Chapter 7: Perceptual decisions

This chapter considers decisions about discrete hidden states based on perceptual evidence. We first formalize the computational problem, and then derive algorithms for solving it. The basic algorithmic motif is evidence accumulation to a decision threshold. We show how this motif can be implemented in neural circuitry, how it can adaptively optimize reward rate, and how it can support downstream computation of decision confidence.

Recall the perceptual discrimination problem from Chapter 4, where the hidden state corresponds to one of several discrete states: $s \in \{1, \ldots, S\}$. The hidden state generates a perceptual signal (x), which an agent uses to discriminate the hidden state. For example, the agent observes a cloud of dots moving either left or right (on average), and must judge the direction of motion. Because the evidence provided by the neural encoding of sensory signals is noisy, a single snapshot will in general not suffice to make a good decision. Rather, the optimal algorithm is to accumulate evidence across time until it crosses a decision threshold. While simple, this algorithm has proven to be a powerful model of perceptual decisions in the brain.

1 *Perceptual discrimination problems*

We will start with the simplest case: two hidden states, $s \in \{A, B\}$, sampled with probability $p(s)$ on each trial. The agent observes a time series of signals $x(t)$, chooses an action $a \in \{A, B\}$, and receives reward $r = R_s$ if $a = s$ (0 otherwise). It will be convenient initially to think about the decision problem in discrete time, where signals are sampled at intervals of length δ; each signal $x(t)$ represents the information acquired over the interval. We will consider two versions of the discrimination problem: (1) the *interrogation paradigm*, where a fixed response deadline is imposed on the agent; and (2) the *free response paradigm*, where the agent chooses when to respond.

Bayes' rule stipulates how to calculate the evidence favoring each state at time t after observing sensory history $X(t) = \{x(t') : t' \leq t\}$. We write the posterior in log-odds form:

$$\log \frac{p(s = A|X(t))}{p(s = B|X(t))} = \log \frac{p(X(t)|s = A)}{p(X(t)|s = B)} + \log \frac{p(s = A)}{p(s = B)}. \quad (1)$$

The first term on the right-hand side is the log likelihood ratio, and the second term is the log prior ratio. When the prior probabilities of the two states are equal, the second term is 0. If the sensory signals

are independent, the log posterior odds can be written in recursive form. Let $L(t)$ denote the log posterior odds at time t. Initializing it to $L(0) = \log \frac{p(s=A)}{p(s=B)}$, the log odds is incremented over a short interval δ according to the momentary evidence supplied by sensory signals:

$$L(t) - L(t-\delta) = \log \frac{p(x(t)|s=A)}{p(x(t)|s=B)}. \tag{2}$$

Assuming the agent's utility is simply the received reward, $u(r) = r$, the expected utility of choosing action a at time t is:

Note that $p(s = A|X(t)) = \sigma(L(t))$ and $p(s = B|X(t)) = \sigma(-L(t))$.

$$\mathbb{E}[u(r)|s, a, L(t)] = \begin{cases} \sigma(L(t))R_s, & \text{if } a = A \\ \sigma(-L(t))R_s, & \text{if } a = B, \end{cases} \tag{3}$$

where $\sigma(L) = 1/(1 + e^{-L})$ is the logistic sigmoid, which transforms log odds to probability. We can use this expression to define a decision variable, the log expected utility ratio:

$$U(t) = \log \frac{\mathbb{E}[u(r)|s, a = A, L(t)]}{\mathbb{E}[u(r)|s, a = B, L(t)]} = 2L(t) + \log \frac{R_A}{R_B}, \tag{4}$$

which is maximized by choosing

$$a^* = \begin{cases} A, & \text{if } U(t) > 0 \\ B, & \text{if } U(t) < 0. \end{cases} \tag{5}$$

This is the Bayes-optimal policy for the interrogation paradigm, where the agent is obligated to make a choice at a particular time. When the reward for a correct response is constant across states, the second term in Eq. 4 is 0, and the decision variable is directly proportional to the log posterior odds. In the general case, the decision variable is just a shifted version of the log posterior odds, where the shift is determined by the asymmetry of rewards across states. Intuitively, the agent will be biased towards high-reward states.

In the free response paradigm, the agent can potentially increase their reward rate by responding more quickly. However, there is a speed-accuracy trade-off: faster responses will tend to yield less reward because the quality of evidence is lower. One way to approach this problem is to set decision thresholds U_A and U_B, such that $a = A$ if $U(t) > U_A$ and $a = B$ if $U(t) < U_B$. If no threshold has been crossed, the agent withholds a response. By setting these thresholds, the agent can determine their preference for speed vs. accuracy. When the threshold separation $\Delta U = U_A - U_B$ is larger, the agent will take longer to respond and the responses will be more accurate on average. We will shortly discuss how to set the thresholds optimally.

2 *The drift-diffusion model*

The previous section introduced a generic class of *evidence accumulation* models for two-alternative perceptual discrimination problems. An important special case is where the signal-generating process is Gaussian:

$$x(t) \sim \mathcal{N}(\delta\mu_s, \delta\nu), \tag{6}$$

where μ_s is the signal strength and ν is the signal variance. The momentary evidence accumulated over δ is also Gaussian distributed (Bogacz et al., 2006; Bitzer et al., 2014):

$$L(t) - L(t-\delta) \sim \mathcal{N}(\delta\theta_s, \delta\kappa^2), \tag{7}$$

with

$$\theta_s = \mu_s \frac{\mu_A - \mu_B}{\nu} \tag{8}$$

$$\kappa = \frac{\mu_A - \mu_B}{\sqrt{\nu}}. \tag{9}$$

In the continuum limit $\delta \to 0$, this becomes the *drift-diffusion model* (DDM; Figure 1), which can be written as a stochastic differential equation:

$$dL = \theta_s dt + \kappa dW, \tag{10}$$

where dW is the differential of the Wiener process. The "drift" corresponds to the deterministic component (controlled by the drift rate θ_s), and the "diffusion" corresponds to the stochastic component (controlled by the noise variance κ).

