

Coupled Decision Processes Update and Maintain Saccadic Priors in a Dynamic Environment

Timothy Doyeon Kim,¹ Mohammad Kabir,² and Joshua I. Gold²

¹Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey 08544, and ²Department of Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Much of what we know about how the brain forms decisions comes from studies of saccadic eye movements. However, saccadic decisions are often studied in isolation, which limits the insights that they can provide about real-world decisions with complex interdependencies. Here, we used a serial reaction time (RT) task to show that prior expectations affect RTs via interdependent, normative decision processes that operate within and across saccades. We found that human subjects performing the task generated saccades that were governed by a rise-to-threshold decision process with a starting point that reflected expected state-dependent transition probabilities. These probabilities depended on decisions about the current state (the correct target) that, under some conditions, required the accumulation of information across saccades. Without additional feedback, this information was provided by each saccadic decision threshold, which represented the total evidence in favor of the chosen target. Therefore, the output of the within-saccade process was used, not only to generate the saccade, but also to provide input to the across-saccade process. This across-saccade process, in turn, helped to set the starting point of the next within-saccade process. These results imply a novel role for functional information-processing loops in optimizing saccade generation in dynamic environments.

Key words: decision-making; eye movement; response times; saccade

Significance Statement

Saccades are the rapid, ballistic eye movements that we make approximately three times every second to scan the visual scene for interesting things to look at. The apparent ease with which we produce saccades belies their computational sophistication, which can be studied quantitatively in the laboratory to provide insights into how our brain manages the interplay between sensory input and motor output. The present work is important because we show for the first time how this interplay operates both within and across saccades to ensure that these eye movements are guided effectively by learned expectations in dynamic environments. More generally, this study shows how sensory-motor decision processes, typically studied in isolation, interact via functional information-processing loops in the brain to produce complex, adaptive behaviors.

Introduction

The generation of goal-directed saccadic eye movements, like other aspects of sensory, cognitive, and motor processing, can be sensitive to expectations (Friston, 2010; O'Reilly, 2013; Seriès and Seitz, 2013; Pezzulo et al., 2015). For example, saccadic reaction

times (RTs) are affected when the expected location of a visual saccade target is cued in advance, such as as in the Posner cueing task (Posner et al., 1980; Posner, 1980), or learned implicitly from the statistical structure of a recent sequence of events, such as in the serial RT task (Nissen and Bullemer, 1987; Reed and Johnson, 1994; Schwarb and Schumacher, 2012). These effects imply a dynamic process that can establish appropriate expectations from one saccade to the next. However, the computational principles that govern this sequential-updating process are not yet fully understood.

To identify these principles, we measured saccadic RT distributions of human subjects performing a dynamic auditory location-discrimination task that required them to look in the direction (left or right) of a virtual sound source that varied probabilistically for sequences of sounds (see Fig. 1A). Our analyses focused on two sets of complementary questions. First, what specific aspect of saccade processing was affected by trial-by-trial, state-dependent expecta-

Received Oct. 3, 2016; revised Jan. 1, 2017; accepted Feb. 1, 2017.

Author contributions: T.D.K. and J.I.G. designed research; T.D.K. and M.K. performed research; T.D.K. analyzed data; T.D.K., M.K., and J.I.G. wrote the paper.

This work was supported by the National Institutes of Health (Grants R01 EY015260 and NSF-NCS 1533623 to J.I.G.) and the University of Pennsylvania (Vagelos Undergraduate Research Grant and University Scholars Grant to T.D.K.). We thank Kenan Saleh for help with data collection and Roger Carpenter, Takahiro Doi, Long Ding, and Kamesh Krishnamurthy for comments.

The authors declare no competing financial interests.

Correspondence should be addressed to Joshua I. Gold, Department of Neuroscience, University of Pennsylvania, D407 Richards Laboratories, Philadelphia, PA 19104-6072. E-mail: jigold@mail.med.upenn.edu.

DOI:10.1523/JNEUROSCI.3078-16.2017

Copyright © 2017 the authors 0270-6474/17/373632-14\$15.00/0

tions? Consistent with previous reports, we found that saccadic RTs were sensitive to state-dependent transition probabilities (in this case, the conditional probabilities that a sound would come from the right or left on the current trial given its location on the previous trials; see inset in Fig. 1A) (Nissen and Bullemer, 1987; Reed and Johnson, 1994; Schwarb and Schumacher, 2012). We extended those results by analyzing the RT distributions using versions of the Linear Approach to Threshold with Ergodic Rate (LATER) model (Carpenter and Williams, 1995; Reddi et al., 2003; Nakahara et al., 2006; Oswal et al., 2007; Noorani and Carpenter, 2016). This kind of “rise-to-threshold” model has been used to analyze saccadic RT data under a broad range of conditions, including how expectations are learned gradually in response to changes in the relative frequencies of occurrence of particular saccade targets (Anderson and Carpenter, 2006; Noorani and Carpenter, 2016). We used it to show that expectations based on learned, state-dependent transition probabilities have consistent effects on the expected (prior) probability of making a particular saccade, but not on how the instructive sensory cue is processed to initiate the saccade (see Fig. 1B).

Second, how were these prior probabilities updated on a trial-by-trial basis? Here, we extended the LATER model to include a dynamic process for updating priors based on normative principles of sequential inference in unpredictable environments (see Fig. 1C) (Glaze et al., 2015). This kind of sophisticated inference process is consistent with how human subjects solve a range of cognitive, perceptual, and motor tasks, but has not yet been related directly to the saccadic system (Courville et al., 2006; Behrens et al., 2007; Nassar et al., 2010; Wilson et al., 2010; Haith and Krakauer, 2013; Payzan-LeNestour et al., 2013; Gallistel et al., 2014; Glaze et al., 2015; Otto et al., 2015). We tested key features of these models, including how estimates of state-dependent transition probabilities are updated according to the strength of evidence provided on each trial about the current state. These analyses show that saccadic priors are updated sequentially according to normative principles that can involve functional information-processing loops that couple within- and across-saccade decision processes (see Fig. 1B,C). These findings provide new insights into, not just how the brain promotes speeded saccadic responses to the most likely target locations in dynamic environments, but also more generally how decision processes, typically studied in isolation, interact to produce complex, adaptive behaviors.

Materials and Methods

Behavioral task

Human subject protocols were approved by the University of Pennsylvania Internal Review Board. Forty-one subjects (15 female and 26 male; age range = 18–42 years) participated in the study after providing informed consent. Data from one subject were excluded from the analyses because of the very high prevalence of short-latency saccades (64.25% of measured RTs were <100 ms). Data from two other subjects were excluded from the analyses because both chose one option almost exclusively (first: 916 leftward choices out of 996 trials; second: 1201 rightward choices out of 1205 trials). Each subject underwent one session of training before participating in one to three experimental sessions.

The task required subjects to indicate whether the virtual source of a sound presented through headphones (Senheiser HD 598 over-ear headphones) was from the left or right side by making a saccadic eye movement to a visual target located on the appropriate side. Subjects were seated and placed their chin in a chin-rest that was ~60 cm from a Tobii T60XL eye tracker and display that sampled the position of both eyes at 60 Hz (see Fig. 1A).

The sounds were obtained from the Listen HRTF database (Institute for Music/Acoustic Research and Coordination). The sound presented on each trial was a set of band-pass-filtered pulses of white noise with a 10 ms gap between the pulses. The duration of each pulse was 50 ms, with

5 ms cosine ramps at the beginning and the end of the pulse. The sounds were played via headphones. The sounds varied in terms of their simulated azimuthal (horizontal) source location, picked randomly on a trial-by-trial basis from 50 different versions of that location chosen from the HRTF database.

For each trial, a fixation dot appeared on the screen, which the subject had to look at for 600 ms without blinking before the sound stimulus was presented. The fixation dot then disappeared and the sound stimulus was played through the headphones. Simultaneous to sound onset, two targets appeared, one ~14° directly to the left of fixation and the other ~14° directly to the right of fixation. The subject was instructed to then make a saccadic eye movement to the target in the perceived direction of the virtual sound source. RT was defined as the time from sound onset (which had a reliable, ~3 ms latency from the time of the software trigger to the time the sound was discernible from the headphones) to the time when the first sampled eye position was >9.3° horizontally from the fixation dot (velocity- and acceleration-based measures yielded nearly identical results). A counter indicating how many trials subjects had done so far was always shown at the bottom of the screen between trials. Each correct choice was worth one point. Subjects were paid based on how many points they earned during the given session. Each session lasted ~40–60 min.

We tested how saccadic RTs and choice behavior on this task were affected by certain combinations of the following four factors: (1) left–right transition probability, (2) stimulus strength, (3) trial-by-trial feedback, and (4) speed–accuracy instructions. The particular combinations that we tested are described below.

Left–right transition probability governed the probability that the simulated sound source was on the left or right side on a given trial given its location on the previous trial. Specifically, sound source location was generated using a Markov chain with two states, C_L (left side) and C_R (right side), and four transition probabilities, H_{LR} (computed as the expected number of transitions per trial, or “generative hazard rate,” corresponding to a sound presented from the left side followed by a sound from the right side), $H_{LL} = 1 - H_{LR}$, H_{RL} , and $H_{RR} = 1 - H_{RL}$. To assess sensitivity to different transition probabilities, we used a range of values (0.15, 0.2, 0.25, 0.35, 0.65, and 0.85) in different combinations. Within a given session, H_{LR} and H_{RL} each took on one such value for the entire session, with $H_{LR} \neq H_{RL}$ in some sessions and $H_{LR} = H_{RL}$ in others (see below). Each trial consisted of the presentation of a single sound, the location of which depended on the current state. State transitions occurred between trials.

Stimulus strength was held constant within a given session at a level defined as either “strong” (S) or “weak” (W). These two levels differed in terms of their left–right discriminability based on their simulated locations determined by the HRTF database. The strong stimuli were always defined as simulated sound sources located directly to the left or right (i.e., $\pm 90^\circ$) from center. The weak stimuli were determined separately for each subject, before performing the main task, using a modified version of the adaptive QUEST procedure (Watson and Pelli, 1983; King-Smith et al., 1994). Using this procedure, the simulated azimuth, in degrees, of the weak stimuli was set to the mean of the threshold probability distribution yielding 70% correct responses in an unbiased environment (i.e., each sound presentation was equally likely to come from the left or right side). Typically, this procedure yielded simulated azimuthal locations of ~3–4° from center.

Trial-by-trial feedback was either given (F) or not (NF) for each full session. When feedback was given, after a correct choice, the correct target turned green and a circle appeared around the target. After an error, the incorrectly chosen target turned red and a circle appeared around the correct target. For these sessions, a second counter at the bottom of the screen showed the running tally of correct trials.

In regard to speed–accuracy instructions, for most sessions, the subjects were instructed to choose as accurately and quickly as possible. For these sessions, the subjects had to earn as many points as possible for a set number of trials. There was no time limit for these sessions. We typically gave 1200 trials for each subject. However, five subjects indicated during the training session that they felt fatigued before finishing, so they were given either 600 trials (a single session with weak stimuli and no feed-

back) or 1000 trials (one session with weak stimuli and no feedback and four sessions with weak stimuli and feedback). Subjects were paid based on how many points they earned during the given session. For the remaining sessions, the subjects were instructed to emphasize speed (“respond as quickly as possible”). We used this manipulation to test whether such an instruction, which is known to reduce the total amount of evidence accumulated to trigger the saccade (equivalent to lowering the decision threshold in Fig. 1B) (Reddi and Carpenter, 2000), also affects the evidence accumulated across trials to establish appropriate priors (see Fig. 1C). In these sessions, subjects had to earn as many points as possible in 40 min. There was no limit on how many trials the subject could complete during that time (minimum = 1092, median = 1452, maximum = 1738 total trials completed per session).

