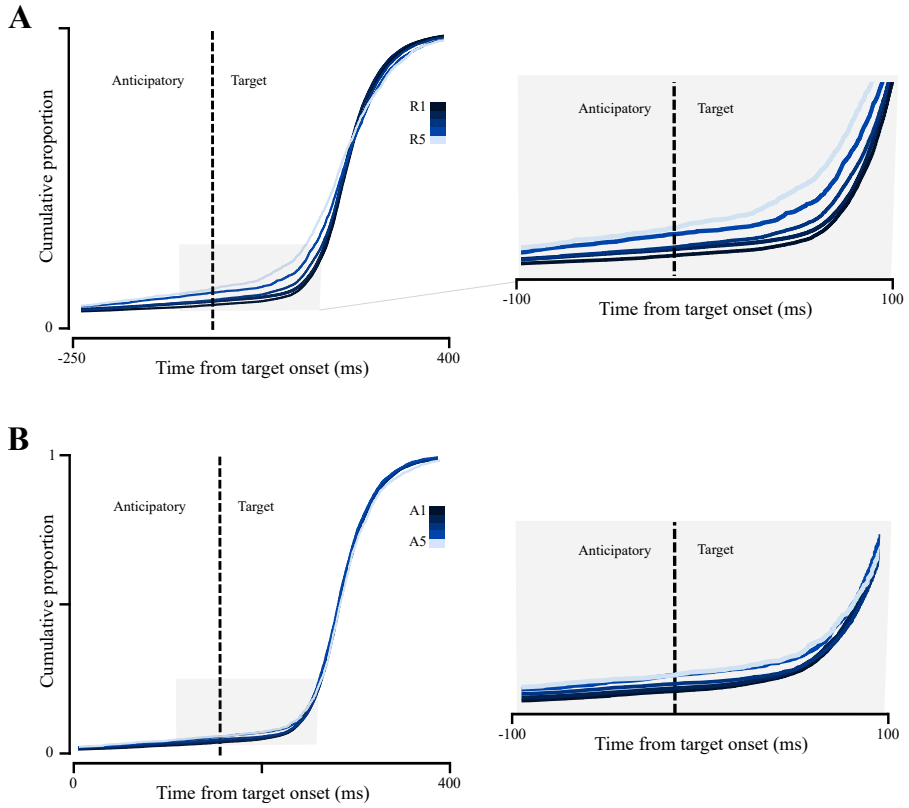

Figure 7: Anticipatory and sensory-guided reaction times are influenced by choice history. **A**: Cumulative probability of reaction times as a function of 10 delay bins, aligned to the onset of the target (left) and trial (right). The onset of the cumulative probability functions deviating from zero in target-aligned responses represent the proportion of anticipatory responses visible in the trial-aligned graph. **B**: Cumulative response functions separated by the number of previous repetitions of the same choice direction. **C**: The same, for alternations.

Choice history jointly influences anticipation and the accumulation of choice evidence

Saccadic timing is often factored into three classes of response: anticipatory responses that occur before stimulus onset, sensory guided responses, and an 'express saccade' regime that lies between the two (Kingstone and Klein 1993). In the psychometric analyses, only choices occurring after sensory onset plus a minimum delay of 70ms were considered (Schmolesky et al. 1998), so as to analyze the role of target onset asynchrony. However this was only a subset of the total behaviour; while participants were asked to try not to anticipate, they were also asked to go as fast as possible. We thus performed an analysis of reaction times throughout the delay interval in order to better understand how these different types of responses covaried.

First, we assessed the influence of the length of the delay interval on anticipatory and sensory-guided responses. *Fig 6* plots the cumulative distributions of reaction times split into 5 equally sized delay bins between 750 and 1250ms. We aligned the distributions in two ways. First, we aligned relative to the onset of the target, and plotted anticipatory responses as negative time relative to this (*Fig 6A*). Next, we aligned the data to the onset of the trial, such that each delay bin will have a mixture of onset times bounded within the delay bin. In each view, it can be seen that anticipatory responses increase, and sensory-guided responses

get faster, when the delay interval was longer. (Reddi and Carpenter 2000; Standage, Wang, and Blohm 2014). Qualitatively, this resulted in a broadening of the distribution as a function of the delay interval, since an increase in anticipatory responses corresponds with an increase in the early tail of the distribution. To quantify this, we randomly sampled from each distribution and bootstrapped the empirical standard deviation, and computed the 95% confidence interval for each. This supported the conclusion that reaction time variance was dependent on the delay interval (*Methods: statistical testing of differences in reaction times*).

We next analyzed reaction times as a function of choice history (sequential repetitions and alternations). *Fig 7* plots the reaction times aligned to the target dependent on how many previous repetitions (A) and alternations (B) preceded the current trial. Here, it can be seen that sequential repetitions were associated with an increase in anticipatory responses, as seen in the cumulative proportion lying before the dashed line indicating target onset, and faster post-target responses. However, this speed advantage was negated for responses in the median range, and even reversed for later responses, consistent with the idea that choice sequence influenced an association with baseline evidence in accumulator models (Nakahara, Nakamura, and Hikosaka 2006; Noorani 2014). This effect was detectable but noticeably diminished in magnitude for alternations, consistent with the idea that sequence effects may operate over different timescales when targets alternate (Cho et al. 2002).