We can use these equations to calculate the expected reward, $\mathbb{E}[r|s]$, and the expected decision time, $\mathbb{E}[T|s]$:

$$\mathbb{E}[r|s] = R_s \sigma(E_s) \tag{11}$$

$$\mathbb{E}[T|s] = \frac{U_s}{\theta_s}\left[2\sigma(E_s) - 1\right] \tag{12}$$

where

$$E_s = \frac{2\theta_s U_s}{\kappa^2} \tag{13}$$

is the log odds of a correct response, and $\sigma(E_s)$ is the corresponding probability (i.e., the accuracy). These expressions make the speed-accuracy trade-off transparent: expected decision time monotonically increases with expected reward. In the absence of noise ($\kappa \to 0$), the accuracy saturates to $\sigma(E_s) = 1$ and the expected decision time simplifies to the ratio between the threshold and the drift rate: $\mathbb{E}[T|s] = \frac{U_s}{\theta_s}$.

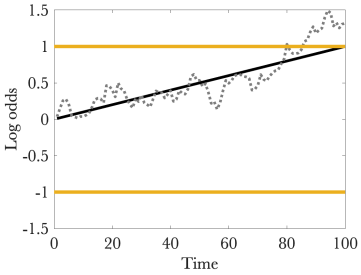

Figure 1: **The drift-diffusion model**. The black line shows the drift component, and the dotted gray line shows the trajectory of the log odds over time (the sum of the drift and diffusion components). The horizontal lines show the decision thresholds.

The Wiener process, also known as *Brownian motion*, is a stochastic process with independent Gaussian increments: $W(t) - W(t-\delta) \sim \mathcal{N}(0, \delta)$.

The empirically measured response time, RT $= T_0 + T$, is decomposed into the signal-dependent decision time T and the signal-independent "non-decision time" T_0.

If we assume that the problem is perfectly symmetric, with $\theta_A = -\theta_B$, $U_A = -U_B$, and $R_A = R_B$, then the equations become independent of s and we write them as $\mathbb{E}[r]$, $\mathbb{E}[T]$, etc.

The DDM has a special place in the pantheon of decision models for several reasons, each of which will be discussed further in subsequent sections. First, it has been empirically successful relative to other models (Ratcliff and Smith, 2004). Second, a number of more biologically detailed models can be formally reduced to the DDM under certain assumptions (Bogacz et al., 2006). Third, it is optimal in the sense that for any desired accuracy (or reward rate), there is a setting of the decision thresholds that will achieve the fastest possible decision time (Bogacz et al., 2006). If the decision time is fixed (i.e., the threshold is determined exogenously, as in the interrogation paradigm), the DDM also achieves the highest possible accuracy.

The optimality result is classically known as Wald's sequential probability ratio test (Wald, 1947).

3 *Bounded evidence accumulation in the brain*

Let's revisit the sequential inference task (Yang and Shadlen, 2007) that we described in Chapter 4. On each trial, the subject (a monkey in this case) is presented with a sequence of abstract shapes $x = (x_1, \ldots, x_N)$, and then makes an eye movement to one of two visual targets (labeled A and B). One of the targets (s) is a correct target, producing a water reward when the monkey selects it. The monkey has been trained extensively to learn that each shape x is associated with a log odds $\omega(x) = \log \frac{p(x|s=\mathrm{A})}{p(x|s=\mathrm{B})}$. The probability that the correct target is red, after observing N stimuli, is given by:

$$p(s = \mathrm{A}|x) = \sigma\left(\sum_{n=1}^{N} \omega(x_n)\right). \tag{14}$$

As summarized in Chapter 4, Yang and Shadlen (2007) found that neurons in LIP increased their firing after each shape in proportion to the corresponding the log odds. This is consistent with these neurons acting as evidence accumulators. Considerable additional evidence for this interpretation has come from other perceptual decision tasks, notably the motion direction discrimination task described above (Gold and Shadlen, 2007).

We now address how the log odds gets translated into a decision. The DDM and related models propose that a decision is made when the accumulated evidence crosses a threshold. Evidence for this threshold-crossing event has been found using a free response version of the sequential inference task (Kira et al., 2015). Instead of observing a fixed number of shapes, monkeys can make a decision following any number of shapes. Monkeys consistently made a decision when the evidence exceeded a fixed threshold (Figure 2). Neurally, decisions coincided with a fixed firing rate of LIP neurons, consistent with earlier findings using a motion direction discrimination task (Roitman and Shadlen, 2002).

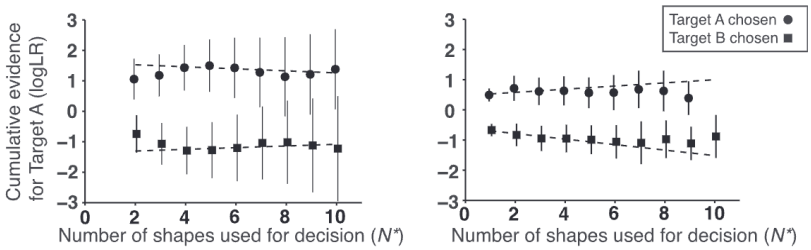

Figure 2: **Cumulative evidence when decisions are made.** The two panels show results for individual monkeys. Reproduced from Kira et al. (2015).

LIP activity recorded by Roitman and Shadlen (2002) is shown in Figure 3. Different motion strength levels (the proportion of coherently moving dots) produced different rates of LIP ramping when activity is aligned to stimulus onset, suggesting that the ramp slope reflects the drift rate in the DDM. Motion-sensitive neurons in area MT, the putative input to the evidence accumulator in LIP for this task, show stronger activity for higher motion strength, but do not ramp up. When LIP activity is aligned to the time of decision, the different ramps converge to a fixed firing rate, suggesting a threshold-crossing event.