We use the following labels in the text to refer to the particular task conditions that we tested, corresponding to combinations of the four factors described above.

Strong stimuli (S) included sessions with: (1) equal, low generative transition probabilities ($H_{LR} = H_{RL} = 0.2$ for 5 sessions, = 0.25 for 3 sessions); (2) unequal, low generative transition probabilities ($H_{LR} = 0.35, H_{RL} = 0.15$ for 6 sessions); or (3) unequal, high generative transition probabilities ($H_{LR} = 0.65, H_{RL} = 0.85$ for 7 sessions). Feedback was provided in all of these sessions except two ($H_{LR} = H_{RL} = 0.2$). Those two sessions were included in this group because performance was high (96.9% and 99.0% correct responses) and comparable to when feedback was provided. The subjects were instructed to choose as accurately and quickly as possible. These 21 sessions were conducted using 21 different subjects.

Weak stimuli with feedback (WF) included instructions to choose as accurately and quickly as possible; 10 sessions with $H_{LR} = H_{RL} = 0.2$ and 2 sessions with $H_{LR} = H_{RL} = 0.25$. These 12 sessions were conducted using 12 different subjects. Of these, three also participated in the S condition, and all 12 also participated in the weak, no feedback (WNF) condition.

WNF stimuli included instructions to choose as accurately and quickly as possible; 10 sessions with $H_{LR} = H_{RL} = 0.2$ and 3 sessions with $H_{LR} = H_{RL} = 0.25$. These 13 sessions were conducted using 13 different subjects. Of these, six also participated in the S condition, and 12 also participated in the WF condition.

WNF stimuli with instructions to emphasize speed (WNF_{sp}) included 10 sessions with $H_{LR} = H_{RL} = 0.2$. These 10 sessions were conducted using 10 different subjects, none of whom participated in any of the other conditions.

LATER analyses

All model-fitting procedures described below were implemented by finding the minimum of the appropriate objective function using the Optimization Toolbox from MATLAB (RRID:SCR_001622).

Basic LATER model

The LATER model describes saccadic RT as the time it takes a linear variable (r) to progress from a starting value (S_0) to a threshold value (S_T) (Carpenter and Williams, 1995; Reddi et al., 2003; Noorani and Carpenter, 2016). Assuming that the rate-of-rise r varies from trial-to-trial with a Gaussian distribution, $r \sim N(\mu_r, \sigma_r)$, RT is distributed as a reciprocal of Gaussian with a heavy tail toward longer values: $(S_T - S_0)/r = \Delta S/r$. We used a basic model that treats σ_r as a scale factor set to one and had two independent parameters: (1) the mean rate of rise (μ_r) and (2) the difference between the starting point and the threshold level (ΔS).

Consistent with previous applications, we always fit the basic LATER model only to RT data from correct choices because the simple version of the model that we used does not account for errors (Carpenter and Williams, 1995; Reddi et al., 2003). For conditions with strong stimuli, this approach was justified because errors were rare. For conditions with weak stimuli, errors were more common, so we also fit the data to a modified model that accounted for errors using two competing processes (see “LATER model with dual accumulators” section below).

Best-fitting parameters of the LATER model were first found by minimizing an objective (likelihood) function M that computed the negative mean of the logarithm of the conditional probabilities of obtaining the

RT data given the LATER model, using three free parameters (μ_r , ΔS , and θ) as follows:

$$M = -\frac{1}{N} \sum_i \log \left[\phi \left(\frac{\Delta S}{RT'_{correct,i}} - \mu_r \right) \right] \tag{1}$$

for all $RT'_{correct,i} > \theta$

where ϕ is the standard normal probability density function, μ_r and ΔS are parameters of the LATER model as described above, N is the number of trials used in the fit, and $RT'_{correct,i}$ is the measured RT on trial i of N in the given session that was from a correct trial and had a value $> \theta$. The parameter θ was used as a criterion for deciding whether a given RT was generated by a separate, fast (e.g., express) process (Fischer and Ramspurger, 1984; Carpenter and Williams, 1995; Reddi et al., 2003). The value of θ was constrained to be between 0 and 200 ms. However, if the best-fitting value under these constraints was > 190 ms, then the model was refit with the upper bound of θ set to the fifth percentile of the RT data used for the given fit, which improved the estimates near 200 ms. The best-fitting value of θ from these fits of the basic LATER model were used to define the set of RTs from the given session that were used for all subsequent fits. Note that for Equation 1, we used the mean instead of the sum of the log-likelihood because the free parameter θ can affect the number of data points (N in Eq. 1) used in each iteration of the fitting procedure (for model comparisons, we multiplied this value by N to obtain the total log-likelihood).

We determined within-subject dependencies of the two basic LATER parameters (ΔS or μ_r) on two sets of task conditions, as follows. In each case, we fit nested models to the set of all $RT'_{correct}$ data from the given session (corresponding to a single subject), as defined by the fits to Equation 1 (i.e., with all the fast saccades removed).

The first set of conditions tested whether or not best-fitting values of ΔS or μ_r depended linearly on the logarithm of the state-dependent, generative transition probability ($H_{C-1,C0} = H_{LR}, H_{LL}, H_{RL},$ and H_{RR}). We used the following two nested models: (1) a four-parameter model in which both ΔS and μ_r were expressed as linear functions of $\log(H_{C-1,C0})$: $\Delta S = \beta_{\Delta S,0} + \beta_{\Delta S,1} * \log(H_{C-1,C0})$ and $\mu_r = \beta_{\mu_r,0} + \beta_{\mu_r,1} * \log(H_{C-1,C0})$, where $\beta_{\Delta S,0}, \beta_{\Delta S,1}, \beta_{\mu_r,0},$ and $\beta_{\mu_r,1}$ were free parameters; and (2) a three-parameter model in which the parameter of interest (ΔS or μ_r) was a single free parameter across conditions and the other parameter was expressed as a linear function of $\log(H_{C-1,C0})$, as above.

The second set of conditions tested whether best-fitting values of ΔS or μ_r depended on the number of trials after a change point. We used the following two nested models, separately for 0, 1, 2, and 3 trials after a change point (denoted t_i ; in addition, t_{4+} denotes 4 or more trials after a change point, corresponding to the steady-state value of the parameter): (1) a four-parameter model, with $\Delta S_{t_i}, \mu_{r,t_i}, \Delta S_{t_{4+}}, \mu_{r,t_{4+}}$ and (2) a three-parameter model in which the parameter of interest (ΔS or μ_r) was a single, free parameter for both t_i and t_{4+} .

For both sets of within-subject comparisons, significance was assessed via a likelihood-ratio test ($D =$ twice the difference in the log-likelihoods of the two fits, with a difference in the degrees of freedom equal to one, and compared with a χ^2 distribution). A significant effect implies that the data were better fit by the first model, in which the given parameter varied as a function of task condition. We also assessed across-subject effects using a Kruskal–Wallis test, with transition probability or trials after a change point as the grouping variable.

LATER model with dual accumulators

Because the basic LATER model does not account for errors, we also used a modified LATER model based on two competing accumulators, one for correct responses and the other for error responses (Reddi et al., 2003). This model assumes that the correct and error processes race toward their respective boundaries on each trial, with the winner determining the choice and RT on that trial. We used this model to fit data from the S, WF, and WNF conditions (see Fig. 3B–E).

This model had five parameters to replace μ_r and ΔS from the basic LATER model: (1) mean rate of rise for correct choices, $\mu_{r,correct}$ (constrained to be > 0), (2) mean rate of rise for the incorrect choices, $-\mu_{r,error}$ (constrained to be < 0), (3) starting point S_0 , (4) threshold mag-

nitude S_T ($+S_T$ for correct choices, $-S_T$ for incorrect choices), and (5) a parameter ε that can add a delay to correct or error RTs relative to the time of boundary crossings in the model. Specifically, we computed the probability of a correct response as follows:

$$p(\text{correct}) = p\left(\frac{1}{RT'_{\text{correct}}} - \frac{1}{RT'_{\text{error}}} > \varepsilon\right) = 1 - \Phi\left(\frac{\varepsilon - (\mu_{r,\text{correct}} \times \Delta S_{\text{correct}}^{-1} - \mu_{r,\text{error}} \times \Delta S_{\text{error}}^{-1})}{\sqrt{\Delta S_{\text{correct}}^{-2} + \Delta S_{\text{error}}^{-2}}}\right) \quad (2)$$

Where ϕ is the cumulative distribution function of the standard normal distribution, $\Delta S_{\text{correct}} = S_T - S_0$, $\Delta S_{\text{error}} = S_T + S_0$, and RT' represents data from non-short-latency saccades, as described above. The parameter ε is equivalent to a criterion bias in signal detection theory in evaluating the difference distribution, $1/RT'_{\text{correct}} - 1/RT'_{\text{error}}$ (Green and Swets, 1966). The value of ε was >0 in all of our fits, which implies that an accumulator with mean $\mu_{r,\text{correct}}$ and an accumulator with mean $-\mu_{r,\text{error}} + \varepsilon \Delta S_{\text{error}}$ race independently toward bounds S_T and $-S_T$, respectively. If S_T is reached first, then a correct choice is made with the RT determined by the time of bound crossing. However, if $-S_T$ is reached first, then an error choice is made and the RT is delayed by the amount $D = \frac{\Delta S_{\text{error}}}{r_{\text{error}}} - \frac{\Delta S_{\text{error}}}{r_{\text{error}} + \varepsilon \Delta S_{\text{error}}}$, where $r_{\text{error}} \sim N(\mu_{r,\text{error}}, \sigma_r)$ is the particular instantiation of the rate of rise on that trial. Therefore, the delay increases as the rate of rise decreases. We also tested a four-parameter model in which the bias parameter ε was set to zero. The four-parameter model captures the RT behavior well, but tended not to capture the choice behavior as well as the five-parameter model (likelihood-ratio test, $p < 0.01$ for all subjects who underwent weak-stimuli conditions).

We fit this model by minimizing the following objective (likelihood) function M that computed the negative sum of the logarithm of the conditional probabilities of obtaining the RT data given the model, using the five parameters describing the dual accumulators:

$$M = -\sum_i \log \left[p(\text{correct}) \times \phi\left(\frac{\Delta S_{\text{correct}}}{RT'_{\text{correct},i}} - \mu_{r,\text{correct}}\right) \right] + -\sum_j \log \left[p(\text{error}) \times \phi\left(\frac{\Delta S_{\text{error}}}{RT'_{\text{error},j}} - \mu_{r,\text{error}}\right) \right], \quad \text{for all } RT'_{\text{correct}} > \theta_{\text{correct}} \text{ and } RT'_{\text{error}} > \theta_{\text{error}} \quad (3)$$

where $p(\text{correct})$ is from Equation 2; $p(\text{error}) = 1 - p(\text{correct})$; and $RT'_{\text{correct},i}$ and $RT'_{\text{error},j}$ are RT data obtained from the i^{th} correct trial and j^{th} error trial, respectively, from the given session and had values $>\theta$. The values of θ_{correct} and θ_{error} were obtained for each session by fits of the same RT data from correct and error trials, respectively, to Equation 1 to remove short-latency values. We assessed within- and across-subject effects as described above for the basic LATER model.

