Article

Neuron

A Neural Implementation of Wald's Sequential Probability Ratio Test

Highlights

- Monkeys can make decisions based on a sequence of symbols
- More reliable symbols have more influence on the decision and neuronal activity
- Choice and reaction time are explained by a threshold on the neural evidence
- The mechanism approximates an optimal statistical procedure for decision making

Authors

Shinichiro Kira, Tianming Yang, Michael N. Shadlen

Correspondence

shadlen@columbia.edu

In Brief

An optimal statistical test for deciding between two hypotheses is to accumulate evidence in units of log odds until reaching a positive or negative criterion. Kira et al. show that neurons in the parietal cortex of monkeys approximate this strategy.

Kira et al., 2015, Neuron *85*, 861–873
February 18, 2015
http://dx.doi.org/10.1016/j.neuron.2015.01.007

CellPress

Neuron
Article

A Neural Implementation of Wald's Sequential Probability Ratio Test

Shinichiro Kira,[1,2,7,8] Tianming Yang,[3,7] and Michael N. Shadlen[2,4,5,6,*]
[1]Neurobiology & Behavior Program, University of Washington, Seattle, WA 98195, USA
[2]Department of Neuroscience, Columbia University, College of Physicians and Surgeons, New York, NY 10032, USA
[3]Institute of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China
[4]Howard Hughes Medical Institute
[5]Zuckerman Mind Brain Behavior Institute
[6]Kavli Institute of Brain Science
Columbia University, College of Physicians and Surgeons, New York, NY 10032, USA
[7]Co-first author
[8]Present address: Department of Neurobiology, Harvard Medical School, 210 Longwood Avenue, Boston, MA 02115, USA
*Correspondence: shadlen@columbia.edu
http://dx.doi.org/10.1016/j.neuron.2015.01.007

SUMMARY

Difficult decisions often require evaluation of samples of evidence acquired sequentially. A sensible strategy is to accumulate evidence, weighted by its reliability, until sufficient support is attained. An optimal statistical approach would accumulate evidence in units of logarithms of likelihood ratios (logLR) to a desired level. Studies of perceptual decisions suggest that the brain approximates an analogous procedure, but a direct test of accumulation, in units of logLR, to a threshold in units of cumulative logLR is lacking. We trained rhesus monkeys to make decisions based on a sequence of evanescent, visual cues assigned different logLR, hence different reliability. Firing rates of neurons in the lateral intraparietal area (LIP) reflected the accumulation of logLR and reached a stereotyped level before the monkeys committed to a decision. The monkeys' choices and reaction times, including their variability, were explained by LIP activity in the context of accumulation of logLR to a threshold.

INTRODUCTION

Complex decision making often requires the collection of multiple pieces of evidence before committing to a choice. Along the way, the brain must evaluate each piece of evidence, combine them together, and determine whether more evidence is needed. The process can be studied at the neural level by training an animal to indicate its decisions with an eye movement. In that case, neurons in the parietal and prefrontal cortex, which are associated with response selection, represent the accumulating evidence during deliberation (Kim and Shadlen, 1999; Shadlen et al., 1996; Shadlen and Newsome, 2001). The same neurons achieve a stereotyped level of firing rate upon completion of the decision (Churchland et al., 2008; Ding and Gold, 2012; Roitman and Shadlen, 2002). Thus, these neurons are thought to participate in the conversion of evidence to a decision variable (DV) suitable for comparison to a threshold (or bound) for terminating the decision process with a choice. Although the underlying neural mechanisms are less well understood, a similar "bounded evidence accumulation" framework explains a variety of perceptual and mnemonic decisions in animals and humans (O'Connell et al., 2012; Ratcliff and McKoon, 2008; Shadlen and Kiani, 2013).

The idea is appealing because the accumulation of evidence might be likened to the evolution of belief in a proposition. However, the concept presupposes that the brain possesses a mechanism to convert sensory evidence into probabilistic values associated with degree of belief (Gold and Shadlen, 2001; Pouget et al., 2013). It has been shown that humans and nonhuman primates rationally combine simultaneous cues in accordance with their reliability (Ernst and Banks, 2002; Fetsch et al., 2012; Jacobs, 1999; Knill, 2007). However, such rational combination of cues has not been studied extensively in the setting of decision making from a sequence of cues that are separated in time. This is because most studies of decision making employ a single stimulus whose reliability is fixed (i.e., statistically stationary) over the course of a decision.