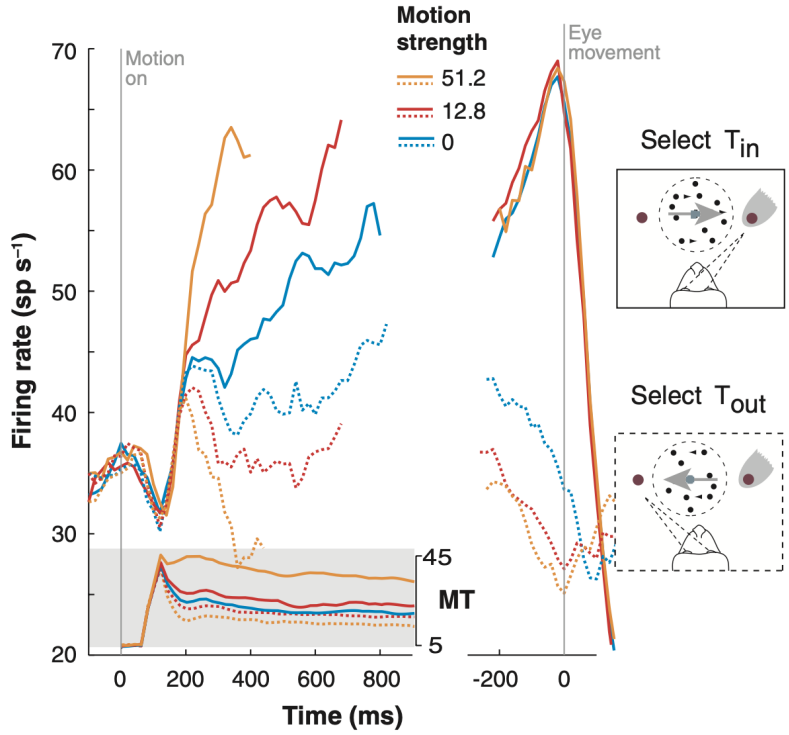

Figure 3: **LIP and MT activity during evidence accumulation and termination.** "T_{in}" (solid lines) denotes trials where the selected target is in the neuron's receptive field; "T_{out}" (dotted lines) denotes trials where the selected target is outside the neuron's receptive field. Data from Roitman and Shadlen (2002) and Britten et al. (1992); figure reproduced from Gold and Shadlen (2007).

How does LIP know when to stop accumulating evidence? This

requires feedback from a downstream area registering the eye movement decision. Among areas responsible for programming eye movements, the superior colliculus is a good candidate for providing the feedback signal because it also sends projections to cortical areas, including LIP. Within the intermediate layer of the superior colliculus, "burst neurons" fire transiently at very high rates immediately before an eye movement (Sparks, 1978). Lo and Wang (2006) proposed that these cells might report threshold-crossing events in LIP (and possibly other cortical areas) and then inhibit the cortical inputs, effectively terminating evidence accumulation. Consistent with this hypothesis, bursts in superior colliculus are triggered by high firing rates in LIP, and inactivation of superior colliculus causes a delay in decision termination (Stine et al., 2023). The inactivation led to paradoxically better performance, because more evidence was accumulated prior to a decision.

The architecture described here is overly simplified; there are several intervening brain areas that mediate the interactions between LIP and superior colliculus.

3.1 Resource constraints on evidence accumulation

In the interrogation paradigm, it is commonly assumed that subjective decision thresholds don't matter; decision time is determined exogenously by the experimenter. One implication of this assumption is that performance should approach perfect accuracy as the decision time gets longer, since eventually the accumulated evidence in favor of the correct decision will overpower the diffusion noise. In fact, human performance often falls short of perfection—why?

One answer to this puzzle is that evidence cannot be accumulated without bound by the brain, for the simple reason that neurons cannot produce arbitrarily high firing rates. At a behavioral level, Ratcliff (2006) has shown that positing "implicit" thresholds provides a good account of accuracy in the interrogation paradigm. The same is true for monkeys (Figure 4). Furthermore, Kiani et al. (2008) showed that LIP neurons do not continue to increase their activity until the response signal, but instead saturate at a fixed level regardless of the decision time.

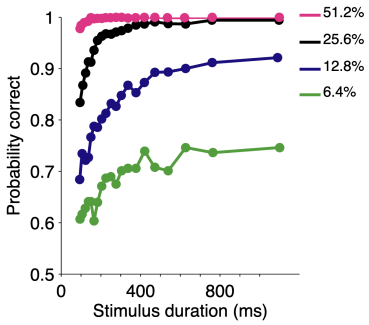

Figure 4: **Performance on motion direction discrimination in the interrogation paradigm.** Reproduced from Kiani et al. (2008).

The idea that the log odds is bounded can also explain many decision phenomena outside of perceptual discrimination tasks (Zhang and Maloney, 2012; Zhang et al., 2020). The general observation is that subjective log odds are attenuated at the extremes, such that people overestimate the probability of rare events and underestimate the probability of frequent events. In essence, all of these phenomena link back to the fundamental constraints on the dynamic range of firing, and hence the information coding capacity of neurons, as discussed in Chapter 3.

3.2 *Leaky, competitive dynamics*

The DDM assumes perfect (albeit noisy) integration—it forgets nothing. Real neurons always have leakage, as reviewed in Chapter 2, because membranes are not perfect insulators. Furthermore, it has been argued that decision circuits in the brain have recurrent connectivity, such that increasing the activity of neurons favoring one decision inhibit (either directly or indirectly) neurons favoring a different decision (Wang, 2008). An influential version of this idea is the leaky competing accumulator (LCA) model (Usher and McClelland, 2001), which models the dynamics of two units, y_1 and y_2, according to:

We can think of the units here as populations of neurons. The activity of a unit represents the pooled firing rate of the population.

$$dy_1 = [-\gamma y_1 - \beta g(y_2) + \mu_1]dt + \kappa dW_1 \tag{15}$$

$$dy_2 = [-\gamma y_2 - \beta g(y_1) + \mu_2]dt + \kappa dW_2, \tag{16}$$

where γ is the leak rate, β controls the mutual inhibition, $g(\cdot)$ is a static nonlinearity, μ_i is the excitatory drive for population i, κ is the noise standard deviation, and $W_i(t)$ is a Wiener process.

This model assumes that the two populations mutually inhibit one another, but it can also approximate a more biologically plausible "pooled inhibition" model (Wang, 2002) in which the two populations send excitatory input to a shared population of inhibitory neurons, which in turn sends inhibition back to the excitatory populations. The pooled inhibition model is more biologically plausible because it obeys *Dale's law*: a neuron performs the same chemical action at all of its synapses. This implies that neurons cannot send excitatory neurotransmitters to some neurons and inhibitory neurotransmitters to others. In this case, the decision populations send excitatory signals to downstream targets in the basal ganglia and subthalamic nucleus, which in turn control motor commands via the superior colliculus and other regions (depending on the response modality). This implies that they cannot directly inhibit one another, though mathematically the mutual and pooled inhibition models can be reduced to a common form under certain assumptions (Bogacz et al., 2006).