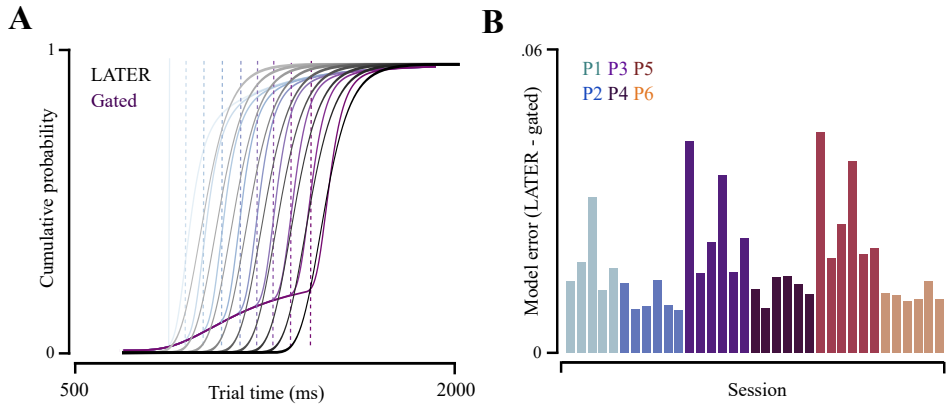

Figure 8: Gated accumulation jointly explains anticipatory and sensory-guided reaction times. **A**: Best-fitting cumulative distributions of reaction times for extended LATER model and gated model to data pooled across participants, split into 10 delay bins. *B:* Model comparison between extended LATER model and gated model for each participant session. Positive values denote worse performance for the LATER model as compared to the gated model

In summation, we have evidence that reaction times in our sample were jointly influenced by choice history and the delay time of the trial. By analyzing reaction times not by summary statistics such as the mean or median, and instead viewing entire distributions relative to different onset times, we were able to conclude that this joint influence changed in magnitude over the course of the trial. This is consistent with the hypothesis made by our model derivation, in which anticipatory responses and baseline evidence are controlled by a common process.

Gated accumulation of sensory timing expectations accounts for choice history and anticipation

Biophysically motivated models of evidence accumulation have suggested that the onset of accumulation is gated in order to prevent prepotent responses to a continuous stream of sensory inputs (Purcell et al. 2012; Schall et al. 2011). The gating of evidence accumulation has been shown to adapt to speed-accuracy trade-offs in stop-signal tasks, a trade-off conceptually similar to the balance between predictive and reactive choice behaviour we observed empirically. We derived a model in which the gating of evidence accumulation is not driven solely delayed sensory information, but also by within-trial fluctuations in baseline dynamics. Here, we demonstrate that this model can explain the statistical distribution of anticipatory and sensory guided responses as a function of delay intervals and choice history.

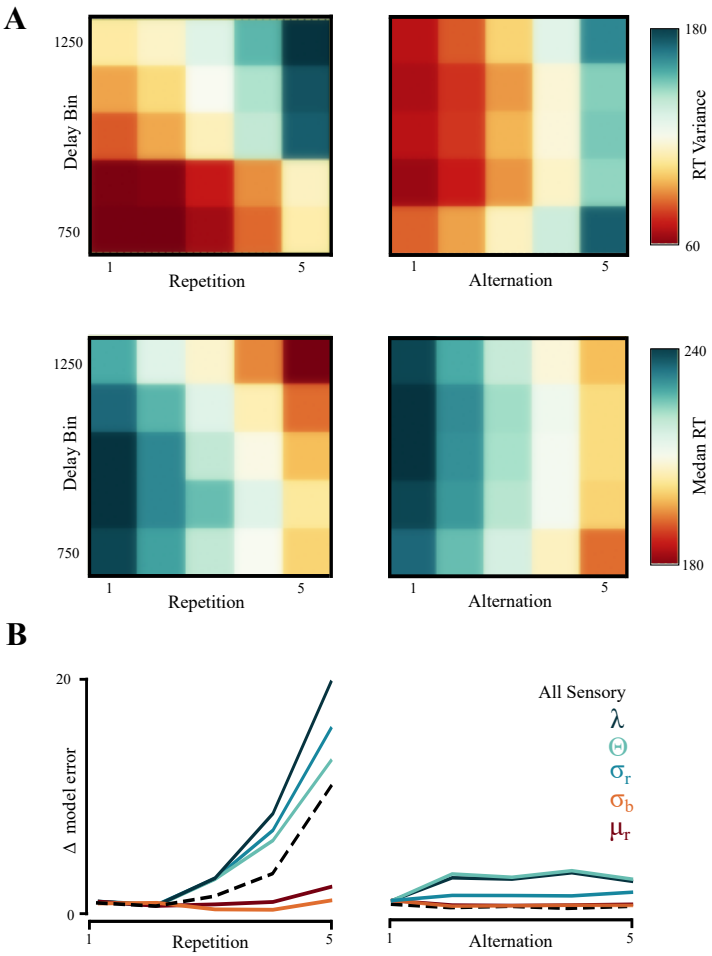

Figure 9: Gated accumulation captures dependence of reaction times on choice sequence and delay time. **A**: Reaction time variance for best-fitting models is plotted for choice sequence and delay time. Each choice sequence model has a unique parameter set that is constant across delay bins. Variance is computed as the empirical standard deviation of 10,000 random draws from the model for a corresponding delay. Alternations are plotted on the left, and repetitions on the right. **B**: The same, for median reaction time.