Adaptive LATER model

To better understand the principles that guided the accumulation process reflected in the saccadic priors across trials, we fit the data to a modified, adaptive LATER model. In this model, the starting value S_0 represents the prior that evolves over time according to normative principles (Glaze et al., 2015). Specifically, the log-prior-odds of the sound being on the right versus left side on trial n (ψ_n) are a function of both the log-posterior probabilities from the previous trial ($q_{R,n-1}$ and $q_{L,n-1} = 1 - q_{R,n-1}$ for right and left, respectively) and the transition probabilities H_{RL} and H_{LR} (Bishop, 2006; Glaze et al., 2015) as follows:

$$\psi_n = \log \left[\frac{(1 - H_{RL})q_{R,n-1} + H_{LR}q_{L,n-1}}{H_{RL}q_{R,n-1} + (1 - H_{LR})q_{L,n-1}} \right] \quad (4)$$

The log-posterior-odds, $L_n = \log(q_{R,n}/q_{L,n}) = \log[q_{R,n}/(1 - q_{R,n-1})]$, are updated on each trial as the sum of the log-prior-odds and the log-likelihood ratio provided by the sensory evidence on that trial (LLR_n ,

which is governed by α , a free parameter associated with the given stimulus strength) as follows:

$$L_n = \psi_n + LLR_n \quad (5)$$

where $LLR_n = -\text{logit}(\alpha)$ for left sources, $+\text{logit}(\alpha)$ for right sources.

We used this process to govern trial-by-trial dynamics of the saccadic prior in the LATER model and then fit this adaptive model to the RT data as described for Equation 1, above. This adaptive model had either six or eight total free parameters, depending on the number of unique, state-dependent, generative transition probabilities (H ; by contrast, the basic LATER model fit to the same data had three free parameters per value of H): either two or four parameters ($\mu_{r,H}$) representing the mean rate of rise per value of H ; three parameters governing the starting value of the LATER process, $S_0 = \beta_n = f(\hat{H}_{RL}, \hat{H}_{LR}, \alpha)$ from Eqs. 4 and 5, where \hat{H}_{RL} and \hat{H}_{LR} are the subjective estimates of the respective objective values for that session; and one parameter representing the ending value of the LATER process, S_T , such that ΔS from Equation 1 = $S_T - S_0$. Note that both μ_r and α reflect stimulus strength, but of the two, only α is used to inform the prior on the subsequent trial and thus is also potentially sensitive to post-saccade feedback. Moreover, by leaving \hat{H}_{RL} and \hat{H}_{LR} as free parameters, this model tests the extent to which subjects can learn and use subjective estimates of the objective, latent values to govern their saccadic behavior, as has been shown for certain choice tasks (Glaze et al., 2015). Finally, we emphasize that this parameterization assumes that the across-trial decision affects the starting point of the LATER process (S_0), but ultimately its effect on saccade generation is via ΔS and, therefore, it could in principle be equivalently applied to the ending value of the LATER process (S_T).

We fit the adaptive LATER model by minimizing the objective (likelihood) function M that computed the negative sum of the logarithm of the conditional probabilities of obtaining the RT data given the model while respecting the sequence of N trials, defined as follows:

$$M = \sum_n h\left(\frac{1}{RT'_{\text{correct},n}}\right) \text{ for } n = 1:N, \quad \text{where } h(x_n) = -\log[\phi(\Delta S_n \times x_n - \mu_{r,n})] \text{ if } RT'_{\text{correct},n} > \theta$$

$$h(x_n) = 0 \text{ otherwise (i.e., an error or short-latency saccade trial).} \quad (6)$$

Here, we used θ obtained per session by fits of the same RT data from correct trials to Equation 1 to remove short-latency values, as described above.

We fit both the basic and the adaptive LATER model to data from each individual session, including all tested conditions (S, WF, WNF, and WNF_{sp}). As noted above, some subjects participated in more than one session and were tested using different combinations of factors in each session. However, participation across sessions by individual subjects was not uniform across conditions (e.g., there was strong overlap in the subjects who participated in the WF and WNF conditions, but not otherwise). We therefore analyzed differences in model fits and parameter estimates across conditions on the group level, which included both within- and across-subject effects.

Integration time index

To quantify the gradual, across-trial updates of saccadic prior probabilities described by the adaptive model fits, we computed an integration time index. We computed the integration time index from each session as the normalized area under the curve of ΔS from the adaptive model fits as a function of trials 1–3 after a change point (see Fig. 3B, C). Specifically, we computed $E[\Delta S_{c,n} - E[\Delta S_{c,m}]]$, where $E[\]$ indicates the expected (mean) value across all values of c and n , c indicates a left or right choice, n is 1–3 trials after a change point, and m is 4 or more trials after a change point. The value of this index is zero if ΔS changes abruptly, larger if ΔS changes gradually after a change point.

Results

We measured saccadic RTs of human subjects performing a dynamic auditory discrimination task. Subjects were required to

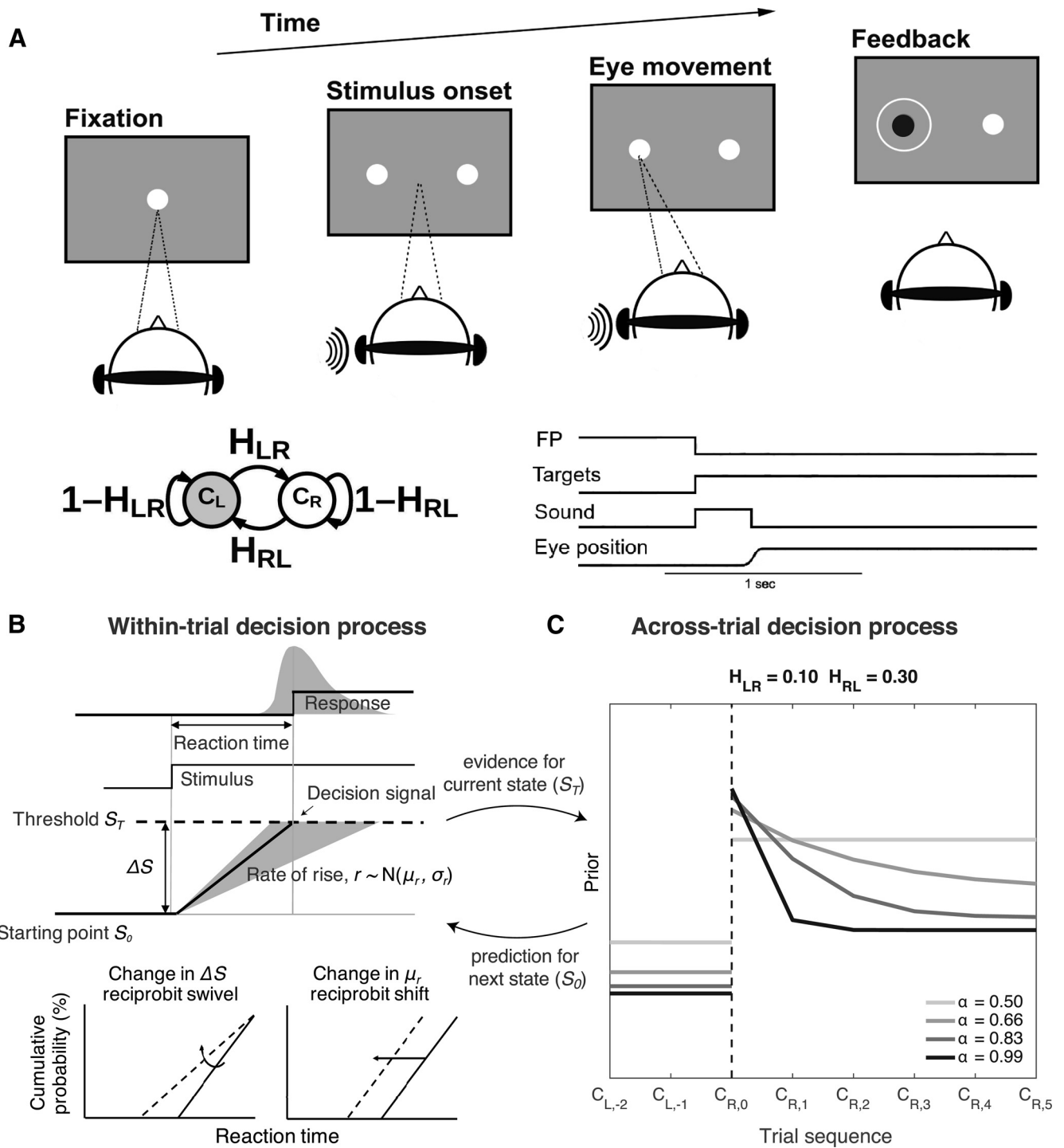


Figure 1. Task and models. **A**, Task schematic and the timing of task events on a single trial. The task required the subject to make a saccadic eye movement to a visual choice target in the direction of the perceived sound (left or right). We used fixed generative hazard rates (H) to control transition probabilities between the two generative states (C_L or C_R , indicating that the sound is generated from the left or right, respectively), as shown on the left. **B**, Basic LATER model, in which saccadic RT distributions are determined by a linear accumulating process with trial-to-trial variability in its rate of rise and predetermined starting (S_0) and ending (S_T) points. Changes in the difference between these points (ΔS) are sensitive to prior expectations, whereas changes in the average rate of rise (μ_r) are sensitive to stimulus difficulty. These effects show up as a “swivel” or “shift” of RT distributions plotted on a reciprobait scale, respectively (Carpenter and Williams, 1995; Reddi et al., 2003). **C**, Predictions of the adaptive LATER model for adaptive changes in priors (specifically, the log-prior-odds of the sound being on the right vs left side, as defined in Eq. 4) as a function of trials relative to a change point from C_L to C_R (the additional, numerical subscripts indicate the trial number relative to the change point at trial 0). According to normative theory, prior expectations should be sensitive to expected transition probabilities (compare values before and after the change point, indicated by the vertical dashed line) and the weight of evidence for the current state (compare the rate of change of values as a function of trials after the change point for the different values of α from Eq. 5) (Glaze et al., 2015). Our findings indicate bidirectional coupling between these within-saccade (**B**) and across-saccade (**C**) processes, such that the total amount of evidence accumulated within a trial to trigger a saccade (the saccade-generation threshold, S_T) determines the weight of evidence for the current state (α) that is accumulated across trials to govern changes in priors that affect S_0 .