Both of these models can be further reduced to the DDM. Here we focus on the mutual inhibition model. As shown by Bogacz et al. (2006), a linearized form of the model, where $g(\cdot)$ is replaced by the identity function, captures the important features of the dynamics if γ and β are sufficiently large. This results in a one-dimensional *Ornstein-Uhlenbeck (O-U) process* (Busemeyer and Townsend, 1993) that depends on the relative excitatory drive and the leak-inhibition difference:

$$dy = [(\mu_1 - \mu_2) + (\beta - \gamma)y]dt + \kappa dW. \tag{17}$$

When leak and inhibition are balanced ($\beta = \gamma$), the model is equivalent to the DDM, with drift rate $\theta = \mu_1 - \mu_2$. As mentioned above, it is precisely in this regime that optimal performance is achieved.

What happens when leak and inhibition are not balanced? If leak is stronger than inhibition ($\gamma > \beta$), the O-U process converges to an equilibrium distribution that is Gaussian-distributed with a mean of $\theta/(\gamma - \beta)$, the fixed point of the dynamics. If inhibition is stronger than leak ($\beta > \gamma$), there is no stable equilibrium; the mean and variance of the process grow exponentially (but note that the process terminates when the threshold is crossed). The leak-dominant regime is interesting because it offers an alternative explanation for suboptimal performance in the interrogation paradigm: if the equilibrium mean is below the decision threshold, then the process relies on noise to cross the threshold. This means that performance will not be perfect even for long decision times.

The static nonlinearity $g(\cdot)$ is potentially more important in the inhibition-dominant regime, because it can stabilize the process.

The leak-dominant regime produces "conservative" behavior in the sense that the dynamics slow down as the activity approaches the fixed point, which lies close to the correct threshold. In contrast, the inhibition-dominant regime produces "risky" behavior in the sense that the dynamics speed up as the activity approaches both the correct and incorrect thresholds. This produces faster responses, but also more errors. Busemeyer and Townsend (1993) suggested that rewards push the imbalance towards inhibition-dominance, whereas punishments push the imbalance towards leak-dominance.

The leak-inhibition imbalance can also affect the influence of sensory evidence at different points in the accumulation process. When leak and inhibition are balanced, evidence at all points in the process are equally weighted. Some studies have found equal weighing of evidence across time in rats and humans (Brunton et al., 2013). When leak is stronger than inhibition, late sensory evidence is more influential (a *recency effect*). When inhibition is stronger than leak, early sensory evidence is more influential (a *primacy effect*). Studies in monkeys have documented a primacy effect: brief motion pulses earlier in the trial have a greater influence on subsequent decisions (Kiani et al., 2008) and LIP activity (Huk and Shadlen, 2005). Primacy has also been reported in studies of humans (Zylberberg et al., 2012; Winkel et al., 2014; Wilming et al., 2020). This can be attributed either to inhibition-dominance or to bounded accumulation. However, neither of these models can explain why early weighting of evidence was present already in MT, the input to the evidence accumulator (Yates et al., 2017). Furthermore, bounded accumulation cannot explain the results of studies finding a recency effect (Usher and McClelland, 2001; Tsetsos et al., 2012; Cheadle et al., 2014).

It's currently unclear how to reconcile these different results. One

possibility is that individual differences are large enough to produce inconsistent results across studies. For example, the studies by Usher and McClelland (2001) and Tsetsos et al. (2012) exhibited substantial heterogeneity across human subjects. The stability of these individual differences across tasks suggests that the heterogeneity cannot be attributed simply to measurement noise (Yoo et al., 2025).

Some studies have identified parametric manipulations that systematically shift evidence weighting between primacy and recency. For example, Carland et al. (2016) found a primacy effect when human subjects were motivated to respond quickly, and a recency effect when subjects were motivated to respond more slowly. This finding is consistent with leak-dominance if one additionally assumes that there is an increasing "urgency" to commit to a decision over the course of a trial (see below for a formalization of this idea). The intuition is that urgency induces reliance on early evidence, producing primacy; with less time pressure, reliance on later evidence increases while the leak causes forgetting of earlier evidence, producing recency.

Monkeys and humans can also shift their temporal weighting strategically, for example when sensory signals are more informative early vs. late in a trial (Levi et al., 2018).

3.3 Why is evidence accumulation leaky?

If the non-leaky DDM achieves optimal performance, then why does evidence accumulation in the brain appear to be (at least sometimes) leaky? One answer is that this is just an unavoidable property of biology. While this might be true at a single-neuron level, networks of neurons can resist leak through recurrent connectivity, producing "reverberating" activity that outlives the time constant of any individual neurons (see for example Wang, 2002).

A different kind of answer is that leak is actually useful. To see why, consider what happens when the hidden state is not fixed over time but instead switches intermittently. In this case, an optimal decision circuit should "forget" the evidence it accumulated prior to a switch (Kilpatrick et al., 2019). In the case where the circuit doesn't have direct access to the switch times, it should forget gradually—precisely what is accomplished by a leaky accumulator. In support of this interpretation, humans (Glaze et al., 2015) and rats (Piet et al., 2018) adopt leakier accumulation when the switch rate is higher.

4 Thresholds that optimize reward rate

A natural optimization objective for the free response paradigm is the reward rate, defined as the ratio of expected reward to expected trial

Here we have assumed perfect symmetry, as described above ($\theta = A = -\theta = B$, $U_A = -U_B$, and $R_A = R_B$).

duration:

$$\bar{r} = \frac{\mathbb{E}[r]}{\mathbb{E}[T] + T_0 + \text{ITI}}, \tag{18}$$

where ITI is the intertrial interval (the time between the response and the next trial. The problem facing the agent is how to set the thresholds to optimize reward rate. While there is no closed-form solution to this problem, we can gain insight into the structure of the solution by inspecting the reward rate as a function of the threshold (Figure 5). Generally speaking, reward rate is optimized by choosing a higher threshold when the evidence is weaker. Intuitively, this is necessary to avoid making erroneous decisions based on noise.