First, we demonstrated that the gated model improved fits to behavioural data relative to the extended

LATER model (Nakahara, Nakamura, and Hikosaka 2006). We fit each model to group-averaged data and independent behavioural sessions across our 6 participants (*Methods: Fitting to free choice saccade data*), and contrasted the model error for each session from the extended LATER model with the gated model. *Fig 8A* plots the resulting distributions for 10 delay bins for group-averaged data in both models. Note qualitatively how the shape of each extended LATER distribution is constant, and does not capture anticipatory responses occurring prior to the target onset associated with each delay bin. To show this was not an effect of participant-averaging, we fit the gated model independent to different sessions of each participant, and computed the difference in performance between each model. Here, we subtracted the model error of the extended LATER model from the gated model, thus a positive value indicates that the gating model better explained the distribution of reaction times for that session. Across all participant sessions, the gating model was better at explaining the data, indicating that it was able to capture variability not expressed in the extended LATER model (*Fig 8B*).

Next, we wanted to demonstrate that the model can capture the sequence effects we previously reported. To this end, we combined participant data together and split the data by choice sequence (repetition/alternation 1-5). Again, we used 5 delay bins, so each choice sequence set minimized the model error over 5 delay bin distributions for a single parameter set. We took the best-fitting models and sampled from them 10,000 times to compute reaction time variance and median (*Fig 9A*). We then plotted the reaction time variance (top) and median (bottom) colour-coded low (red) to high (blue). By looking at the diagonal trend from bottom left to top right, it can be seen that the repetition models have a higher joint effect of delay time and choice sequence number on reaction times, consistent with our analysis of the behavioural data. In contrast, the effect of alternations appears to be less dependent on the delay interval.

Once we confirmed the model can capture the data, we fixed individual parameters in the model fits to understand which parameters were necessary to describe the changes in reaction times obesrved. To this end, we fixed all parameters in the best-fitting model averaged across all data, and iteratively fit the the model to each choice sequence while allowing each individual parameter to vary. This resulted in a total of 25 model fits for repetitions, and 25 model fits for alternations (5 sequence x 5 parameter variations). In *Fig 9B*, the resulting model error, computed as the difference in Kullback-Leibler divergence between the free and fixed model, is plotted as a function of choice sequence number (repetitions and alternations plotted separately). In this plot a positive change in model error indicates a decrease in model performance after fixing the corresponding parameter. Here it can be seen that 3 parameters, the mean-reversion baseline parameter θ, the baseline rate parameter κ, and the variance of the sensory integration parameter σ_r resulted in worse model performance as a function of choice repetition number. These three parameters resulted in larger model errors in alternations, but reduced in magnitude, consistent with the smaller empirical change in reaction times. This suggests that variability in two parameters, the baseline dynamics variance σ_b, and the sensory integration rate μ_r, were independently able to explain how reaction times varied when conditioned on choice sequence. This effect was strongest in repetitions, where the strongest choice sequence effect was present, and reduced in alternations. In contrast, fixing the parameters and allowing the mean-reversion baseline parameter θ, the baseline rate parameter κ, and the variance of the sensory integration parameter σ_r to vary resulted in reduced model performance, suggesting that isolated changes to each parameter could not explain this variability.

From this, we can see that two parameters, the baseline dynamics variance σ_b, and the sensory integration rate μ_r, were able to independently explain variability in the dataset when conditioned on choice sequence. This effect was strongest in repetitions, where the strongest choice sequence effect was present, and reduced in alternations. In contrast, fixing the parameters and allowing the mean-reversion baseline parameter θ, the baseline rate parameter κ, and the variance of the sensory integration parameter σ_r to vary resulted in reduced model performance. The observation that a parameter of the baseline dynamics, σ_b, and the sensory integration rate μ_r were independently sufficient to capture variance in reaction times suggests that the model may provide multiple solutions that result in similar reaction time profiles. To confirm this, we next fixed all the parameters of the baseline dynamics, and allowed only the sensory integration parameters to vary. As can be seen in the dashed black line, fixing these parameters resulted in an incremental change between the parameters that were and were not able to explain choice sequence-dependent variability. From this combination of results, we can conclude that isolated parametric changes to baseline dynamics as a function of choice sequence were sufficient to explain reaction time variability.

Discussion

In this paper, we used a model-based approach to ask how expectations of sensory timing influence the accumulation of evidence during a free choice saccade task. Our model extends a gated accumulator (Schall et al. 2011; Purcell et al. 2012) to capture variability in the delay period of tasks when sensory information is withheld. In doing so, we provide a technical advancement on the accumulator framework by explicitly modelling reaction times as a product of anticipatory responses and statistical distributions of delay intervals, and experimental evidence for a link between phenomena in choice tasks – anticipation, urgency, history effects – that are often considered independently.

Anticipating the timing of sensory information is useful when interacting with recurring patterns in the environment (Nobre and Ede 2018), particularly in the presence of delays. In this study, we concluded that anticipation reflected a balance between lowering the amount of evidence required to trigger a saccade and inhibiting movement generation prior to the appearance of sensory information. The onset of evidence accumulation has been suggested to be under active control and subject to task-specific modulation (Teichert, Grinband, and Ferrera 2016; Brink et al. 2021), but in the analysis of choice tasks, anticipatory responses are usually treated as outliers to be discard – as such, the relationship between anticipation and evidence accumulation is rarely discussed. When it has, anticipation has been explained as arising from an independent and parallel process (Hernández-Navarro et al. 2021), supporting claims for a mechanistic dualism between proactive and reactive processing (Hikosaka and Isoda 2010; Evans 2008). This paper does not endorse or refute this. While we do support the conclusion that a trade-off between anticipation and fast motor responses can result from gating the accumulation of sensory timing expectations, we consider the model statistical in nature – it says nothing absolute about the degree to which anticipatory responses emerge from a distinct physical mode.