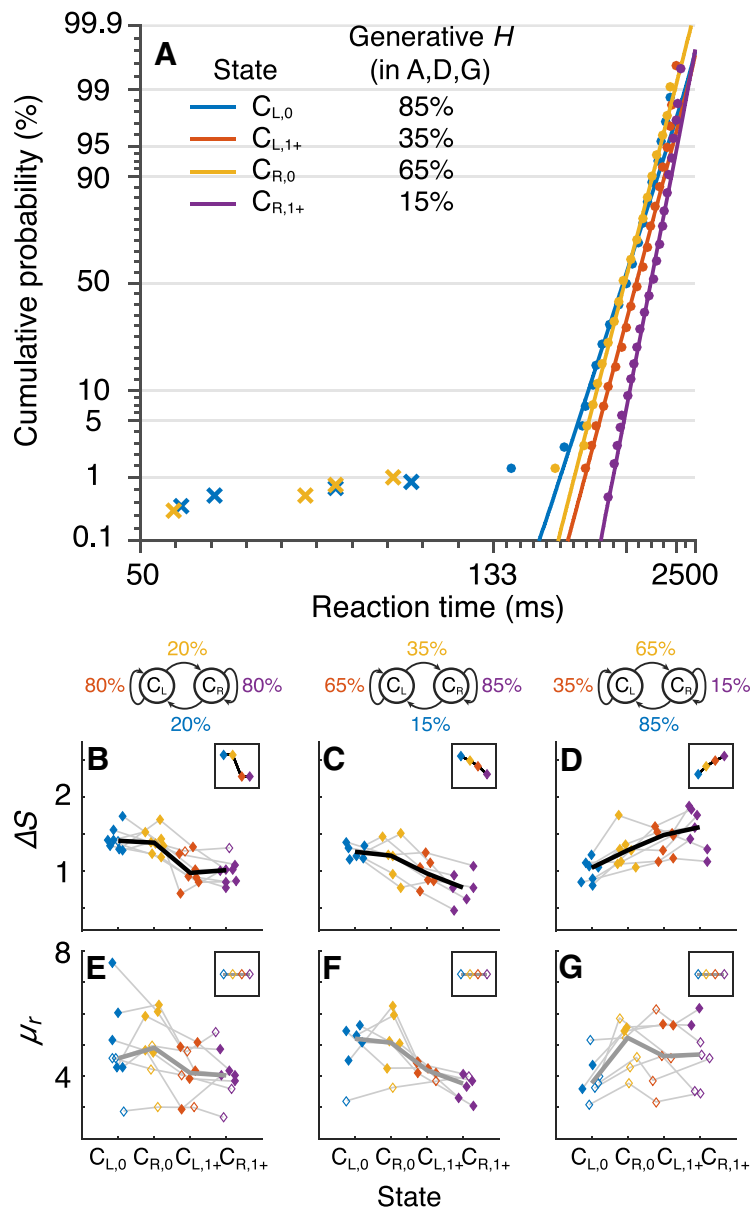


Figure 2. Saccadic priors reflect state-dependent transition probabilities. **A**, Example saccadic RT distributions from a single subject (points) plotted on a reciprob scale: the abscissa shows reciprocal RT and the ordinate shows cumulative probability on a probit scale (Noorani and Carpenter, 2016). Data are separated by color, as indicated, according to the current state, conditionalized on the previous state: $C_{L,0}/C_{R,0}$ is a correct left/right choice on a change point (switch) trial, and $C_{L,1+}/C_{R,1+}$ is a correct left/right choice one or more trials after a change point (stay). These conditionalized states are associated with different transition probabilities from the generative process, as indicated. Lines are fits of the basic LATER model, computed separately for each condition, to regular saccades (circles) but not short-latency saccades (crosses). **B–G**, Best-fitting values from the basic LATER model of ΔS (**B–D**) and μ_r (**E–G**) per condition (abscissa) for different generative transition probabilities, as indicated at the top of each column. Colors are associated with the four conditionalized states, as in **A**. Points are filled if, for the given subject, the given parameter depended on the appropriate transition probability [likelihood-ratio test indicating that a model in which the given parameter was a linear function of $\log(\text{transition probability})$ outperformed a model in which the given parameter was a constant across transition probabilities, $p < 0.01$]. The thin lines connect data from a single session. The thick lines connect the median values across subjects and are shown in black if the medians differed as a function of generative transition probability (Kruskal–Wallis test, $p < 0.01$) and in gray otherwise. The insets show predictions of a LATER model in which prior probability (ΔS) but not rate-of-rise (μ_r) depends on the appropriate transition probability. All fit RT data were obtained from correct trials only in the strong-stimulus (S) condition.

make a saccadic eye movement to look at a visual target in the right or left visual field in response to a brief noise burst played through headphones from a virtual sound source located to the right or left of central gaze, respectively (Fig. 1A). Sound source location was generated using a Markov chain with two states, C_L

(left) and C_R (right), and four generative transition probabilities, H_{LR} (the probability of a sound presented from the left side followed by a sound from the right side), $H_{LL} = 1 - H_{LR}$, H_{RL} , and $H_{RR} = 1 - H_{RL}$. Each trial consisted of the presentation of a single sound, the location of which depended on the current state. State transitions occurred between trials.

Saccadic priors reflected state-dependent transition probabilities

Previous studies that analyzed saccadic RT data using the LATER model used a state- (history-) independent process that is equivalent to $H_{LR} + H_{RL} = 1$; that is, the probability that the location of the sound is to the left is $H_L = H_{RL} = H_{LL}$ and to the right is $H_R = H_{LR} = H_{RR}$ regardless of the outcome of the previous trial (Carpenter and Williams, 1995; Reddi et al., 2003). To study how expectations affect saccadic processing in more structured environments, we used state-dependent transition processes in which the probability that the sound is to the left or right depends on its prior location; i.e., $H_{LR} + H_{RL} \neq 1$. For the first set of tests, we also used strong, easily discriminable stimuli [the median (interquartile, or IQR) error rate per session was 2.56% (range = 2.11–3.91%)] and feedback on each trial to minimize uncertainty about the previous state.

Under these conditions, we found that saccadic priors were updated dynamically according to the appropriate, state-dependent transition probabilities (H ; Fig. 2). Specifically, we tested three sets of conditions: (1) relatively low, equal values of H_{LR} and H_{RL} (0.2 or 0.25), implying a stable environment in which target locations tended to repeat (Fig. 2B,E); (2) relatively low, unequal values of H_{LR} (0.35) and H_{RL} (0.15), implying a more stable environment following right versus left locations (Fig. 2C,F); and (3) relatively high, unequal values of H_{LR} (0.65) and H_{RL} (0.85), implying an unstable environment in which target locations tended to switch slightly more after right versus left locations (Fig. 2A,D,G). Figure 2A shows example RT distributions for trials associated with the high, unequal values of H_{LR} (0.65) and H_{RL} (0.85) on reciprob axes, with reciprocal RT on the x-axis and percentage cumulative frequency on a probit scale on the y-axis. Consistent with the assumptions of the LATER model (Fig. 1B), each distribution plotted in this way followed an approximately straight line except for the small percentage of very short-latency saccades that likely reflect an express saccade or other fast process (Fischer and Ramsperger, 1984; Carpenter, 2012; Noorani and Carpenter,

each distribution plotted in this way followed an approximately straight line except for the small percentage of very short-latency saccades that likely reflect an express saccade or other fast process (Fischer and Ramsperger, 1984; Carpenter, 2012; Noorani and Carpenter,

2016). Critically, the reciprob lines tended to show a qualitative “swivel” as a function of the different transition probabilities, consistent with changes in priors (Fig. 1B) (Carpenter and Williams, 1995).

To quantify these effects, we fit the basic LATER model to the interleaved conditions associated with each state transition. These fits yielded changes in the height of the rise-to-threshold process (ΔS in Fig. 1B and Eq. 1) that depended appropriately on the state-dependent transition probability when measured both for individual subjects (filled circles in Fig. 2B–D; likelihood-ratio test, $p < 0.01$ in all cases except one subject in Fig. 2B) and median values across subjects (Kruskal–Wallis test, $p < 0.01$ in all three cases). In contrast, best-fitting values of the mean rate of rise of the rise-to-threshold process (μ_r in Fig. 1B and Eq. 1), which is typically interpreted in terms of the strength of the stimulus instructing the saccadic eye movement, were not affected as consistently by these manipulations: data from individual subjects showed some dependencies on the generative transition probability (possibly reflecting a priming effect thought to modulate μ_r under some conditions; Cho et al., 2002; Goldfarb et al., 2012), but median values across subjects did not (Figs. 2E–G; $p > 0.01$ in all three cases). Therefore, saccadic RTs reflected learned expectations about state-dependent transition probabilities via changes in the height of a rise-to-threshold saccade generation process, representing a form of prior probability of making a particular saccade.

Saccadic priors reflected the accumulation of uncertain evidence over trials

The data shown in Figure 2 correspond to conditions in which the auditory stimuli were easily identifiable as coming from the left or right side and were followed by explicit visual feedback about the correct location. Under those conditions, the expected probabilities of a left or right choice on the current trial depended on the previous state (i.e., whether the sound on the previous trial came from the left or the right), but not the states before that. Here, we used less-discriminable stimuli and no feedback (WNF; see Fig. 5A). Under these conditions, there was uncertainty about the previous state. According to normative theory, this uncertainty should lead to more gradual updating across trials, representing a form of evidence accumulation (Fig. 1C and Eq. 4 and 5) (Glaze et al., 2015).

We found that, consistent with normative theory, saccadic priors changed gradually over multiple trials in the presence of state uncertainty (Fig. 3). Specifically, when weak stimuli were used and no feedback was given, the best-fitting value of ΔS from the LATER model was largest just after a change point, then decreased gradually over several trials (Fig. 3C, cf. dark gray curves in Fig. 1C). These effects were not evident when using strong stimuli, which reliably indicated the current state and therefore corresponded to abrupt and complete adjustments to ΔS after a change point (Fig. 3B). The gradual updates of ΔS in the presence of state uncertainty were not evident for μ_r , which, on average, stayed relatively constant across trials for the three task conditions (Fig. 3D,E).

One potential concern about these fits is that the basic LATER model does not account for the errors that occurred in the weak stimulus condition, which were therefore not included in the fits (Carpenter and Williams, 1995; Reddi et al., 2003). To address this concern, we fit the full datasets from the WNF condition to an extended model that used competing accumulators to account for both correct and error choices (Reddi et al., 2003). This model captured empirical error rates (Pearson’s r comparing empirical and simulated error rates per session was 0.99, $p < 0.0001$) via a compet-

ing process that tended to have equal or smaller best-fitting values of ΔS and μ_r than those obtained from correct trials for both repeated (ΔS : median value = 0.97 for error trials vs 0.88 for correct trials, Wilcoxon signed-rank test, $p = 0.48$; μ_r : 2.47 vs 2.85, respectively, $p < 0.001$) and switched choices (ΔS : 0.77 vs 1.35, respectively, $p < 0.00001$; μ_r : 2.54 vs 3.01, respectively, $p = 0.0034$). These fits imply that the subjects were using consistent strategies throughout these conditions, and errors tended to occur when the noisy, trial-by-trial evidence was particularly weak. Critically, the best-fitting values of ΔS and μ_r for correct trials from this model were nearly identical to those obtained from the basic LATER fits to correct trials only (Pearson’s r comparing best-fitting values of both μ_r and ΔS from the two models per session was 0.99, $p < 0.0001$). Accordingly, the effects of trial history using these fits were also nearly identical to those reported above: ΔS , but not μ_r , for correct choices was adjusted gradually after a change point on trials with weak stimuli and no feedback (Fig. 3B–E, magenta symbols; note that there were too few error trials to conduct a similar analysis to identify their dependence on trial history).

Saccadic priors were updated according to predictions of a normative model

The preceding analyses demonstrated that changes in saccadic prior probabilities affected the total amount of evidence accumulated per saccade (ΔS in the basic LATER model) and not the rate of rise of the decision variable (μ_r in the basic LATER model), which is consistent with previous findings (Carpenter and Williams, 1995). To better understand the principles that guided how the saccadic priors were updated across trials, we fit the data to a modified, adaptive LATER model that describes such updates according to normative principles (Glaze et al., 2015). According to this model, the starting point S_0 represents the prior that is updated on each trial based on three parameters: the subject’s estimates of H_{LR} and H_{RL} and a parameter α that represents the weight of evidence provided by the stimulus about the current state and thus strongly governs the time course of prior updates, particularly after change points (Fig. 1C and Eqs. 4–5). The model included several other parameters as in the basic LATER model: one decision threshold, S_T , plus separate values of μ_r for each generative transition probability (models using a single value of μ_r across conditions yielded similar results). Unlike the basic model, which accounted for differences in ΔS as a function of generative transition probability (Fig. 2) or trials after change point for weak stimuli (Fig. 3) only by explicitly using separate parameters for each such condition, a single version of the adaptive model was sufficiently flexible to account for all of these effects (Fig. 4).