To overcome this limitation, we previously trained monkeys to observe a sequence of shape cues that furnished probabilistic evidence bearing on a binary decision (Yang and Shadlen, 2007). This study showed that the monkeys based their decisions on the combined evidence from four cues, giving more weight to the more reliable cues. Moreover, as the shapes appeared sequentially during a trial, the firing rates of neurons in area LIP tracked the running sum of the evidence, in units proportional to log likelihood ratio (logLR), for and against the choice alternatives. This suggests that the brain can optimally combine cues from sequential samples. However, two aspects of this study preclude a direct connection to the "bounded evidence accumulation" mechanism mentioned above. First, there was no measure of decision termination (e.g., reaction time), because four shape cues were shown on each trial. Second, although the

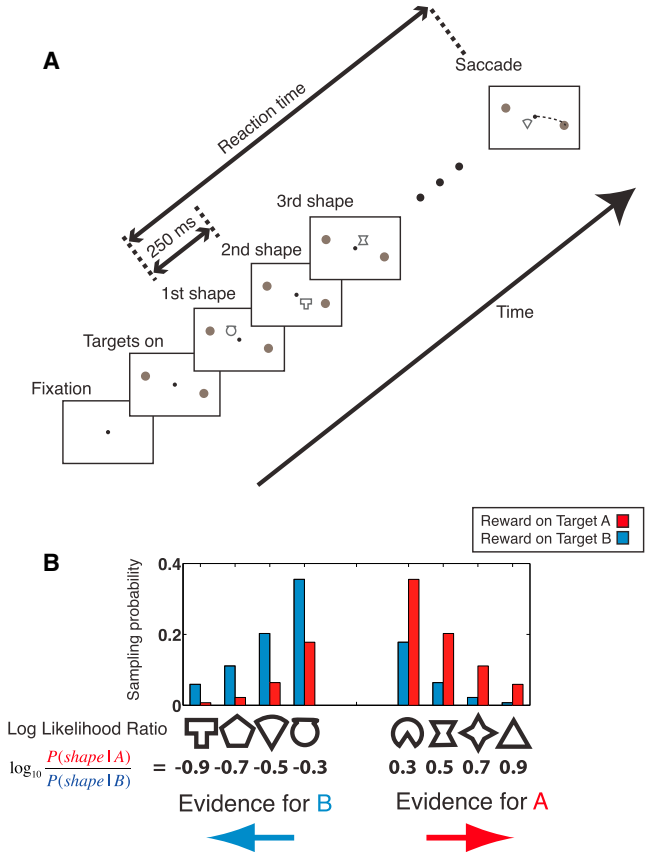

Figure 1. Sequential Inference Task

(A) Trial flow. The monkey stares at a central fixation point at the beginning of a trial, and two choice targets appear, followed by a sequence of shapes, which are shown in succession every 250 ms. Each shape supplies evidence bearing on whether a reward is associated with one or the other choice target. The sequence continues until the monkey initiates an eye movement to a choice target. The reaction time is the interval from the onset of the first shape to initiation of the saccadic choice.

(B) The shapes shown on a trial are independent, random draws from either of the two discrete sampling distributions (indicated by color). On each trial, the designation of the correct (rewarded) choice is randomized to Target A or B, and the shapes are sampled accordingly. The ratio of sampling probabilities associated with each shape implies that half of the shapes support one or the other choice. Target A refers to the left choice target for monkey E and the red choice target for monkey J.

animal based its choices on the cumulative evidence from the four shapes, there was no actual requirement to integrate evidence in time. This is because each shape remained visible from the time it was presented until the monkey made a decision. Thus, it was possible that the monkeys based their decision on the combination of the four cues present at the end of each trial.

Here we employed a modified version of this probabilistic classification task in which a sequence of shape-cues are presented transiently, until the monkey terminates the sequence with a decision. The task makes explicit demands on working memory and evidence accumulation. Moreover, to perform this task optimally, the monkey should terminate decisions when the accumulated logLR reaches a threshold level, or bound. This process, termed the sequential probability ratio test (SPRT) (Barnard, 1946; Good, 1979; Wald, 1947), is optimal in the sense that it requires the least number of samples, on average, to achieve any given level of accuracy (Wald and Wolfowitz, 1948). We hypothesized that the primate brain approximates such a stratagem, and we demonstrate neural correlates of this process in the firing rates of neurons in area LIP.