We also report urgency effects in our task, that is a decrease in reaction time as a function of the length of the delay period, likely reflecting the increasing probability of a target appearing in our task as time goes on. In accumulator frameworks, urgency has classically been associated with a change in the rate of

evidence integration (Reddi and Carpenter 2000)). An additional class of models eschews the accumulator framework altogether, instead arguing that urgency is represented as an independent signal that is combined with sequentially-sampled sensory information (Cisek, Puskas, and El-Murr 2009). Differentiating these models empirically has been a matter of debate (Carland, Thura, and Cisek 2015; Winkel et al. 2014; Evans et al. 2017). In light of this, we wish to be clear that within-trial changes to baseline evidence, in combination with an anticipatory response distribution, was merely a sufficient statistical explanation of reaction time variability in our task. As we have not performed a formal model comparison of an urgency gating model that accounts for anticipatory responses, if one were to exist, it is unclear to what degree the models can be dissociated empirically. Future work may address if these models make distinct statistical predictions, or simply provide a different lens through which to generate mechanistic hypotheses.

Part of the difficulty in dissociating models and model parameters comes from an additional degree of freedom that must be allowed when one is forced to consider non-stationarity as an additional source of variability. Previously, it has been suggested that baseline evidence in an accumulator framework can be thought of as analogous to a prior belief over the likelihood of one of a limited number of choice alternatives (Carpenter and Williams 1995; Gold and Shadlen 2007), motivating the adoption of reinforcement learning approaches to model trial-trial variability in saccadic reaction times (Kim, Kabir, and Gold 2017; Pedersen, Frank, and Biele 2017; McDougle and Collins 2021). We showed in behavioural data that anticipatory and sensory-guided reaction times are jointly influenced by choice history. In our model, we can account for this dependency by allowing the parameters of the baseline dynamic to remain free across fits to choice sequence data. This is a claim for sufficiency of trial-trial variability to capture sequence effects for a specific data-set, however it does not rule out history-dependent variance in sensory integration, as has been reported elsewhere (Urai et al. 2019; Steingroever, Wabersich, and Wagenmakers 2021), and suggested by the variance in sensory integration influencing our choice sequence fitting. Nor does it make any claim about what the statistical form of the sequential dependencies may be. Without access to the ground-truth of a system's dynamics, it is difficult to estimate the precise form of trial-trial variance in model parameters, simply because the hypothesis space over potential models becomes intractable when forced to consider each individual trial as a singularly-observable outcome of a model that has stochasticity as a starting axiom (Carpenter 1999).

Because of this, we wish to elaborate on the distinction between our analysis of choice history, and any conclusions that may be drawn with respect to the functional form of underlying history-dependence. What we do report is a dependence on the outcome of previous trials on anticipatory and sensory-guided reaction times for a particular task. From the perspective of the instructions of the task, history effects may be considered to denote some form of suboptimality in the decision process – as the task statistics were randomized, information from the previous trial contained no evidentiary value beyond providing a single sample of the task statistics. Arguments for *why* history effects persist in different tasks are numerous. Some have argued that history effects in randomized experiments (often binary series) reflect a capacity limitation of the system (Baddeley 1966; Tune 1964; Kahneman and Tversky 1972. Others reframe the question in the context of optimality by considering these history effects as carry-over from rational behaviour in other contexts where statistical regularity is more prevalent, or question if bias is truly present (Lopes 1982; Ayton, Hunt, and Wright 1989; Kareev 1992). From our perspective, there is little to be concluded about *why* history effects are present in our task and dataset, simply because the dimensionality of sufficient explanations is

such that redundancy is likely, and falsification is a challenge. However, we can conclude that saccadic choice to randomized asynchronous targets was not independent of previous trials, and this dependency could be explained by a model where expectations of sensory timing could trigger the onset of evidence accumulation to minimize the consequences of delays in sensory processing.

References

Abrahamyan, A., et al. 2016. "Adaptable history biases in human perceptual decisions". *Proceedings of the National Academy of Sciences of the United States of America* 113 (25): E3548–3557. doi:10.1073/pnas.1518786113.

Acerbi, L., D. M. Wolpert, and S. Vijayakumar. 2012. "Internal representations of temporal statistics and feedback calibrate motor-sensory interval timing". *PLOS Computational Biology* 8 (11): e1002771. doi:10.1371/journal.pcbi.1002771.

Ayton, P., A. J. Hunt, and G. Wright. 1989. "Psychological conceptions of randomness". *Journal of Behavioral Decision Making* 2 (4): 221–238. doi:10.1002/bdm.3960020403.

Baddeley, A. D. 1966. "The capacity for generating information by randomization". *Quarterly Journal of Experimental Psychology* 18 (2): 119–129. doi:10.1080/14640746608400019.

Bakan, P. 1960. "Response-tendencies in attempts to generate random binary series". *The American Journal of Psychology* 73:127–131. doi:10.2307/1419124.

Bar-Hillel, M., and W. A. Wagenaar. 1991. "The perception of randomness". *Advances in Applied Mathematics* 12 (4): 428–454. doi:10.1016/0196-8858(91)90029-I.

Basso, M. A., and R. H. Wurtz. 2002. "Neuronal activity in substantia nigra pars reticulata during target selection". *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 22 (5): 1883–1894. doi:10.1523/JNEUROSCI.22-05-01883.2002.

Bitzer, S., et al. 2014. "Perceptual decision making: drift-diffusion model is equivalent to a Bayesian model". *Frontiers in Human Neuroscience* 8:1–17. doi:10.3389/fnhum.2014.00102.