We compared fits to the adaptive and basic models for data collected under several different conditions that were likely to affect how priors were updated across trials. These conditions, which led to different patterns of accuracy (Fig. 5A) and RTs (Fig. 5B) for each saccadic decision, were as follows: (1) S stimuli, corresponding to the data presented in Fig. 2, which tended to have the highest accuracy and short RTs; (2) WF stimuli, which tended to have slightly lower accuracy and longer RTs; (3) WNF stimuli, corresponding to the data presented in Fig. 3C,E, which tended to have slightly lower accuracy and shorter RTs than in the WF condition; and (4) WNF_{sp} stimuli, which tended to have the lowest accuracy and short RTs.

The adaptive model tended to provide better fits to the data than the basic models, particularly when weak stimuli were used and the subjects were not emphasizing speed (the model comparisons were quantified using the Bayesian information criterion,

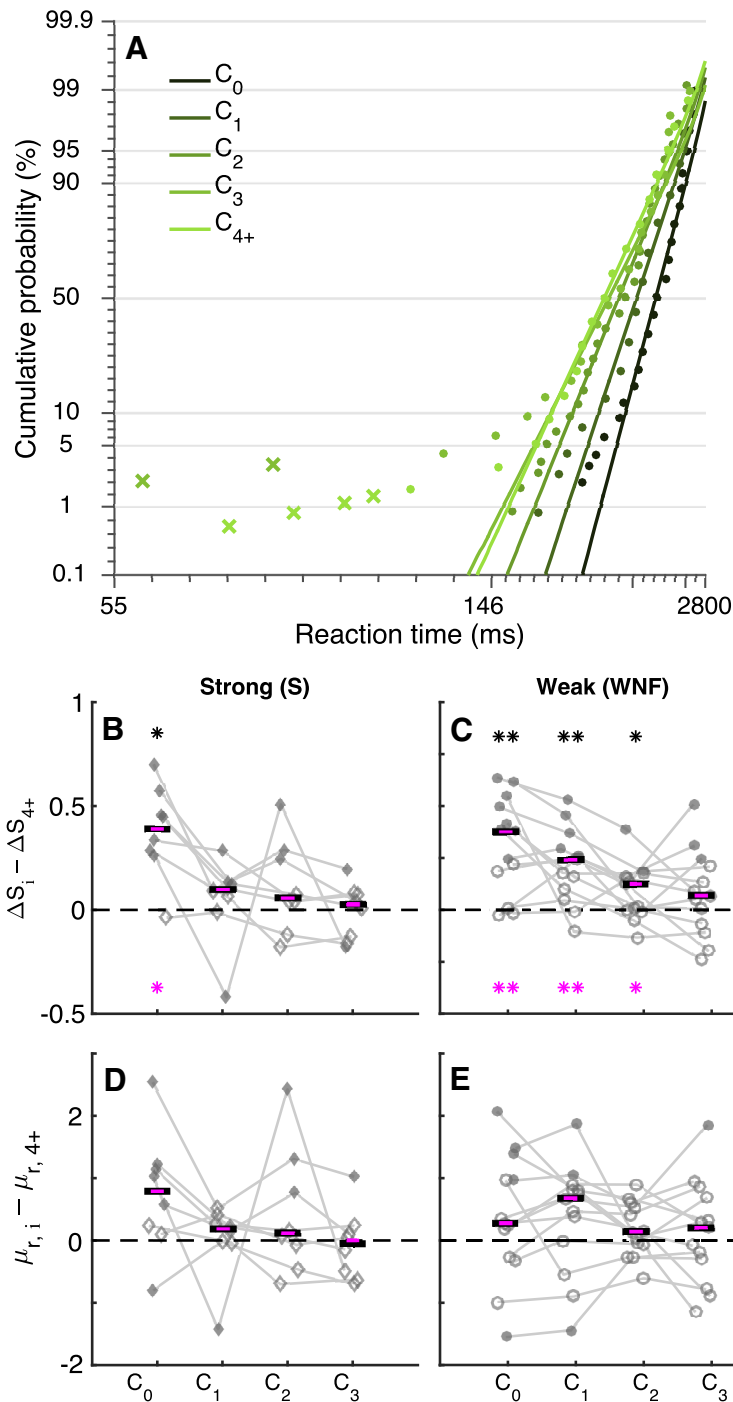


Figure 3. Saccadic priors reflect the accumulation of weak evidence over trials for individual subjects. **A**, Example saccadic RT distributions from a single subject (points) plotted on a reciprob scale for weak stimuli with feedback. Data are separated by color, as indicated, based on the trial number after a change point: C₀ is a choice on a change point (switch) trial and C_i is a choice on the ith trial after a change point (stay). Lines are fits of the LATER model, computed separately for each condition, to regular saccades (circles) but not short-latency saccades (crosses). **B–E**, Best-fitting values of ΔS (**B**, **C**) or μ_r (**D**, **E**) from the basic LATER model, fit to correct trials only, standardized to their steady-state values (computed from 4+ trials after a change point). These values are plotted as a function of trials 0–3 after a change point (indicated on the abscissa) computed per subject (connected points). **B**, **D**, Strong stimuli (n = 8 subjects). **C**, **E**, Weak stimuli with no feedback (n = 13 subjects). Filled points indicate that the given value is significantly nonzero (likelihood-ratio test, p < 0.01). Black lines indicate across-subject medians per task condition; asterisks indicate whether the median deviates significantly from zero (i.e., with respect to their steady-state, prechange point values; two-sided Wilcoxon rank-sum test, *p < 0.05; **p < 0.01; ***p < 0.001). Magenta lines (and asterisks) indicate medians of best-fitting values from the corresponding parameters (corresponding to correct choices) of the model with dual accumulators, which was fit to both correct and error choices. In all cases, H_{LR} = H_{RL} = 0.2 or 0.25.

which accounted for differences in the numbers of parameters between the two sets of non-nested models; Fig. 5C). Under these conditions, priors tended to change gradually after change points (e.g., Fig. 3C), as predicted by the adaptive model (Fig. 1C) but not the basic model. In contrast, the two sets of models tended to provide similar fits for conditions in which priors were predicted by the adaptive model to either change abruptly (S) or not change (WNF_{sp}) after change points (Fig. 1C), both of which could be accounted for using the basic model applied separately to conditions with different generative transition probabilities. However, there was also considerable individual variability across conditions in the relative quality of the fits of the two models. This variability was strongly dependent on variability in the time course of prior updates. In particular, subjects better fit by the adaptive model tended to update their priors more gradually across trials (Fig. 5D). This result emphasizes the flexibility of the adaptive model and its general applicability to all of the conditions we tested.

This flexibility reflected task-dependent differences in best-fitting parameter values from the adaptive model (Fig. 6). The parameter μ_r, representing the mean slope of the rising process that is sensitive to the strength of the instructive sensory signal (Reddi et al., 2003), was stronger in the strong versus the weak stimulus conditions (Fig. 6A; Wilcoxon rank-sum tests comparing S separately with WF, WNF, and WNF_{sp}, p < 0.001 in each case, and comparing WF separately with WNF and WNF_{sp} and WNF with WNF_{sp}, p > 0.24 in each case) and in all cases nearly identical to best-fitting values from the basic models (Pearson’s r comparing best-fitting μ_r values from the two models, across all conditions, was 0.96, p < 0.00001), as expected. The parameter α, which governs the weight of evidence used across trials to infer the current state, was generally larger for strong versus weak stimuli unless feedback about the current state was given after each saccade (Wilcoxon rank-sum tests comparing S and WF, p = 0.38; S and WNF, p < 0.01; S and WNF_{sp}, p < 0.001). The parameter S_T, representing the threshold of the LATER process, was lowest in the speeded condition (Fig. 6C; Wilcoxon rank-sum tests comparing WNF_{sp} separately with S, WF, and WNF, p < 0.01 in each case). This result is consistent with the idea that this parameter governs the speed–accuracy trade-off (Gold and Shadlen, 2007; Bogacz et al., 2010), which