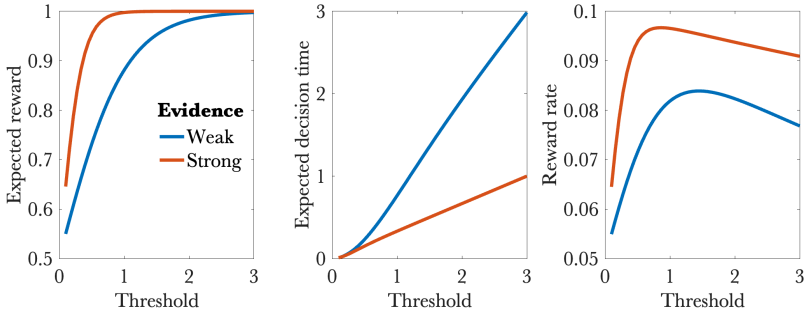

Figure 5: **The effect of threshold on reward, decision time, and reward rate**. Weak evidence corresponds to $\theta = 1$; strong evidence corresponds to $\theta = 3$. This simulation assumed the following parameters: ITI$= 10$, $T_0 = 0$, $\nu = 1$, $R = 1$.

4.1 *Do humans optimize reward rate?*

Bogacz et al. (2006) derived a performance characterization that does not depend on the choice of parameters (drift rate, noise variance)—what they termed the *optimal performance curve*:

$$\frac{\mathbb{E}[T]}{T_0 + \text{ITI}} = \left[\frac{1}{P_e \log \frac{1-P_e}{P_e}} + \frac{1}{1-2P_e} \right]^{-1}, \tag{19}$$

where P_e is the error rate (the proportion of trials on which the wrong decision is made). The quantity on the left-hand-side is the *normalized decision time*. The optimal performance curve predicts that decision times will be fastest when the error rate is very large and when it is very small. When the evidence is very weak relative to the noise level, the threshold should be set low because the stimulus carries little information about the optimal decision, making it more advantageous to move as quickly as possible to the next trial. When the evidence is very strong relative to the noise level, the threshold should again be set low, in this case because the stimulus carries substantial information about the optimal decision—the decision is easy.

Bogacz et al. (2010) compared human performance on perceptual discrimination tasks to the optimal performance curve, finding that the 30% of subjects achieving the lowest error rates matched the theoretical optimum quite well (Figure 6). Why were most subjects suboptimal?

One possibility is that the optimization process is limited by timing uncertainty. It is well known that the perception of elapsed time is subject to irreducible noise. In the context of reward rate optimization, timing uncertainty means that humans may not have direct access to temporal variables such as the ITI, the non-decision time, and decision time. Zacksenhouse et al. (2010) proposed that some subjects adopt a *maximin* strategy, attempting to maximize the worst reward rate given their level of timing uncertainty. This yields a rescaled optimal performance curve, with longer decision times compared to reward rate optimization without timing uncertainty. The maximin strategy does a good job explaining why most subjects tended to have surprisingly long decision times.

Consistent with the timing uncertainty theory, Balci et al. (2011) showed that that individual differences in timing ability (assessed by an independent test) are correlated with deviations from the optimal performance curve. This suggests that individuals take into account their abilities when setting thresholds.

Threshold optimization could be implemented through an iterative adjustment process, where errors induce increments in the threshold and correct decisions induce decrements in the threshold (see Simen et al., 2006, for a more sophisticated version of this idea). Indeed, humans become increasingly close to the optimal performance curve with practice (Balci et al., 2011). Iterative adjustment also explains more fine-grained patterns across trials: after an error, humans tend to slow down on the next trial, making them more accurate (e.g., Rabbitt, 1966; Purcell and Kiani, 2016).

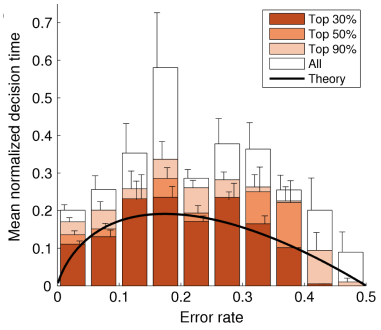

Figure 6: **The optimal performance curve compared to human performance**. Data from Bogacz et al. (2010); figure reproduced from Zacksenhouse et al. (2010).

4.2 *Collapsing thresholds and urgency gating*

The optimality analysis conducted by Bogacz et al. (2006) models the case where task difficulty and other parameters do not vary across trials, and thus the agent (with sufficient training) can be expected to know these parameters. In many experimental tasks (and in real life) these parameters do in fact vary and are not necessarily known in advance. Drugowitsch et al. (2012) developed a theoretical framework for the more general case where the agent has uncertainty about θ, expressed as a posterior probability distribution $p(\theta|X(t))$, where $X(t)$ represents all of the evidence up to time t. The posterior probability over alternatives, $b(t) = p(s = A|X(t))$, can then be obtained

by:

$$b(t) = p(\theta > 0 | X(t)) = \int_0^\infty p(\theta | X(t)) d\theta. \tag{20}$$

Here again we are making symmetry assumptions, as described above.

Assuming a Gaussian prior on the drift rate, $\theta \sim \mathcal{N}(0, \chi)$, the belief state is a sigmoidal function of the accumulated evidence $L(t)$, with a slope that depends on elapsed time (this will become important later when we discuss confidence), and the optimal threshold for reward rate maximization becomes a function of time:

$$b(t) = \Phi\left(\frac{L(t)}{\sqrt{t + \chi^{-1}}}\right) \tag{21}$$

$$U^*(t) = \sqrt{t + \chi^{-1}} \Phi^{-1}(b(t)), \tag{22}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard Gaussian distribution and $\Phi^{-1}(\cdot)$ is its inverse. We can see that the threshold collapses over time. The intuition is that on really difficult trials, it's better to move on to the next trial rather than try to collect sufficient evidence for the current trial. The collapsing threshold forces the agent to make a decision with less evidence.

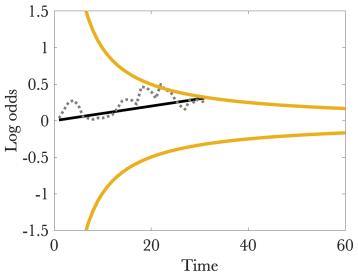

Figure 7: **The drift-diffusion model with collapsing thresholds**.