Bogacz, R., et al. 2006. "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks". *Psychological Review* 113 (4): 700–765. doi:10.1037/0033-295X.113.4.700.

Brink, R. L. van den, et al. 2021. "Temporal expectation hastens decision onset but does not affect evidence Qquality". *The Journal of Neuroscience* 41 (1): 130–143. doi:10.1523/JNEUROSCI.1103-20.2020.

Brown, J. W., D. Bullock, and S. Grossberg. 2004. "How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades". *Neural Networks* 17 (4): 471–510. doi:10.1016/j.neunet.2003.08.006.

Busemeyer, J. R., and J. T. Townsend. 1993. "Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment". *Psychological Review* 100 (3): 432–459. doi:10.1037/0033-295X.100.3.432.

Carland, M. A., D. Thura, and P. Cisek. 2015. "The urgency-gating model can explain the effects of early evidence". *Psychonomic Bulletin & Review* 22 (6): 1830–1838. doi:10.3758/s13423-015-0851-2.

Carpenter, R. H. S. 1999. "A neural mechanism that randomises behaviour". *Journal of Consciousness Studies* 6 (1): 13–22.

Carpenter, R. H. S., and M. L. L. Williams. 1995. "Neural computation of log likelihood in control of saccadic eye movements". *Nature* 377 (6544): 59–62. doi:10.1038/377059a0.

Cho, R. Y., 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. doi:10.3758/cabn.2.4.283.

Cisek, P., G. A. Puskas, and S. El-Murr. 2009. "Decisions in changing conditions: the urgency-gating model". *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29 (37): 11560–11571. doi:10.1523/JNEUROSCI.1844-09.2009.

Dzhafarov, E. N. 1993. "Grice-representability of response time distribution families". *Psychometrika* 58 (2): 281–314. doi:10.1007/BF02294577.

Evans, J. S. B. T. 2008. "Dual-processing accounts of reasoning, judgment, and social cognition". *Annual Review of Psychology* 59:255–278. doi:10.1146/annurev.psych.59.103006.093629.

Evans, N. J., et al. 2017. "The computations that support simple decision-making: A comparison between the diffusion and urgency-gating models". *Scientific Reports* 7 (1): 16433. doi:10.1038/s41598-017-16694-7.

Fan, Y., J. I. Gold, and L. Ding. 2020. "Frontal eye field and caudate neurons make different contributions to reward-biased perceptual decisions". *eLife* 9. doi:10.7554/eLife.60535.

Gershman, S. J., E. J. Horvitz, and J. B. Tenenbaum. 2015. "Computational rationality: A converging paradigm for intelligence in brains, minds, and machines". *Science* 349 (6245): 273–278. doi:10.1126/science.aac6076.

Gilovich, T., R. Vallone, and A. Tversky. 1985. "The hot hand in basketball: On the misperception of random sequences". *Cognitive Psychology* 17 (3): 295–314. doi:10.1016/0010-0285(85)90010-6.

Gisiger, T., and M. Boukadoum. 2011. "Mechanisms gating the flow of information in the cortex: What they might look like and what their uses may be". *Frontiers in Computational Neuroscience* 5:1–15. doi:10.3389/fncom.2011.00001.

Glaze, C. M., J. W. Kable, and J. I. Gold. 2015. "Normative evidence accumulation in unpredictable environments". *eLife* 4:e08825. doi:10.7554/eLife.08825.

Gold, J. I., and M. N. Shadlen. 2007. "The neural basis of decision making". *Annual Review of Neuroscience* 30 (1): 535–574. doi:10.1146/annurev.neuro.29.051605.113038.

Griffiths, T. L., and J. B. Tenenbaum. 2006. "Optimal predictions in everyday cognition". *Psychological Science* 17 (9): 767–773. doi:10.1111/j.1467-9280.2006.01780.x.

Hanes, D. P., and J. D. Schall. 1996. "Neural control of voluntary movement initiation". *Science*, no. 5286 (274): 427–30. doi:10.1126/science.274.5286.427.

Hernández-Navarro, L., et al. 2021. "Proactive and reactive accumulation-to-bound processes compete during perceptual decisions". *Nature Communications* 12 (1): 7148. doi:10.1038/s41467-021-27302-8.

Hikosaka, O., and M. Isoda. 2010. "Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms". *Trends in Cognitive Sciences* 14 (4): 154–161. doi:10.1016/j.tics.2010.01.006.

Jarvik, M. E. 1951. "Probability learning and a negative recency effect in the serial anticipation of alternative symbols". *Journal of Experimental Psychology* 41 (4): 291–297. doi:10.1037/h0056878.

Jazayeri, M., and M. N. Shadlen. 2010. "Temporal context calibrates interval timing". *Nature Neuroscience* 13 (8): 1020–1026. doi:10.1038/nn.2590.

Kahneman, D., and A. Tversky. 1972. "Subjective probability: A judgment of representativeness". *Cognitive Psychology* 3 (3): 430–454. doi:10.1016/0010-0285(72)90016-3.

Kareev, Y. 1992. "Not that bad after all: Generation of random sequences." *Journal of Experimental Psychology: Human Perception and Performance* 18 (4): 1189–1194. doi:https://doi.org/10.1037/0096-1523.18.4.1189.

Kim, T. D., M. Kabir, and J. I. Gold. 2017. "Coupled decision processes update and maintain saccadic priors in a dynamic environment". *The Journal of Neuroscience* 37 (13): 3632–3645. doi:10.1523/JNEUROSCI.3078-16.2017.