RESULTS

Behavior

Monkeys made decisions between two peripheral targets based on a sequential presentation of shapes (Figure 1A, see Experimental Procedures). Each presented shape conferred probabilistic evidence that differed in reliability in predicting which target (A or B) would furnish a reward. Although each of the eight shapes could be shown on any trial, they were sampled differently depending on the reward-associated target (Figure 1B). Four shapes favored one of the targets ("Target A") because they were more likely to appear when it was assigned the reward. The other four shapes favored the other target. Each shape's evidence can be quantified by logLR: the sign indicates a favored target and the magnitude represents the reliability of evidence. While the monkey maintained fixation, shapes appeared parafoveally every 250 ms, one at a time, until the monkey indicated its decision with a saccadic eye movement to one of the two targets (Figure 1A). The monkey was rewarded for correct choices after a fixed or variable delay. For the two monkeys, target designations "A" and "B" refer to left and right (monkey E) or red and green (monkey J).

As shown in Figure 2A, both monkeys displayed a wide range of reaction times (RTs) on this task. The serrated quality of the histograms is consistent with the rate at which the shapes were shown (4 Hz), although there is considerable variability that blurs this discrete schedule. Evidently, monkey J made some decisions based on only a few shapes, but both monkeys often delayed their responses until after viewing several shapes. We therefore wished to ascertain whether they used information from all the shapes or waited for a particular shape or perhaps used information from a favored time period.

To address this empirically, we used logistic regression to estimate the leverage of shapes presented at the beginning through the end of the trial on the choice. To isolate the leverage of shapes presented at particular times, we considered all the other shapes shown in the trial as confounders (Equations 1 and 2). The leverage is summarized by a single weighting coefficient (β_1, Equation 2), displayed in Figure 2B. The left half of these graphs displays the effect of evidence presented at the beginning of the trial on the monkey's ultimate choice. Because the shapes appeared (and disappeared) every 250 ms, this portion of the graph contains only three discrete values, corresponding to the onset times of the first three shapes. Shapes presented in each of these epochs had coefficients significantly greater than zero, indicating that they affected the monkeys' choices ($p < 0.05$ up to the tenth shape for each monkey).

The right half of the graph plots the effect of the shapes on choice at the time the shapes appeared relative to the saccadic choice. Since saccade initiation is not fixed to the time of

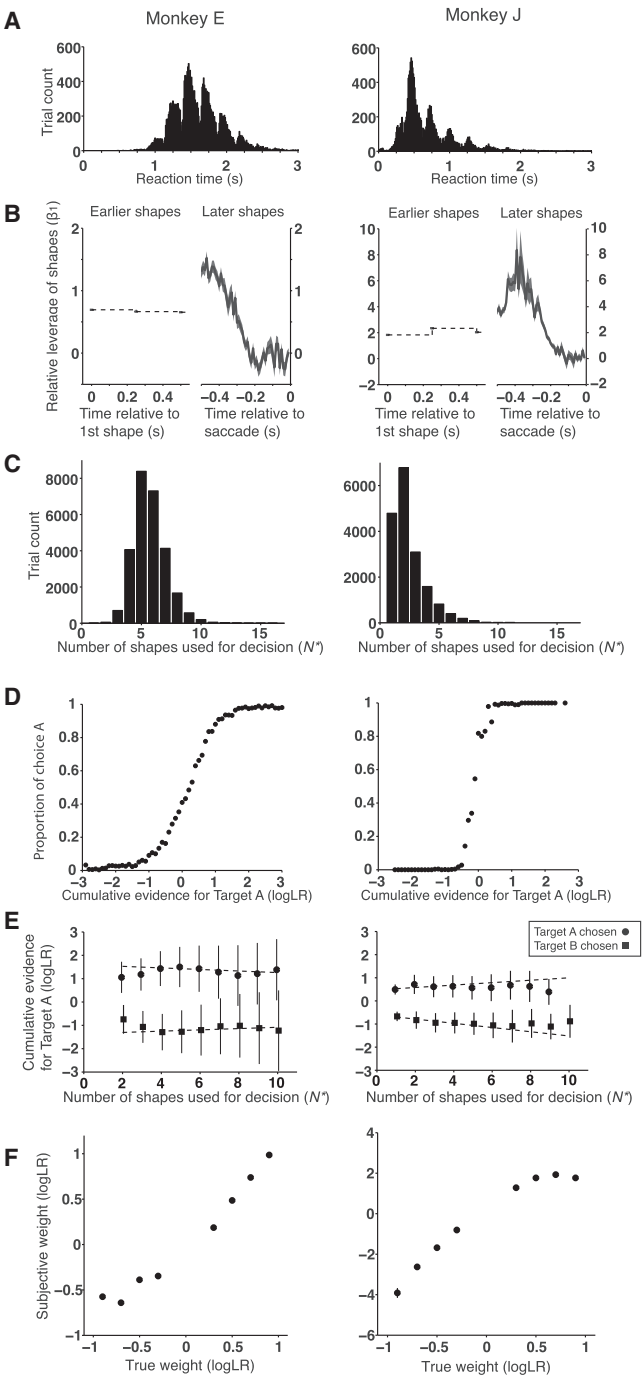

Figure 2. Behavior

Data from the two monkeys are depicted in the left and right columns.