Fudenberg et al. (2018) analyzed a closely related problem which allowed them to obtain analytical results. They model an agent maximizing expected reward penalized by a linear time cost:

$$\pi^* = \underset{\pi}{\operatorname{argmax}}\, \mathbb{E}[r - \alpha T | \pi], \tag{23}$$

where $\alpha > 0$ is a time cost coefficient. They derive an asymptotic approximation of the optimal threshold for large t and κ (Figure 7):

Recall that κ is the diffusion noise standard deviation.

$$U^*(t) \approx \frac{1}{2\alpha(\chi^{-1} + \kappa^{-2} t)}. \tag{24}$$

According to this expression, the threshold is a hyperbolic function of time (i.e., declining with rate $1/t$), collapsing more quickly when signals are noisier (large κ). This fits with the intuition that more difficult trials should induce a lower threshold. Two other features of this expression are worth noting. First, thresholds are lower when the time cost is greater (large α), capturing an agent's disinclination to deliberate. Second, thresholds are lower when the agent is more certain *a priori* that the drift rate is close to the prior mean of 0 (small χ), again capturing the intuition that more difficult decisions should induce a lower threshold.

The empirical data supporting collapsing thresholds is mixed. Based on quantitative model fits to behavioral data, some studies have found support for collapsing thresholds (Palestro et al., 2018; Bhui, 2019), and others have not (Hawkins et al., 2015; Voskuilen

et al., 2016). In some cases, collapsing thresholds can be incentivized by emphasizing decision speed, but emphasizing reward rate maximization appears to be insufficient (Evans et al., 2019).

One reason for these mixed results might be that the optimality analysis is more subtle than it first appears. Malhotra et al. (2018) showed that collapsing thresholds are only optimal when some trials are very difficult; if these trials are absent, the optimal thresholds are constant or can even increase. Moreover, even when collapsing thresholds are theoretically optimal, there are many decision problems in which a fixed threshold can achieve a comparable reward rate (Boehm et al., 2020).

A more fundamental objection to time-varying thresholds is that they seem to be inconsistent with the neural evidence (reviewed above) that decisions are made when firing rates putatively representing accumulated evidence cross a fixed threshold is crossed. One possible reconciliation is to convert the temporal dynamics of the threshold into a separate time-varying "urgency" signal $\Omega(t)$ that modulates the evidence accumulator $L(t)$, effectively pushing it to terminate earlier (Ditterich, 2006; Churchland et al., 2008; Standage et al., 2011; Thura et al., 2012). The urgency signal does not depend on the evidence itself, only the passage of time. For example, a linear urgency signal takes the form $\Omega(t) = m_0 + mt$ with parameters m and m_0. In some models (e.g., Ditterich, 2006) the urgency signal modulates the accumulator additively, $\tilde{L}(t) = \Omega(t) + L(t)$, whereas in others (e.g., Standage et al., 2011; Thura et al., 2012) it modulates the accumulator multiplicatively, $\tilde{L}(t) = \Omega(t)L(t)$. Smith and Ratcliff (2021) showed that the multiplicative urgency model is (under certain assumptions) equivalent to a collapsing threshold model. Thus, multiplicative urgency may be a neurally plausible implementation of optimal time-varying thresholds which maintains the core ideas of the DDM (accumulation of a scalar decision variable to a threshold).

As with studies of collapsing thresholds, the behavioral evidence for urgency models is also mixed (see Trueblood et al., 2020; Smith and Ratcliff, 2021).

The urgency signal is thought to take the form of a temporal ramp, similar to the representation of elapsed time by some models of interval timing (e.g., Durstewitz, 2004; Simen et al., 2011), and attested by neural recordings of some time-encoding neurons (e.g., Niki and Watanabe, 1979; Jazayeri and Shadlen, 2015; Cao et al., 2024). Urgency models require the evidence accumulator to be tuned by this temporal ramp (either additively or multiplicatively). Thura and Cisek (2014) identified neurons in motor areas that appear to reflect an urgency-modulated evidence accumulator: neural activity increased in response to sensory evidence, while for any fixed evidence level the activity was higher closer to the time of decision. This pattern appears most consistent with an additive urgency signal.

Studies of LIP neurons have also identified an evidence-independent

urgency component (Churchland et al., 2008; Hanks et al., 2014). This component was extracted by averaging the neural time series across all stimulus categories, revealing a build-up activity across time that terminates when a decision is made. Hanks et al. (2014) showed that emphasizing accuracy over speed (by slightly delaying reward delivery) reduced the slope of the build-up. Another reflection of a putative urgency signal is the reduced build-up slope following an error (Purcell and Kiani, 2016)—a neural analogue of the behavioral post-error slowing phenomenon described earlier. These findings support the hypothesis that an urgency signal can adaptively optimize reward rate under different conditions by modulating evidence accumulation.

Note that extracting urgency by averaging can only identify an additive component.

See Heitz and Schall (2012) for evidence of slope modulation by speed vs. accuracy in the frontal eye field, a prefrontal area involved in programming of eye movements.

5 *Confidence*

Accumulated evidence can support judgments of confidence, which serve a variety of functions, including calibrating time/energy investment, managing risk preferences, and controlling learning rates. When evidence strength is fixed, the posterior probability that a decision is correct is a sigmoidal function of the accumulated evidence $L(t)$ at the time of decision (see section 1). The *Bayesian confidence hypothesis* states that subjective judgments of confidence reflect the posterior probability of a correct decision (Aitchison et al., 2015; Meyniel et al., 2015).

If the Bayesian confidence hypothesis is correct, then the state of the evidence accumulator at the time of a decision should monotonically predict confidence reports. To study this question in monkeys, Kiani and Shadlen (2009) developed a post-decision wagering task in which monkeys were first shown a standard random dot motion stimulus and then either chose between two motion directions or were additionally given a "safe" option which effectively allowed them to opt out of the discrimination task on that trial. Monkeys should choose the safe option when their confidence is low. Consistent with this interpretation, the probability of choosing the safe option decreased with stimulus duration and increased with motion coherence—two factors that should increase confidence. The probability of choosing the safe option was also associated with firing rates in LIP that were intermediate between the firing rates associated with the two motion direction choices, consistent with the hypothesis that low confidence reflects a level of accumulated evidence that is far from the two decision thresholds.

Despite its parsimony and simplicity, the Bayesian confidence hypothesis has been challenged in studies of human perceptual decisions. For example, Adler and Ma (2018) found that human con-

fidence reports are better described by heuristic models that use sensory uncertainty suboptimally. These models apply a linear or quadratic transformation of the sensory signals and uncertainty into confidence, approximating Bayes-optimal confidence reports but nonetheless deviating from them quantitatively.

See Xue et al. (2024) for qualitative behavioral evidence against the Bayesian confidence hypothesis.