Kingstone, A., and R. M. Klein. 1993. "What are human express saccades?" *Perception & Psychophysics* 54 (2): 260–273. doi:10.3758/BF03211762.

Kuss, M., F. Jäkel, and F. A. Wichmann. 2005. "Bayesian inference for psychometric functions". *Journal of Vision* 5 (5): 8–8. doi:10.1167/5.5.8.

LaPlace, P. 1951. "A Philosophical Essay on Probabilities (English Translation)". *Dover.*

Lee, J., T. R. Darlington, and S. G. Lisberger. 2020. "The neural basis for response latency in a sensory-motor behavior". *Cerebral Cortex* 30 (5): 3055–3073. doi:10.1093/cercor/bhz294.

Lee, J., et al. 2016. "Signal, noise, and variation in neural and sensory-motor latency". *Neuron* 90 (1): 165–176. doi:10.1016/j.neuron.2016.02.012.

Lipton, A., and V. Kaushansky. 2018. "On the first hitting time density of an Ornstein-Uhlenbeck process". *arxiv*.

Lo, C.-C., and X.-J. Wang. 2006. "Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks". *Nature Neuroscience* 9 (7): 956–963. doi:10.1038/nn1722.

Lopes, L. L. 1982. "Doing the impossible: A note on induction and the experience of randomness". Place: US Publisher: American Psychological Association, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 8 (6): 626–636. doi:10.1037/0278-7393.8.6.626.

McDougle, S. D., and A. G. E. Collins. 2021. "Modeling the influence of working memory, reinforcement, and action uncertainty on reaction time and choice during instrumental learning". *Psychonomic Bulletin & Review* 28 (1): 20–39. doi:10.3758/s13423-020-01774-z.

Mookherjee, D., and B. Sopher. 1994. "Learning behaviour in an experimental matching pennies game". *Games and Economic Behaviour* 7 (1).

Mulder, M. J., et al. 2012. "Bias in the brain: a diffusion model analysis of prior probability and potential payoff". *Journal of Neuroscience* 32 (7): 2335–2343. doi:10.1523/JNEUROSCI.4156-11.2012.

Nakahara, H., K. Nakamura, and O. Hikosaka. 2006. "Extended LATER model can account for trial-by-trial variability of both pre- and post-processes". *Neural Networks: The Official Journal of the International Neural Network Society* 19 (8): 1027–1046. doi:10.1016/j.neunet.2006.07.001.

Nash, J. F. 1950. "Equilibrium points in n-person games". *Proceedings of the National Academy of Sciences* 36 (1): 48–49. doi:10.1073/pnas.36.1.48.

Nickerson, R. S. 2002. "The production and perception of randomness". *Psychological Review* 109 (2): 330–357. doi:10.1037/0033-295X.109.2.330.

Nijhawan, R. 2008. "Visual prediction: psychophysics and neurophysiology of compensation for time delays". *The Behavioral and Brain Sciences* 31 (2): 179–198, discussion 198–239. doi:10.1017/S0140525X08003804.

Nobre, A. C., and F. van Ede. 2018. "Anticipated moments: temporal structure in attention". *Nature Reviews Neuroscience* 19 (1): 34–48. doi:10.1038/nrn.2017.141.

Noorani, I. 2014. "LATER models of neural decision behavior in choice tasks". *Frontiers in Integrative Neuroscience* 8. doi:10.3389/fnint.2014.00067.

Pedersen, M. L., M. J. Frank, and G. Biele. 2017. "The drift diffusion model as the choice rule in reinforcement learning". *Psychonomic Bulletin & Review* 24 (4): 1234–1251. doi:10.3758/s13423-016-1199-y.

Purcell, B. A., et al. 2012. "From salience to saccades: mltiple-alternative gated stochastic accumulator model of visual search". *The Journal of Neuroscience* 32 (10): 3433–3446. doi:10.1523/JNEUROSCI.4622-11.2012.

Ratcliff, R., and G. McKoon. 2008. "The diffusion decision model: theory and data for two-choice decision tasks". *Neural computation* 20 (4): 873–922. doi:10.1162/neco.2008.12-06-420.

Ratcliff, R., and P. L. Smith. 2004. "A comparison of sequential sampling models for two-choice reaction time". *Psychological Review* 111 (2): 333–367. doi:10.1037/0033-295X.111.2.333.

Reddi, B. a., and R. H. Carpenter. 2000. "The influence of urgency on decision time." *Nature neuroscience* 3 (8): 827–830. doi:10.1038/77739.

Richards, B. A., and P. W. Frankland. 2017. "The persistence and transience of memory". *Neuron* 94 (6): 1071–1084. doi:10.1016/j.neuron.2017.04.037.

Schall, J. D. 2019. "Accumulators, neurons, and response time". *Trends in Neurosciences* 42 (12): 848–860. doi:10.1016/j.tins.2019.10.001.

Schall, J. D., et al. 2011. "Neural mechanisms of saccade target selection: gated accumulator model of the visual-motor cascade." *The European Journal of Neuroscience* 33 (11): 1991–2002. doi:10.1111/j.1460-9568.2011.07715.x.

Schmolesky, M. T., et al. 1998. "Signal timing across the Macaque visual system". *Journal of Neurophysiology* 79 (6): 3272–3278. doi:10.1152/jn.1998.79.6.3272.

Schütt, H. H., et al. 2016. "Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data". *Vision Research* 122:105–123. doi:10.1016/j.visres.2016.02.002.

Simen, P. 2012. "Evidence accumulator or decision threshold – which cortical mechanism are we observing?" *Frontiers in Psychology* 3. doi:10.3389/fpsyg.2012.00183.