(A) Reaction time (RT) distributions (histogram bin width = 10 ms).

(B) All shapes, except the last, influenced the monkeys' decisions. Graphs show the relative influence of shapes that appeared within ±10 ms of the time indicated on the abscissa. The leverage is estimated from logistic regression (β_1 in Equation 2), which incorporates the other shapes shown on a trial as potential confounders. On the left portion of the graph, time is aligned to the onset of the first shape; shapes appear every 250 ms and onsets are shown by the short thick lines (shading indicates SE). On the right portion of the graph, time is aligned to saccade initiation. The leverage is calculated every 10 ms (shading indicates SE), using all trials in which a shape appeared within ±10 ms of the time indicated.

(C) Distribution of the number of shapes that affected the decision in each trial. This is a discrete version of the RT distribution, adjusting for late shapes that did not influence the decision. We refer to this number of shapes as N^* (see Experimental Procedures).

(D) Choice probabilities were governed by the cumulative evidence supplied by the shapes. The cumulative evidence is the sum of the logLR assigned to the first through N^{*th} shapes shown on each trial.

(E) Cumulative evidence at the end of decisions. Points are the means of the cumulative logLR for trials that ended with a total of N^* shapes leading to choice A (circles) or B (squares). Error bars represent SD to highlight the diversity of the values (SEM are about half the size of the symbols). Dashed lines are fitted lines (weighted least-squares).

(F) Comparison of the assigned weights to the apparent subjective weights assigned by the monkeys to each unique shape. Assigned weights are in units of logLR. The weights are the coefficients from logistic regression ($\beta_{1\text{-}8}$, Equation 3). Error bars represent SE. Most error bars are smaller than the data symbols.

See also Figure S1.

appearance of a shape, we applied logistic regression at finer time steps. At each time point, we extracted the leverage of shapes that were presented within ±10 ms of the time point displayed on the graph. Importantly, shapes that were presented long before the saccade had leverage on the choice (up to ~1.5 s and ~0.75 s for monkeys E and J, respectively; $p < 0.05$), indicating that early shapes affected the choice even when they were followed by multiple shapes. The analysis also demonstrates that shapes that are shown within the last 200–300 ms had negligible effects on the choice (last 270 ms and 180 ms for monkeys E and J, respectively). This observation implies that, on most of trials, the monkeys had committed to a decision before the evidence from the last shape was incorporated into the deliberation process. For monkey E, this non-decision interval was long enough to allow for the possibility that on some trials even the penultimate shape appeared after commitment (i.e., if a saccade occurred <20 ms after onset of the final shape).

We applied these adjustments to generate the "number of samples" histograms in Figure 2C. These discrete RT distributions depict more clearly the number of shapes the monkeys used to reach their decision, which we term N^*. Although monkey E used more samples than monkey J (mean 5.7 ± 1.4 versus 2.4 ± 1.5 shapes, respectively), both monkeys exhibited considerable variance in the number of shapes they relied upon. A possible explanation will be clear in a moment. Importantly, when either monkey's decision was based on more than one shape, the entire succession of shapes, from the first to the N^{*th}, was influential. From here on, we perform all analyses on these shapes.