A more fundamental problem for the Bayesian confidence hypothesis, at least in the context of models like the DDM, is that if decision are always made when a threshold is crossed, then the accumulated evidence should always have the same value and hence reported confidence should always be the same. This is contradicted by the empirical finding that confidence is higher when responses are faster (Festinger, 1943; Vickers and Packer, 1982). However, it is important to remember that most experiments use mixed levels of evidence strength across trials, which means that subjects have uncertainty about evidence strength on each trial. In this case, the posterior probability that a decision is correct depends on elapsed time, such that a longer response time indicates lower Bayesian confidence, even for identical levels of accumulated evidence (Calder-Travis et al., 2024). Thus, the Bayesian confidence hypothesis can be at least partially rescued by considering the more complex nature of the decision problem in mixed evidence tasks.

6 Conclusion

The study of two-alternative perceptual decisions has been an extremely fruitful area of research, fostering the development of highly accurate quantitative models. Nonetheless, it is important to consider some limitations and challenges of this setup. Many decision problems involve additional complexities: time-varying hidden states, larger action spaces, state-dependent rewards, and so on. Some of these complexities have been addressed theoretically and experimentally; in the next chapter we will begin to develop a more general theoretical framework.

Another challenge is that some of the canonical computations discussed above, such as the role of LIP as an evidence accumulator, run into stiff empirical contradictions. For example, one study (Katz et al., 2016) found that inactivation of LIP surprisingly had no effect on decision making performance. Another study (Latimer et al., 2015) argued that many LIP neurons do not ramp up gradually with evidence, but instead exhibit discrete jumps. While the interpretation of these findings continue to be hotly debated, it should not be shocking that the simplest evidence accumulation models fail to be comprehensive accounts of decision making in the brain. Their usefulness, as is the case for all simple models, lies in exposing a set of

computational principles which we can use as the foundation for more comprehensive accounts.

Study questions

1. Compare how the DDM, leaky competing accumulator model, and urgency gating model account for the speed-accuracy trade-off. What are the distinctive computational and neural mechanisms each emphasizes?
2. Studies report saturation in LIP firing rates. How does this observation provide a neural basis for "implicit" thresholds, and how might this connect to resource-rational models?
3. Some studies found that inactivation of LIP (Katz et al., 2016) or discrete jumps (Latimer et al., 2015) in neural activity contradict the ramping evidence accumulator view of LIP. How might computational neuroscientists reconcile these discrepancies?

References

Adler, W. T. and Ma, W. J. (2018). Comparing Bayesian and non-Bayesian accounts of human confidence reports. *PLoS Computational Biology*, 14:e1006572.

Aitchison, L., Bang, D., Bahrami, B., and Latham, P. E. (2015). Doubly Bayesian analysis of confidence in perceptual decision-making. *PLoS Computational Biology*, 11:e1004519.

Balci, F., Simen, P., Niyogi, R., Saxe, A., Hughes, J. A., Holmes, P., and Cohen, J. D. (2011). Acquisition of decision making criteria: reward rate ultimately beats accuracy. *Attention, Perception, & Psychophysics*, 73:640–657.

Bhui, R. (2019). A statistical test for the optimality of deliberative time allocation. *Psychonomic Bulletin & Review*, 26:855–867.

Bitzer, S., Park, H., Blankenburg, F., and Kiebel, S. J. (2014). Perceptual decision making: drift-diffusion model is equivalent to a Bayesian model. *Frontiers in Human Neuroscience*, 8:102.

Boehm, U., van Maanen, L., Evans, N. J., Brown, S. D., and Wagenmakers, E.-J. (2020). A theoretical analysis of the reward rate optimality of collapsing decision criteria. *Attention, Perception, & Psychophysics*, 82:1520–1534.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113:700–765.

Bogacz, R., Hu, P. T., Holmes, P. J., and Cohen, J. D. (2010). Do humans produce the speed–accuracy trade-off that maximizes reward rate? *Quarterly Journal of Experimental Psychology*, 63:863–891.

Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, 12:4745–4765.

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

Busemeyer, J. and Townsend, J. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100:432–459.

Calder-Travis, J., Charles, L., Bogacz, R., and Yeung, N. (2024). Bayesian confidence in optimal decisions. *Psychological review*, 131:1114–1160.

Cao, R., Bright, I. M., and Howard, M. W. (2024). Ramping cells in the rodent medial prefrontal cortex encode time to past and future events via real Laplace transform. *Proceedings of the National Academy of Sciences*, 121(38):e2404169121.

Carland, M. A., Marcos, E., Thura, D., and Cisek, P. (2016). Evidence against perfect integration of sensory information during perceptual decision making. *Journal of Neurophysiology*, 115(2):915–930.

Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., De Gardelle, V., Castañón, S. H., and Summerfield, C. (2014). Adaptive gain control during human perceptual choice. *Neuron*, 81:1429–1441.

Churchland, A. K., Kiani, R., and Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience*, 11:693–702.

Ditterich, J. (2006). Evidence for time-variant decision making. *European Journal of Neuroscience*, 24:3628–3641.

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

Durstewitz, D. (2004). Neural representation of interval time. *Neuroreport*, 15:745–749.

Evans, N., Hawkins, G., and Brown, S. (2019). The role of passing time in decision-making. *Journal of Experimental psychology. Learning, Memory, and Cognition*, 46:316–326.

Festinger, L. (1943). Studies in decision: I. Decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. *Journal of Experimental Psychology*, 32:291–306.

Fudenberg, D., Strack, P., and Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108:3651–3684.

Glaze, C. M., Kable, J. W., and Gold, J. I. (2015). Normative evidence accumulation in unpredictable environments. *Elife*, 4:e08825.

Gold, J. I. and Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30:535–574.

Hanks, T., Kiani, R., and Shadlen, M. N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. *Elife*, 3:e02260.

Hawkins, G. E., Forstmann, B. U., Wagenmakers, E.-J., Ratcliff, R., and Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience*, 35:2476–2484.

Heitz, R. P. and Schall, J. D. (2012). Neural mechanisms of speed-accuracy tradeoff. *Neuron*, 76:616–628.

Huk, A. C. and Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience*, 25:10420–10436.

Jazayeri, M. and Shadlen, M. N. (2015). A neural mechanism for sensing and reproducing a time interval. *Current Biology*, 25:2599–2609.