Standage, D., D.-H. Wang, and G. Blohm. 2014. "Neural dynamics implement a flexible decision bound with a fixed firing rate for choice: a model-based hypothesis". *Frontiers in Neuroscience* 8. doi:10.3389/fnins.2014.00318.

Steingroever, H., D. Wabersich, and E.-J. Wagenmakers. 2021. "Modeling across-trial variability in the Wald drift rate parameter". *Behavior Research Methods* 53 (3): 1060–1076. doi:10.3758/s13428-020-01448-7.

Stone, E. 1960. "Models for choice reaction time". *Psychometrika* 25 (3): 251–260. doi:https://doi.org/10.1007/BF02289729.

Taatgen, N., and H. van Rijn. 2011. "Traces of times past: representations of temporal intervals in memory". *Memory & Cognition* 39 (8): 1546–1560. doi:10.3758/s13421-011-0113-0.

Teichert, T., J. Grinband, and V. Ferrera. 2016. "The importance of decision onset". *Journal of Neurophysiology* 115 (2): 643–661. doi:10.1152/jn.00274.2015.

Tune, G. S. 1964. "A brief survey of variables that influence random-generation". *Perceptual and Motor Skills* 18 (3): 705–710. doi:10.2466/pms.1964.18.3.705.

Urai, A. E., et al. 2019. "Choice history biases subsequent evidence accumulation". *eLife* 8. doi:10.7554/eLife.46331.

Usher, M., and J. L. McClelland. 2001. "The time course of perceptual choice: The leaky, competing accumulator model". *Psychological Review* 108 (3): 550–592. doi:10.1037//0033-295X.108.3.550.

Winkel, J., et al. 2014. "Early evidence affects later decisions: Why evidence accumulation is required to explain response time data". *Psychonomic Bulletin & Review* 21 (3): 777–784. doi:10.3758/s13423-013-0551-8.

Wong, K.-F., and X.-J. Wang. 2006. "A Recurrent Network Mechanism of Time Integration in Perceptual Decisions". *Journal of Neuroscience* 26 (4): 1314–1328. doi:10.1523/JNEUROSCI.3733-05.2006.

Methods

Code and data availability

Python code for the gated accumulator model, data analysis, and model fitting can be found on `https://github.com/BJCaie/Gated-Accumulation`

Participants

Six participants took part in the study. All had normal or corrected-to-normal vision. The participant gave written consent prior to undergoing the procedure. The study was conducted with an experimental protocol approved by Health Sciences Research Ethics Board, which adheres to the principles of the Canadian Tri-council Policy Statement on Ethical Conduct for Research Involving Humans and the principles of the Declaration of Helsinki (2013).

Task

The subject began by fixating on a central target projected onto a screen (20" Mitsubishi Diamond Pro [16x12inches] 1280x1024 pixels, 60Hz, contrast 75.2%, brightness 0%). Two targets, one in each visual field, were then presented with randomly assigned temporal asynchrony values (16ms, 33ms, 66ms, 99ms) and onset times drawn from $U(750ms, 1250ms)$. The subject was instructed to look at either target, so long as it was done as fast as possible, without anticipating. The subjects were informed prior to the beginning of the experiment that the targets would come up at different times relative to each other. The subject was instructed to blink in between trials in order to minimize disruption of stimulus perception. The script for the task was written in MATLAB using Psychtoolbox.

Data acquisition and filtering

We collected a total of 80,550 trials from 52 sessions across our 6 participants. During each trial, eye movements were tracked via EyeLink 1000 Tower Mount (SR Research, Mississauga, Canada) with a 1000Hz infrared camera that tracked retinal position. Retinal position was calibrated prior to each session. The EyeLink apparatus was placed 60cm from the screen containing the saccade targets. Saccades were calculated offline using a saccade detection algorithm with a velocity criterion of 50°/s, and were individually verified through a custom visual marking program. Trials where the tracker lost the eye (either through tracking error or participant lapse) were excluded. Trials were also excluded if the eye position deviated more than 15 degrees from the horizontal. The first saccade following target onset, or the first saccade within 3 degrees of the future target position prior to target onset, were taken to calculate the saccadic reaction time and endpoint error. Saccades with a positive endpoint horizontal position were given a choice value of 1, and saccades with a negative endpoint horizontal position a choice value of 0.

Psychometric fitting

All data analysis was performed using custom scripts in Python. Psychometric functions were fit using the Psignifit toolbox with a Bayesian Inference fitting procedure (Kuss, Jäkel, and Wichmann 2005). We fit psychometric functions to a logistic function

$$F(x; m, w) = (1 + exp(-\frac{z(a)}{w}(x - m))^{-1}, z(\alpha) = 2 * log(\frac{1}{\alpha - 1})$$

and a fixed lapse rate of 1%. Choices occuring prior to 70ms (sensory delay cutoff) after target onset were discounted from psychometric fitting. To compare the slopes of the psychometric functions across choice sequences, we bootstrapped the marginal distribution that estimates the slope parameter, and used the

position in the choice sequence as an ordinal variable to fit a linear regression model. The model therefore tests the hypothesis that there is a linear relationship between the slope of the psychometric function and the position in the choice sequence, against the null model that they are independent.