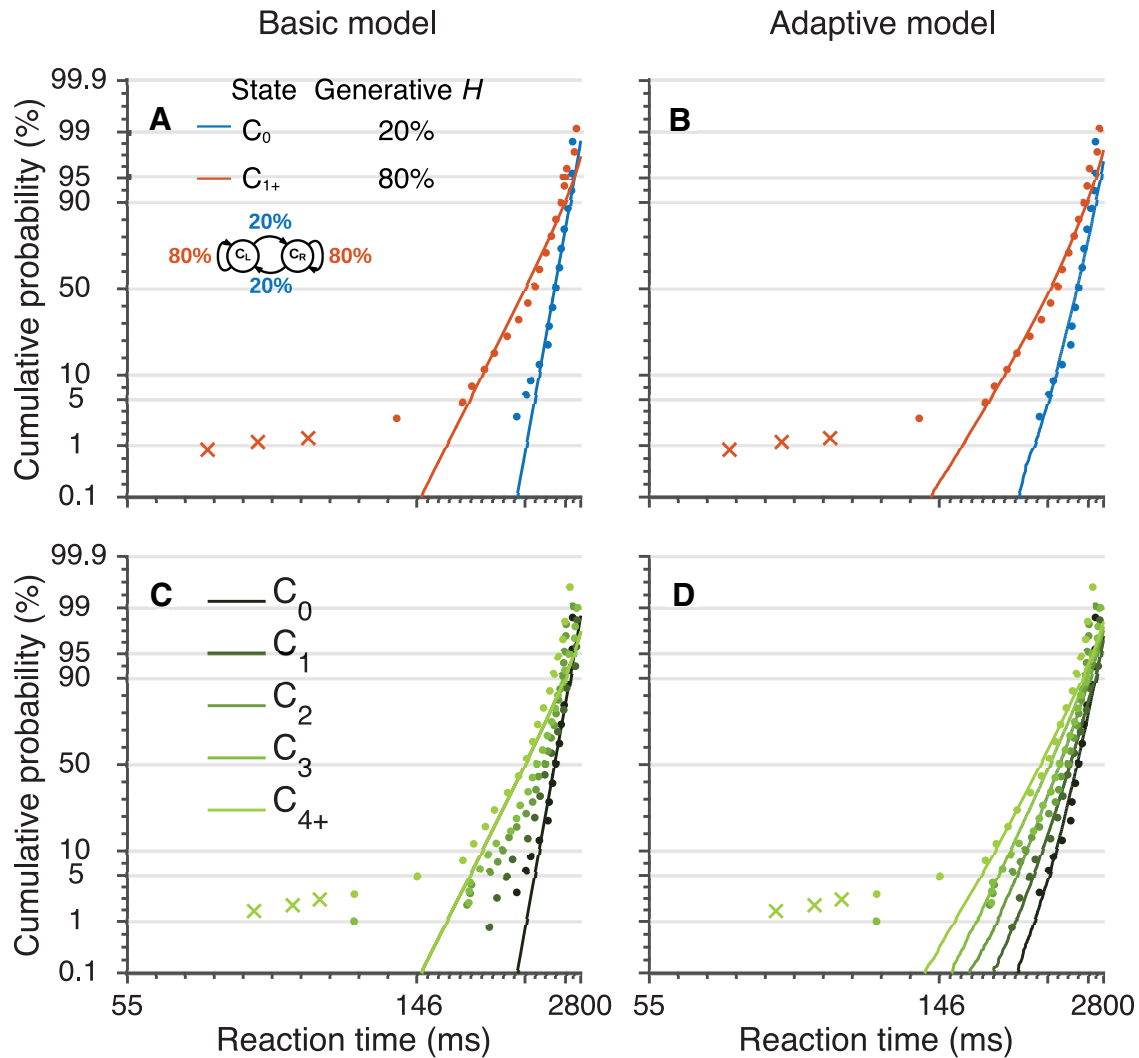


Figure 4. Adaptive LATER model that accounts for trial-by-trial, normative evidence accumulation matches data better than the basic LATER model. Example saccadic RT distributions (points for saccades fit to the model, crosses for short-latency saccades not fit to the model) for a single subject performing the task using WF stimuli. The same data are shown in each panel, separated by either change point (blue) versus nonchange point (red) trials (**A, B**) or by trials after change point (**C, D**). C_i refers to the state on the i^{th} trial after a change point. **A, C**, Lines are fits of the basic LATER model to all correct, non-short-latency saccades grouped by the generative transition probability. Because there were only two unique, generative transition probabilities in these sessions (one for all C_0 , or switch, trials and one for all C_{1+} , or stay, trials; see inset in **A**), there are only two fit lines in each panel. **B, D**, Lines are fits of the adaptive LATER model to the full sequence of correct, non-short-latency saccades. Because the adaptive model is sensitive to both the generative transition probability and the number of trials after a change point, there are different numbers of lines in each panel.

for our subjects tended to emphasize accuracy (particularly in the WF condition) unless otherwise instructed (i.e., WNF_{sp}).

The final two parameters governed subjective estimates of the two generative transition probabilities, H_{LR} and H_{RL} . Best-fitting values of these parameters were correlated with their corresponding objective values across the relatively broad range that we tested in the strong stimulus condition (Fig. 6D). This result indicates that the subjects learned and maintained approximately appropriate estimates of the generative transition probabilities based on experience. However, subjective estimates of either parameter alone tended to be biased toward values of 0.5 and, accordingly, their sum tended to be biased toward values of 1.0 (Fig. 6E). These biases were in the direction of history-independent values, as has been reported previously (Glaze et al., 2015). This bias was not simply driven by uncertainty about change points, because it was most reliably present when strong stimuli or weak stimuli with feedback were used (Fig. 6D). Therefore, in general, subjects tended to implicitly assume that the probability of the

sound coming from a particular direction was more independent of the previous trials than it actually was.

Saccadic priors reflected interactions between within- and across-saccade processes

These model fits also yielded insights into further interactions between the within-trial rise-to-threshold process that leads to a saccade and the across-trial inference process that establishes appropriate expectations about the state-dependent transition probabilities. When feedback was given (including all subjects in the WF condition and all but two subjects in the S condition), it provided unambiguous evidence for the current state. Accordingly, under these conditions, the within- and across-saccade decision processes were not strongly coupled. Specifically, the best-fitting values of α from the adaptive model, representing the weight of evidence used across trials to infer the current state, and S_T , representing the total prior and sensory evidence in the LATER process used to infer the current state before feedback,

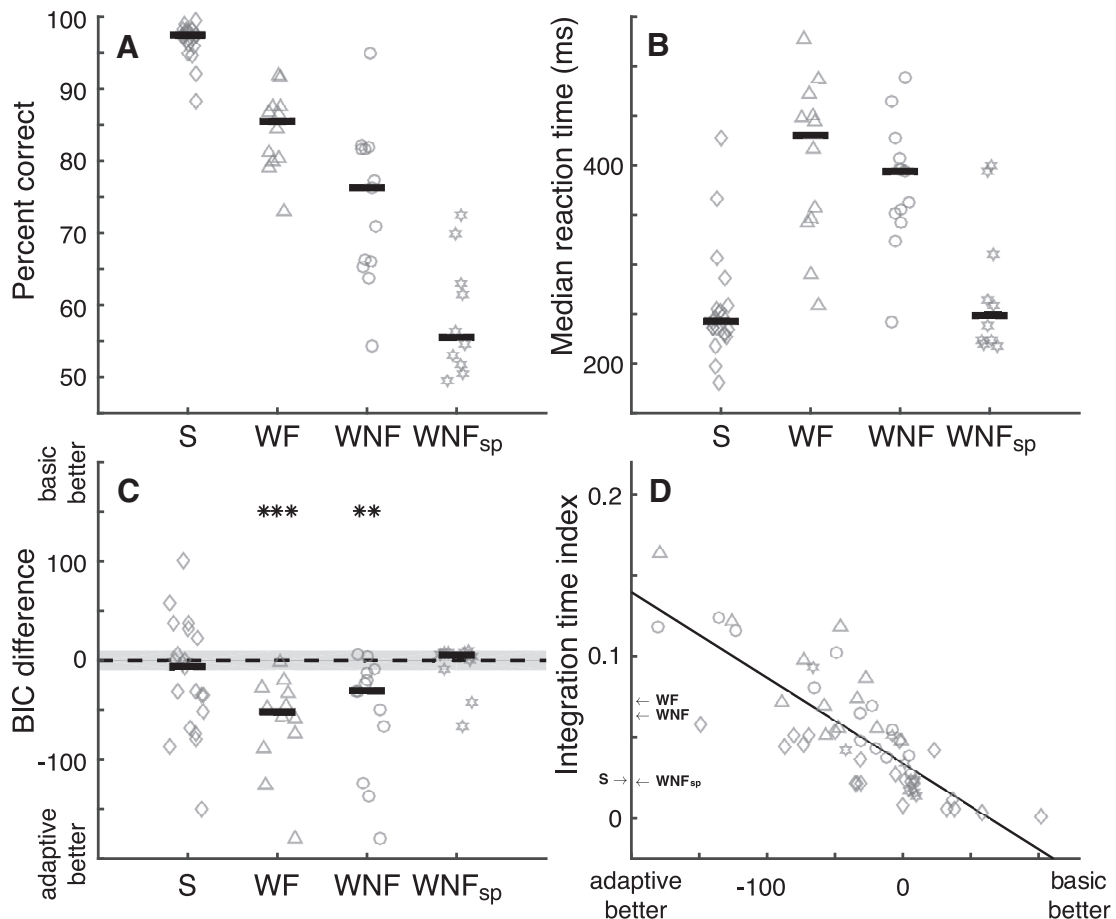


Figure 5. Summary of performance and adaptive LATER model fits, for four task conditions: (1) S stimuli (diamonds); (2) WF stimuli (triangles); (3) WNF stimuli (circles); and (4) WNF_{sp} stimuli (stars). In each panel, points are data computed for a single subject and grouped by condition on the abscissa. **A**, Overall accuracy. **B**, Median RT. **C**, Bayesian information criterion (BIC) difference between the basic LATER model and the adaptive LATER model per condition (abscissa). Negative/positive differences indicate support for the adaptive/basic model. The gray area indicates BIC differences of ± 10 . Asterisks indicate two-sided Wilcoxon rank-sum test for nonzero median. ** $p < 0.01$; *** $p < 0.001$. **D**, Degree of gradual updating of priors across trials (larger numbers imply more gradual updating; see Materials and Methods) as a function of the BIC difference for the two models (as in **C**). The line is a linear fit; Spearman's $\rho = -0.81$, $p < 0.001$. Labeled arrows along the ordinate indicate median values per condition.

were unrelated to each other across subjects (Pearson's $r = -0.21$, $p = 0.40$; Fig. 7A). In contrast, when feedback was not given (the WNF and WNF_{sp} conditions), the total accumulated evidence for the within-saccade decision appeared to drive the across-saccade process. Specifically, there was an approximately linear relationship between best-fitting values of logit α and S_T across subjects ($r = 0.76$, $p < 0.0001$; Fig. 7B). This relationship reflected effects that occurred both across task conditions that explicitly manipulated S_T (median values of both α and S_T were smaller in the WNF_{sp} versus WNF conditions; Wilcoxon rank-sum tests, $p < 0.01$ in both cases) and within task conditions (WNF: $r = 0.78$, $p = 0.0019$, WNF_{sp}: $r = 0.80$, $p = 0.0063$; although only the former, which included strong integration across trials, held for a rank-order test, WNF: Spearman's $\rho = 0.65$, $p = 0.0196$, WNF_{sp}: $\rho = 0.40$, $p = 0.26$). We did not find any relationship between α and μ_r , which governed the within-trial evidence, for any task condition ($p > 0.05$ for ρ in each case). This result implies that, in the absence of other feedback, the total, predefined level of evidence accumulated in each trial can be used by the across-trial inference process to update expectations about state-dependent transition probabilities.

Discussion

We used an auditory discrimination task to determine how sequential, saccadic RTs reflect learned expectations in dy-

namic environments. We presented three main, novel findings. First, learned state-dependent transition probabilities affect the prior probability of making a particular saccade. In the context of saccadic RTs, these effects can be modeled effectively as changes in the starting point of a rise-to-threshold saccade-generation process. This finding is consistent with both experimental and theoretical work implicating a similar computational basis for certain effects of expectations on RT and choice behavior using other tasks (Carpenter and Williams, 1995; Bogacz et al., 2006; Gold and Shadlen, 2007; Mulder et al., 2012; Glaze et al., 2015). Second, these adaptive adjustments follow principles of normative evidence accumulation, including gradual updates across trials (saccades) when there is uncertainty about the current state (Wilson et al., 2010; Glaze et al., 2015). Third, this across-saccade evidence-accumulation process depends on a confidence-like signal (Kiani and Shadlen, 2009; Purcell and Kiani, 2016), which in our task was linked to the within-saccade decision threshold that was also sensitive to feedback. Together, these results imply that saccade generation is governed by sophisticated and interacting forms of inference that operate on multiple timescales. Below we discuss in more detail how these findings relate to previous studies and how they can help guide our understanding of the underlying neural mechanisms.