Katz, L. N., Yates, J. L., Pillow, J. W., and Huk, A. C. (2016). Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature*, 535:285–288.

Kiani, R., Hanks, T. D., and Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *Journal of Neuroscience*, 28:3017–3029.

Kiani, R. and Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, 324:759–764.

Kilpatrick, Z. P., Holmes, W. R., Eissa, T. L., and Josić, K. (2019). Optimal models of decision-making in dynamic environments. *Current Opinion in Neurobiology*, 58:54–60.

Kira, S., Yang, T., and Shadlen, M. N. (2015). A neural implementation of Wald's sequential probability ratio test. *Neuron*, 85:861–873.

Latimer, K. W., Yates, J. L., Meister, M. L., Huk, A. C., and Pillow, J. W. (2015). Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science*, 349:184–187.

Levi, A. J., Yates, J. L., Huk, A. C., and Katz, L. N. (2018). Strategic and dynamic temporal weighting for perceptual decisions in humans and macaques. *ENeuro*, 5.

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

Malhotra, G., Leslie, D. S., Ludwig, C. J., and Bogacz, R. (2018). Time-varying decision boundaries: insights from optimality analysis. *Psychonomic Bulletin & Review*, 25:971–996.

Meyniel, F., Sigman, M., and Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, 88:78–92.

Niki, H. and Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Research*, 171:213–224.

Palestro, J. J., Weichart, E., Sederberg, P. B., and Turner, B. M. (2018). Some task demands induce collapsing bounds: Evidence from a behavioral analysis. *Psychonomic Bulletin & Review*, 25:1225–1248.

Piet, A. T., El Hady, A., and Brody, C. D. (2018). Rats adopt the optimal timescale for evidence integration in a dynamic environment. *Nature Communications*, 9:4265.

Purcell, B. A. and Kiani, R. (2016). Neural mechanisms of post-error adjustments of decision policy in parietal cortex. *Neuron*, 89:658–671.

Rabbitt, P. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, 71:264–272.

Ratcliff, R. (2006). Modeling response signal and response time data. *Cognitive Psychology*, 53:195–237.

Ratcliff, R. and Smith, P. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111:333–367.

Roitman, J. D. and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, 22:9475–9489.

Simen, P., Balci, F., deSouza, L., Cohen, J. D., and Holmes, P. (2011). A model of interval timing by neural integration. *Journal of Neuroscience*, 31:9238–9253.

Simen, P., Cohen, J. D., and Holmes, P. (2006). Rapid decision threshold modulation by reward rate in a neural network. *Neural Networks*, 19:1013–1026.

Smith, P. and Ratcliff, R. (2021). Modeling evidence accumulation decision processes using integral equations: Urgency-gating and collapsing boundaries. *Psychological Review*, 129:235–267.

Sparks, D. L. (1978). Functional properties of neurons in the monkey superior colliculus: coupling of neuronal activity and saccade onset. *Brain Research*, 156:1–16.

Standage, D., You, H., Wang, D.-H., and Dorris, M. C. (2011). Gain modulation by an urgency signal controls the speed–accuracy trade-off in a network model of a cortical decision circuit. *Frontiers in Computational Neuroscience*, 5:7.

Stine, G. M., Trautmann, E. M., Jeurissen, D., and Shadlen, M. N. (2023). A neural mechanism for terminating decisions. *Neuron*, 111:2601–2613.

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

Thura, D. and Cisek, P. (2014). Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making. *Neuron*, 81:1401–1416.

Trueblood, J., Heathcote, A., Evans, N., and Holmes, W. (2020). Urgency, leakage, and the relative nature of information processing in decision-making. *Psychological Review*, 128:160–186.

Tsetsos, K., Gao, J., McClelland, J. L., and Usher, M. (2012). Using time-varying evidence to test models of decision dynamics: bounded diffusion vs. the leaky competing accumulator model. *Frontiers in Neuroscience*, 6.

Usher, M. and McClelland, J. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, 108:550–592.

Vickers, D. and Packer, J. (1982). Effects of alternating set for speed or accuracy on response time, accuracy and confidence in a unidimensional discrimination task. *Acta Psychologica*, 50:179–197.

Voskuilen, C., Ratcliff, R., and Smith, P. L. (2016). Comparing fixed and collapsing boundary versions of the diffusion model. *Journal of Mathematical Psychology*, 73:59–79.

Wald, A. (1947). Sequential analysis.

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36:955–968.

Wang, X.-J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, 60:215–234.

Wilming, N., Murphy, P. R., Meyniel, F., and Donner, T. H. (2020). Large-scale dynamics of perceptual decision information across human cortex. *Nature Communications*, 11:5109.

Winkel, J., Keuken, M. C., van Maanen, L., Wagenmakers, E.-J., and Forstmann, B. U. (2014). Early evidence affects later decisions: Why evidence accumulation is required to explain response time data. *Psychonomic Bulletin & Review*, 21:777–784.

Xue, K., Shekhar, M., and Rahnev, D. (2024). Challenging the Bayesian confidence hypothesis in perceptual decision-making. *Proceedings of the National Academy of Sciences*, 121:e2410487121.

Yang, T. and Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, 447:1075–1080.

Yates, J. L., Park, I. M., Katz, L. N., Pillow, J. W., and Huk, A. C. (2017). Functional dissection of signal and noise in MT and LIP during decision-making. *Nature Neuroscience*, 20:1285–1292.

Yoo, M., Bahg, G., Turner, B., and Krajbich, I. (2025). People display consistent recency and primacy effects in behavior and neural activity across perceptual and value-based judgments. *Cognitive, Affective, & Behavioral Neuroscience*, pages 1–18.

Zacksenhouse, M., Bogacz, R., and Holmes, P. (2010). Robust versus optimal strategies for two-alternative forced choice tasks. *Journal of Mathematical Psychology*, 54:230–246.

Zhang, H. and Maloney, L. T. (2012). Ubiquitous log odds: a common representation of probability and frequency distortion in perception, action, and cognition. *Frontiers in Neuroscience*, 6:1.

Zhang, H., Ren, X., and Maloney, L. T. (2020). The bounded rationality of probability distortion. *Proceedings of the National Academy of Sciences*, 117:22024–22034.

Zylberberg, A., Barttfeld, P., and Sigman, M. (2012). The construction of confidence in a perceptual decision. *Frontiers in Integrative Neuroscience*, 6:79.