Statistical testing of differences in reaction times

Qualitatively, we observed an increase in specific portions of the distribution of reaction times as a function of choice history and delay intervals. From the perspective of our modelling framework, causal change in the underlying process may effect certain portions of the observed reaction time distribution without influencing typical summary statistics such as the mean or, preferably for heavy-tailed distributions like reaction times, the median. Thus, standard parametric statistics were not appropriate to analyze the effects we observed qualitatively. In order to avoid this issue, we chose to take the standard deviation as a proxy measure of changes to early reaction times (be them anticipatory or express saccades).

The logic for this is as follows: because classical reaction time distributions, with responses filtered before target onset or a cutoff chosen based on estimates of the delay time of the sensory system, are heavy-tailed, the introduction of more early responses will increase the probability density on the other side of the heavy-tail. Thus, an increase in early responses should influence the standard deviation of the distribution. Although this is not a perfect metric, as variance could increase from other processes as well, we took this to be good enough to quantify what is readily observable from visually inspecting the distributions.

To get an estimate of the confidence interval around the standard deviation, we bootstrapped each reaction time distribution 1000 times and computed the 95% confidence interval based on this distribution. We then compared each successive choice sequence and determined if the intervals overlapped. For repetitions, a progressive increase in the mean estimate of standard deviation was observed, with an overlap in the confidence intervals only occurring for repetitions 3-5.

$$R1: 103.05-109.09, R2: 116.98-125.88, R3: 141.04-154.76, R4: 150.05-169.71, R5: 155.92-178.11$$

The change in means was smaller for alternations: while a progressive increase was observed, there was still considerable overlap when comparing a single trial back

$$A1: 100.00-106.44, A2: 102.61-111.23, A3: 108.74-121.53, A4: 110.45-126.26, A5: 120.19-141.20$$

Fitting to free choice saccade data

In the previous section, we described the computation of a single probability density function describing 1. anticipatory reaction times occurring during the delay period and 2. reaction times resulting from evidence accumulated from a free choice saccade target appearing at time $t = y$. However in our dataset, the onset time between trial and target is not fixed, but varies according to a uniform distribution between 750 and 1250ms following the appearance of a fixation cross (*Methods: Task*). Therefore, each reaction time we observe can be computed in two ways: 1. relative to the onset of the target, and 2: relative to beginning of the trial (appearance of the fixation cross).

The first case is how traditional evidence accumulation models are computed, as they typically are not concerned with explaining either behaviour (anticipatory responses) or model correlates (baseline neural activity) that occur in between trials. For our purposes, however, by aligning data to target onset, we lose all information about the relative timing of any anticipatory responses. Take for example a trial where the participant directs a saccade 50ms prior to the onset of a target. Depending on the delay time of the trial

in question, this could correspond to very different points in time relative to the fixation cross onset. So our gated accumulation model with within-trial baseline dynamics is ill-posed to fit data aligned to target onset when the delay period is variable. Thus, we need to align our distributions to fixation onset.

However, this introduces a new problem: for a given choice, while our model computes a probability distribution of reaction times over this delay time, we only have a single observation to fit this model to. We could then recompute this data-limited problem for each individual observation, finding the best model to explain a single observation, but this approach assumes that each observation is representative of an unbiased sample from our best-fitting generative model, and would be computationally expensive (requiring a generative model to be fit to each data observation). Instead, we took an intermediate approach, and binned a given data set into evenly spaced delay time distributions before computing separate fixation-aligned reaction time distributions for each bin. This preserves the timing of the anticipatory response of the data relative to the model, and reduced computational complexity.

We then sought to find the model that best explained *all* of these delay time-binned reaction time distributions as a function of a single parameter set for our gated accumulator model. As such, for model fitting, we generated 10 evenly spaced delay bins between 750 and 1250ms, and took all trials where the true delay time fell between these bins, then realigned the reaction time data to the onset time of that bins first value. For example, all trials with a delay time between 750 and 805 ms were realigned so $t = y = 750$, and all trials between 806 and 861 realigned so $t = y = 806$. Our task for fitting then was to find the best combination of evidence accumulation and baseline parameters that combined to explain these 10 reaction time distributions.

We used differential evolution, a stochastic optimization technique in the *Scipy* optimization toolbox, to fit a parameter set for each delay bin. This algorithm allowed us to search large parameter spaces and was chosen because it generally performed well (early iterations of model code were often stuck in local minima). We used the 'bestlbin' search method, with a maximum of 1000 iterations performed. For boundary conditions on the parameters, we fixed $_0$ to 1, and set

$$\theta : (0.4, 0.6), \kappa : (0.8, 2), \sigma_{OU} : (.05, .1), \mu_r : (2/1000, 6/1000), \sigma_r : (.6/1000, 1/1000) \tag{10}$$

We used a cumulative Kullback-Leibler divergence summed across each delay bin as a metric for the algorithm to minimize. Since reaction time data is discretely sampled, we binned each dataset into 100 bins between the beginning and end of the trial period ($t = 0$ and $t = 2000$) before computing the KL divergence.

$$\sum_{i=1}^{N} D_{\mathrm{KL}}(P_i \| Q_i) \tag{11}$$

$$D_{\mathrm{KL}}(P_i \| Q_i) = \sum_{x \in \mathcal{X}} P(x) \log\left(\frac{P(x)}{Q(x)}\right) * \lambda \tag{12}$$

where i is the delay bin start, generated by N equal spacings between 750 and 1250 ms, $P_i(x)$ is the reaction time model probability for delay bin i and reaction time bin x, $Q_i(x)$ is the same for the data probability distribution, and λ was a constant scaling term to prevent numerical underflow. The discrete model probability was calculated from the full probability density function by sampling 10,0000 times from the continuous model over the same set of 100 discrete reaction time bins as the data histogram was computed from.