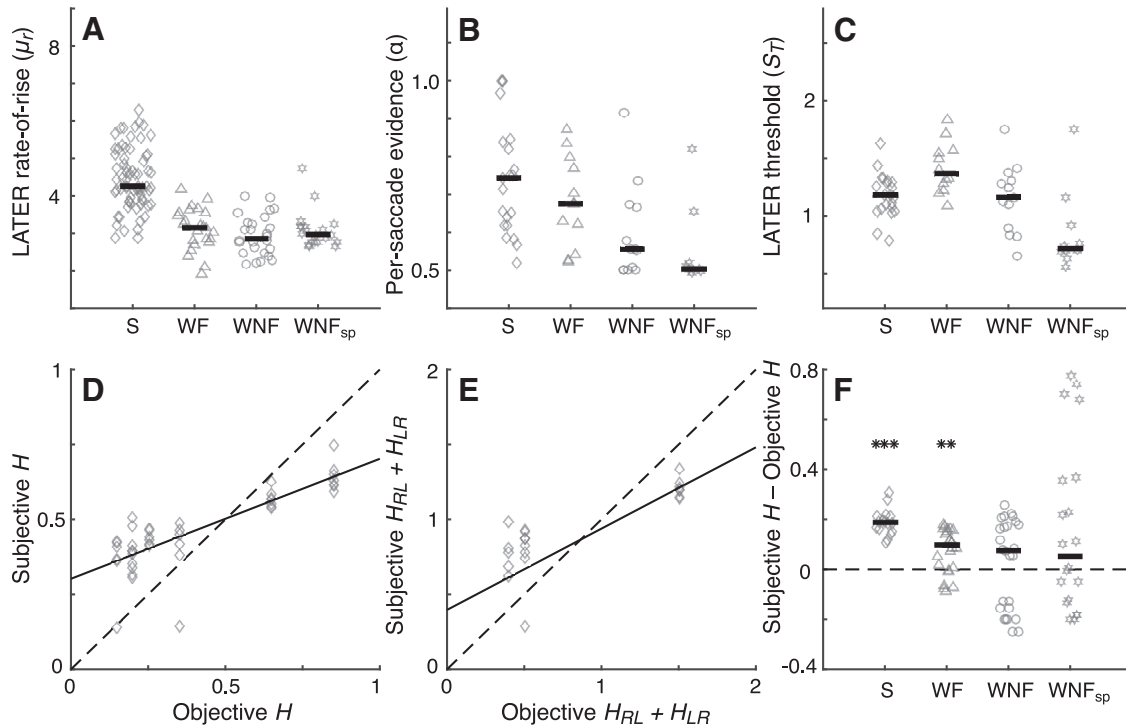


Figure 6. Summary of adaptive LATER model parameters, for four task conditions: (1) S stimuli (diamonds); (2) WF stimuli (triangles); (3) WNF stimuli (circles); and (4) WNF_{sp} stimuli (stars). **A–C**, Best-fitting values of μ_r (**A**), α (**B**), and S_T (**C**) from the adaptive model computed per subject (points) and grouped by condition on the abscissa. **D**, Best-fitting values of subjective transition probabilities from the adaptive model plotted against the corresponding objective (generative) transition probabilities for the strong stimuli condition. The solid black line is the least-square linear regression (slope = 0.40; $\rho = 0.80$, $p < 0.001$). **E**, Best-fitting sum of subjective $H_{LR} + H_{RL}$ against the sum of the corresponding objective values. Solid black line is the least-square linear regression (slope = 0.54; $\rho = 0.78$, $p < 0.001$). **F**, Difference between subjective and objective H per subject for each task condition. Only objective values of 0.2 or 0.25 are included for consistency across conditions. Conventions are as in **A–C**. Asterisks indicate two-sided Wilcoxon rank-sum test for nonzero median. ** $p < 0.01$, *** $p < 0.001$.

Relationship to other work

Our findings provide new insights into the computational basis for sequential structure in saccadic RTs, as have been studied extensively using versions of the serial RT task (Nissen and Bullemer, 1987; Reed and Johnson, 1994; Schwarb and Schumacher, 2012). In particular, we showed that this sequential structure can reflect specific, saccade-by-saccade adjustments to normative, rise-to-threshold decision processes. These findings put RT measurements, long recognized as window into higher brain function (Luce, 1986; Noorani and Carpenter, 2016), on more even footing with other saccade metrics that have been linked to principles of normative decision making in uncertain environments. For example, in visual search tasks, visual saccadic inhibition of return reflects trial-by-trial updating of subjective prior beliefs based on recent experience and contextual cues in the environment (Ludwig et al., 2012). Similar to our findings, these effects can be accounted for via changes in the baseline value of a LATER-like model of saccadic decision making (Farrell et al., 2010). Saccadic accuracy in visual search, along with saccadic choices on a cued discrimination task, are similarly modulated by expectations in a manner that is consistent with optimal, probabilistic inference (Shimozaki et al., 2003; Eckstein et al., 2006). Together, these results imply that the timing, selection, and accuracy of saccadic

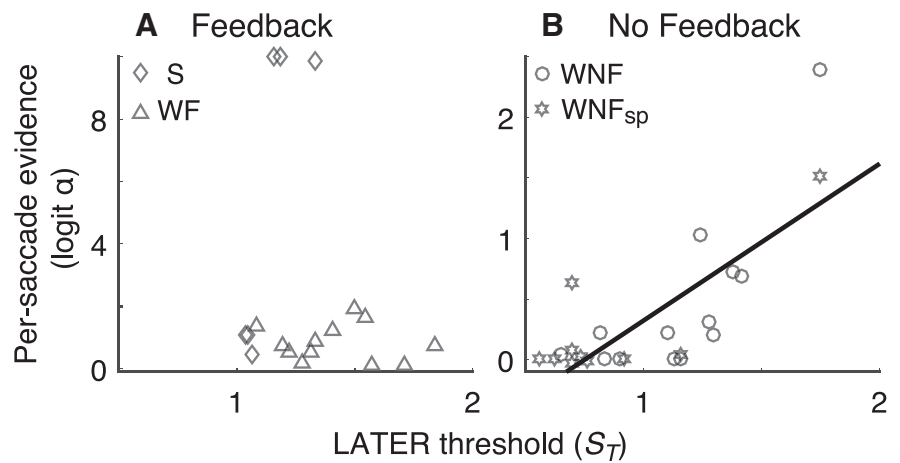


Figure 7. Relationship between S_T and α . **A**, Conditions with feedback: S stimuli (diamonds) and WF stimuli (triangles). **B**, Conditions with no feedback: WNF stimuli (circles) and WNF_{sp} stimuli (stars). Points are best-fitting values of the adaptive model to data from individual subjects. Line in **B** is a linear fit ($p < 0.001$).

eye movements can yield important insights into the complex computations that the brain employs to use expectations to optimize behavior.

These RT-sensitive computations include certain principles of optimal inference that are needed to make effective predictions in environments that are both uncertain and can undergo fundamental, unexpected changes (Behrens et al., 2007; Nassar et al., 2010; Payzan-LeNestour et al., 2013; Gallistel et al., 2014; Glaze et al., 2015). Somewhat similar principles appear to govern sequential RT effects measured in key-press tasks, although previous studies using

those tasks focused only on conditions with little or no state uncertainty. Specifically, under those conditions, sequential key-press RTs appear to involve trial-by-trial updates of the starting or ending point of a rise-to-threshold decision process, following both normative and non-normative principles (Goldfarb et al., 2012; Zhang et al., 2014). We showed that similar updates to a normative decision process that governs saccadic RTs are sensitive to state-dependent transition probabilities, albeit with an apparent predisposition to assume history-independent values of ~ 0.5 that may minimize computational costs (Drugowitsch et al., 2012; Shenhav et al., 2013). Moreover, in the presence of uncertainty, these processes can flexibly accumulate evidence to identify the current state of the environment. Both of these features of saccadic processing are hallmarks of high-level inference processes and thus support the notion that the saccadic system is a useful substrate for understanding, not just decisions that lead to immediate actions, but also how those decisions depend on learned behavioral contexts or models of the world (Gold and Shadlen, 2007; Otto et al., 2015; Purcell and Kiani, 2016).

Possible neural mechanisms

A primary advantage of studying the saccadic system is our extensive knowledge of the underlying neural mechanisms. Our findings suggest several novel extensions to existing ideas about these mechanisms that can be used to guide future experiments. For example, the within-saccade decisions that in our task drive the selection and timing of the impending saccade relate directly to numerous lines of research involving measurements of neural activity in nonhuman primates performing relatively simple sensory-driven saccade tasks. These studies have identified correlates of the kind of rise-to-threshold processes described by the LATER model in a network of interconnected brain regions including the lateral intraparietal area (LIP), frontal eye field (FEF), superior colliculus (SC), and basal ganglia (Hanes and Schall, 1996; Horwitz and Newsome, 1999; Kim and Shadlen, 1999; Roitman and Shadlen, 2002; Ding and Gold, 2010; Ding and Gold, 2012). Consistent with the prior-driven changes in starting value of the LATER process that we found, baseline activity in each of these areas that occurs before the onset of visual cues instructing particular saccades can under certain conditions reflect the probabilities of making those saccades (Basso and Wurtz, 1997; Dorris and Munoz, 1998; Coe et al., 2002; Lauwereyns et al., 2002; Ding and Hikosaka, 2006; Rao et al., 2012). Our results suggest that these baseline modulations may be driven by across-saccade decision processes that are sensitive to stimulus discriminability, feedback, and change point dynamics.

There is less known about how and where in the brain these kinds of across-saccade processes are implemented. Nevertheless, neural correlates of some of the key computational components have been identified. In the context of saccadic eye movements, signals related to feedback monitoring have been identified in several brain regions, including parts of the medial frontal and lateral prefrontal cortex (Stuphorn et al., 2000; Scangos et al., 2013; Teichert et al., 2014). Signals related to the detection and use of environmental change point dynamics for inference problems have been found in the cingulate cortex and may also involve more general arousal-related modulations that include a primary recipient of cingulate output, the locus ceruleus–norepinephrine system (Behrens et al., 2007; Nassar et al., 2012; Tervo et al., 2014; Joshi et al., 2016). Integration of these signals across saccades may involve brain regions that encode history-dependent information related to saccade choices, which in-

cludes parts of parietal, cingulate, and prefrontal cortex (Sugrue et al., 2004; Bernacchia et al., 2011). This integration process also likely depends on working memory signals that involve interactions between the hippocampus and cortex (Shadlen and Shohamy, 2016), which would be intriguing targets of future studies using this kind of task.

We further propose that the interactions that we found between these within- and across-saccade decision processes may also have a compelling mapping onto a known neural circuit: the information-processing loops involving the cortex and basal ganglia. In the oculomotor version of this loop, cortical areas FEF and LIP both project to the caudate in the basal ganglia. Caudate output is divided between direct and indirect pathways through the basal ganglia that ultimately converge in the substantia nigra, pars reticulata (SNr). SNr output, in turn, goes both to the SC and subsequently to brainstem circuits that control movement and, via the thalamus, back up to cortex (Hikosaka et al., 2000). Recent work has emphasized this circuit's role in integrating sensory and other information to form decisions that guide saccadic behavior (Bogacz and Gurney, 2007; Ding and Gold, 2013; Wiecki and Frank, 2013). In part, these functions may involve adjusting the decision threshold, which is equivalent to S_T in our LATER model, but in the brain, may involve multiple changes in the dynamics of the within-saccade rise-to-threshold process (Lo and Wang, 2006; Forstmann et al., 2008; Heitz and Schall, 2012).

Together, these findings suggest that our task may provide a useful model system for studying functional information-processing loops in this cortical-basal ganglia system. In principle, these loops may involve the following flow of information. First, biases determined across trials are fed from cortex into the basal ganglia. Second, these biases are then integrated there with incoming sensory information to form a saccadic decision variable. Third, this decision variable terminates when reaching a predetermined bound that mediates the speed-accuracy trade-off. Fourth, this termination causes a release of inhibition that both facilitates the saccade plan encoded in the superior colliculus and causes information related to the decision bound to be sent back up to cortex, via the thalamus, to be used to determine subsequent biases. Among the novel hypotheses suggested by this scheme is that disruption of the basal ganglia–thalamus–cortex feedback pathway may hinder the ability of the brain to make decisions across saccades that depend on the amount of sensory evidence accumulated for each saccade. Our findings may provide a strong behavioral and theoretical basis for pursuing this kind of study in the future.

References

- Anderson AJ, Carpenter RH (2006) Changes in expectation consequent on experience, modeled by a simple, forgetful neural circuit. *J Vis* 6:822–835. [Medline](#)
- Basso MA, Wurtz RH (1997) Modulation of neuronal activity by target uncertainty. *Nature* 389:66–69. [CrossRef Medline](#)
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214–1221. [CrossRef Medline](#)
- Bernacchia A, Seo H, Lee D, Wang XJ (2011) A reservoir of time constants for memory traces in cortical neurons. *Nat Neurosci* 14:366–372. [CrossRef Medline](#)
- Bishop CM (2006) *Pattern recognition and machine learning*. New York: Springer.
- Bogacz R, Gurney K (2007) The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Comput* 19:442–477. [CrossRef Medline](#)
- Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced choice tasks. *Psychol Rev* 113:700–765. [CrossRef Medline](#)

- Bogacz R, Wagenmakers EJ, Forstmann BU, Nieuwenhuis S (2010) The neural basis of the speed-accuracy tradeoff. *Trends Neurosci* 33:10–16. [CrossRef Medline](#)
- Carpenter RH, Williams ML (1995) Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377:59–62. [CrossRef Medline](#)
- Carpenter RHS (2012) Analysing the detail of saccadic reaction time distributions. *Biocybernetics and Biomedical Engineering* 32:49–63. [CrossRef](#)
- Cho RY, Nystrom LE, Brown ET, Jones AD, Braver TS, Holmes PJ, Cohen JD (2002) Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cogn Affect Behav Neurosci* 2:283–299. [CrossRef Medline](#)
- Coe B, Tomihara K, Matsuzawa M, Hikosaka O (2002) Visual and anticipatory bias in three cortical eye fields of the monkey during an adaptive decision-making task. *J Neurosci* 22:5081–5090. [Medline](#)
- Courville AC, Daw ND, Touretzky DS (2006) Bayesian theories of conditioning in a changing world. *Trends Cogn Sci* 10:294–300. [CrossRef Medline](#)
- Ding L, Gold JI (2010) Caudate encodes multiple computations for perceptual decisions. *J Neurosci* 30:15747–15759. [CrossRef Medline](#)
- Ding L, Gold JI (2012) Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. *Cereb Cortex* 22:1052–1067. [CrossRef Medline](#)
- Ding L, Gold JI (2013) The basal ganglia's contributions to perceptual decision making. *Neuron* 79:640–649. [CrossRef Medline](#)
- Ding L, Hikosaka O (2006) Comparison of reward modulation in the frontal eye field and caudate of the macaque. *J Neurosci* 26:6695–6703. [CrossRef Medline](#)
- Dorris MC, Munoz DP (1998) Saccadic probability influences motor preparation signals and time to saccadic initiation. *J Neurosci* 18:7015–7026. [Medline](#)
- Drugowitsch J, Moreno-Bote R, Churchland AK, Shadlen MN, Pouget A (2012) The cost of accumulating evidence in perceptual decision making. *J Neurosci* 32:3612–3628. [CrossRef Medline](#)
- Eckstein MP, Drescher BA, Shimozaki SS (2006) Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychol Sci* 17:973–980. [CrossRef Medline](#)
- Farrell S, Ludwig CJ, Ellis LA, Gilchrist ID (2010) Influence of environmental statistics on inhibition of saccadic return. *Proc Natl Acad Sci U S A* 107:929–934. [CrossRef Medline](#)
- Fischer B, Ramsperger E (1984) Human express saccades: extremely short reaction times of goal directed eye movements. *Exp Brain Res* 57:191–195. [Medline](#)
- Forstmann BU, Dutilh G, Brown S, Neumann J, von Cramon DY, Ridderinkhof KR, Wagenmakers EJ (2008) Striatum and pre-SMA facilitate decision-making under time pressure. *Proc Natl Acad Sci U S A* 105:17538–17542. [CrossRef Medline](#)
- Friston K (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11:127–138. [CrossRef Medline](#)
- Gallistel CR, Krishan M, Liu Y, Miller R, Latham PE (2014) The perception of probability. *Psychol Rev* 121:96–123. [CrossRef Medline](#)
- Glaze CM, Kable JW, Gold JI (2015) Normative evidence accumulation in unpredictable environments. *Elife* 4.
- Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574. [CrossRef Medline](#)
- Goldfarb S, Wong-Lin KF, Schwemmer M, Leonard NE, Holmes P (2012) Can post-error dynamics explain sequential reaction time patterns? *Front Psychol* 3.
- Green DM, Swets JA (1966) *Signal detection theory and psychophysics*. New York: Wiley.
- Haith AM, Krakauer JW (2013) Model-based and model-free mechanisms of human motor learning. *Adv Exp Med Biol* 782:1–21. [CrossRef Medline](#)
- Hanes DP, Schall JD (1996) Neural control of voluntary movement initiation. *Science* 274:427–430. [CrossRef Medline](#)
- Heitz RP, Schall JD (2012) Neural mechanisms of speed-accuracy tradeoff. *Neuron* 76:616–628. [CrossRef Medline](#)
- Hikosaka O, Takikawa Y, Kawagoe R (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol Rev* 80:953–978. [Medline](#)
- Horowitz GD, Newsome WT (1999) Separate signals for target selection and movement specification in the superior colliculus. *Science* 284:1158–1161. [CrossRef Medline](#)
- Joshi S, Li Y, Kalwani RM, Gold JI (2016) Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron* 89:221–234. [CrossRef Medline](#)
- Kiani R, Shadlen MN (2009) Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324:759–764. [CrossRef Medline](#)
- Kim JN, Shadlen MN (1999) Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci* 2:176–185. [CrossRef Medline](#)
- King-Smith PE, Grigsby SS, Vingrys AJ, Benes SC, Supowit A (1994) Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision Res* 34:885–912. [CrossRef Medline](#)
- Lauwereyns J, Watanabe K, Coe B, Hikosaka O (2002) A neural correlate of response bias in monkey caudate nucleus. *Nature* 418:413–417. [CrossRef Medline](#)
- Lo CC, Wang XJ (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat Neurosci* 9:956–963. [CrossRef Medline](#)
- Luce RD (1986) *Response times: their role in inferring elementary mental organization*. New York: OUP.
- Ludwig CJ, Farrell S, Ellis LA, Hardwicke TE, Gilchrist ID (2012) Context-gated statistical learning and its role in visual-saccadic decisions. *J Exp Psychol Gen* 141:150–169. [CrossRef Medline](#)
- Mulder MJ, Wagenmakers EJ, Ratcliff R, Boekel W, Forstmann BU (2012) Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *J Neurosci* 32:2335–2343. [CrossRef Medline](#)
- Nakahara H, Nakamura K, Hikosaka O (2006) Extended LATER model can account for trial-by-trial variability of both pre- and post-processes. *Neural Netw* 19:1027–1046. [CrossRef Medline](#)
- Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 30:12366–12378. [CrossRef Medline](#)
- Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012) Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat Neurosci* 15:1040–1046. [CrossRef Medline](#)
- Nissen MJ, Bullemer P (1987) Attentional requirements of learning: evidence from performance measures. *Cognitive Psychology* 19:1–32. [CrossRef](#)
- Noorani I, Carpenter RH (2016) The LATER model of reaction time and decision. *Neurosci Biobehav Rev* 64:229–251. [CrossRef Medline](#)
- O'Reilly JX (2013) Making predictions in a changing world-inference, uncertainty, and learning. *Front Neurosci* 7:105. [CrossRef Medline](#)
- Oswal A, Ogden M, Carpenter RH (2007) The time course of stimulus expectation in a saccadic decision task. *J Neurophysiol* 97:2722–2730. [CrossRef Medline](#)
- Otto AR, Skatova A, Madlon-Kay S, Daw ND (2015) Cognitive control predicts use of model-based reinforcement learning. *J Cogn Neurosci* 27:319–333. [CrossRef Medline](#)
- Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP (2013) The neural representation of unexpected uncertainty during value-based decision making. *Neuron* 79:191–201. [CrossRef Medline](#)
- Pezzulo G, Rigoli F, Friston K (2015) Active Inference, homeostatic regulation and adaptive behavioural control. *Prog Neurobiol* 134:17–35. [CrossRef Medline](#)
- Posner MI (1980) Orienting of attention. *Q J Exp Psychol* 32:3–25. [CrossRef Medline](#)
- Posner MI, Snyder CR, Davidson BJ (1980) Attention and the detection of signals. *J Exp Psychol* 109:160–174. [CrossRef Medline](#)
- Purcell BA, Kiani R (2016) Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. *Proc Natl Acad Sci U S A* 113:E4531–E4540. [CrossRef Medline](#)
- Rao V, DeAngelis GC, Snyder LH (2012) Neural correlates of prior expectations of motion in the lateral intraparietal and middle temporal areas. *J Neurosci* 32:10063–10074. [CrossRef Medline](#)
- Reddi BA, Carpenter RH (2000) The influence of urgency on decision time. *Nat Neurosci* 3:827–830. [CrossRef Medline](#)
- Reddi BA, Asrress KN, Carpenter RH (2003) Accuracy, information, and response time in a saccadic decision task. *J Neurophysiol* 90:3538–3546. [CrossRef Medline](#)
- Reed J, Johnson P (1994) Assessing implicit learning with indirect tests: determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20:585. [CrossRef](#)
- Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intra-

- parietal area during a combined visual discrimination reaction time task. *J Neurosci* 22:9475–9489. [Medline](#)
- Scangos KW, Aronberg R, Stuphorn V (2013) Performance monitoring by pre-supplementary and supplementary motor area during an arm movement countermanding task. *J Neurophysiol* 109:1928–1939. [CrossRef Medline](#)
- Schwarb H, Schumacher EH (2012) Generalized lessons about sequence learning from the study of the serial reaction time task. *Adv Cogn Psychol* 8:165–178. [CrossRef Medline](#)
- Seriès P, Seitz AR (2013) Learning what to expect (in visual perception). *Front Hum Neurosci* 7:668. [CrossRef Medline](#)
- Shadlen MN, Shohamy D (2016) Decision making and sequential sampling from memory. *Neuron* 90:927–939. [CrossRef Medline](#)
- Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79:217–240. [CrossRef Medline](#)
- Shimozaki SS, Eckstein MP, Abbey CK (2003) Comparison of two weighted integration models for the cueing task: linear and likelihood. *J Vis* 3:209–229. [Medline](#)
- Stuphorn V, Taylor TL, Schall JD (2000) Performance monitoring by the supplementary eye field. *Nature* 408:857–860. [CrossRef Medline](#)
- Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304:1782–1787. [CrossRef Medline](#)
- Teichert T, Yu D, Ferrera VP (2014) Performance monitoring in monkey frontal eye field. *J Neurosci* 34:1657–1671. [CrossRef Medline](#)
- Tervo DG, Proskurin M, Manakov M, Kabra M, Vollmer A, Branson K, Karpova AY (2014) Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell* 159:21–32. [CrossRef Medline](#)
- Watson AB, Pelli DG (1983) QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys* 33:113–120. [CrossRef Medline](#)
- Wiecki TV, Frank MJ (2013) A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychol Rev* 120:329–355. [CrossRef Medline](#)
- Wilson RC, Nassar MR, Gold JI (2010) Bayesian online learning of the hazard rate in change point problems. *Neural Comput* 22:2452–2476. [CrossRef Medline](#)
- Zhang S, Huang HC, Yu AJ (2014) Sequential effects: a bayesian analysis of prior bias on reaction time and behavioral choice in *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (Québec City, QC).