By basing their decisions on more than one shape, the monkeys improved their performance. Had they based their decisions on just one shape, even perfect knowledge of the assigned logLR would achieve only 63% correct (i.e., rewarded) choices, owing to the overlap of the sampling distributions (Figure 1B). However, monkeys E and J achieved 85% and 80% correct choices, respectively. Perhaps a more telling observation is that the fraction of trials in which the monkey chose the option

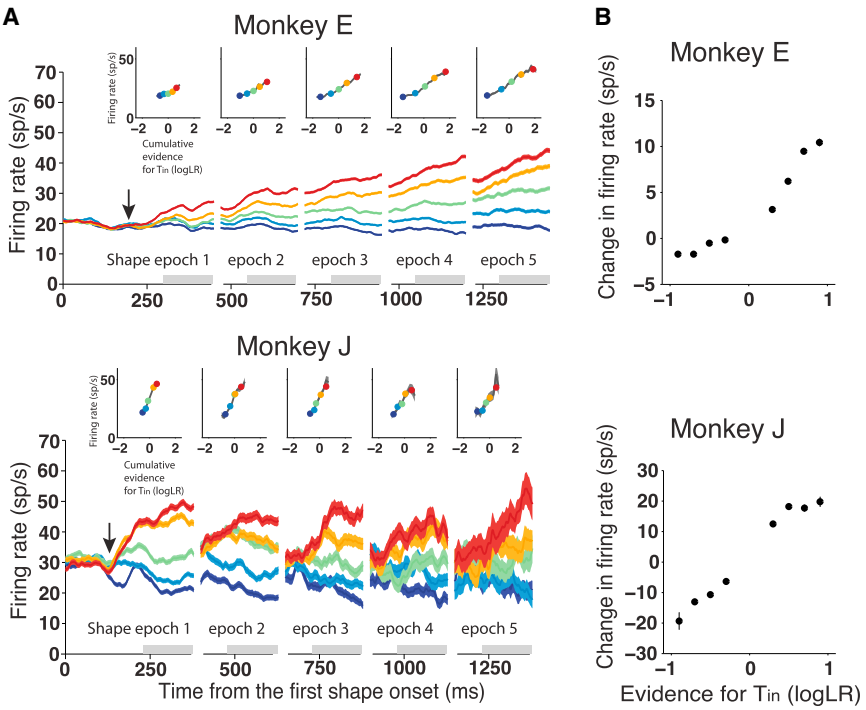

Figure 3. LIP Neural Activity Accompanying Decision Formation

(A) Average firing rates reflect the cumulative evidence from the sequence of shapes. Average firing rates (lower row) are plotted as a function of time from the onset of the first shape. The five curves in each panel correspond to five quantiles grouped by the cumulative logLR in favor of T_{in}. Quantiles are redefined in each epoch, denoted by the break between panels, which respects the sensory delay (τ_s) evident in the leftmost panel (arrow). Notice also the ~100 ms rise time of the response after the arrow. Firing rates in each epoch were calculated in a 150 ms window (gray bars on the abscissa; see Experimental Procedures). Curve thickness depicts SEM. The last effective shape in each trial was excluded from the analysis. Insets show the mean firing rate at the end of the corresponding epoch, plotted for all unique values of cumulative logLR in each epoch (thickness depicts SEM). Colored symbols correspond to the quantiles in the graphs below the inset. All panels include both correct and error trials.

(B) The average change in firing rate induced by each of the eight shapes. Most error bars (SEM) are smaller than the symbols.

For monkey J, only trials with the red target in the RF are shown. See also Figures S2 and S3 and Movies S1 and S2.

supported by the cumulative logLR from the shapes—regardless of what was actually rewarded—was 90% and 97% of trials (E and J, respectively; Figure 2D).

More striking yet, the cumulative logLR was remarkably similar, on average, whether the monkeys used a few or many shapes to make a decision (Figure 2E). The consistency of these averages, as a function of the number of shapes, suggests that the decisions might terminate when the representation of cumulative logLR reaches a threshold level. Accordingly, the monkeys tended to use fewer shapes on trials in which the first few shapes conferred stronger evidence (Figure S1A). We infer that the representation is inexact, however, because the SDs (error bars in Figure 2E) are substantial, a point we will elaborate later. Also, this putative threshold level may not be time invariant (i.e., flat), as evidenced by the dashed lines (Figure 2E; $p < 10^{-4}$, weighted regression). However, the similarity of logLR at decision termination rules out the main alternative to a threshold on accumulated evidence. Were the reaction times controlled by some other criterion, independent of the level of evidence, the average logLR would be the product of the number of shapes times the expectation of the logLR, but this is clearly refuted by the data (Figure S1B).

Finally, we assessed how much weight the monkeys assigned to each of the eight shapes (Figure 2F). For this analysis, we again used logistic regression, but instead of using the sum of the assigned logLR, we estimated the relative weighting of each of the unique shapes (fitted coefficients β_{1-8}, Equation 3). For both monkeys, the estimated weights differ from the assigned weights, but they have the correct sign, and they are ordered appropriately with just one exception for each monkey (Pearson r = 0.98 and 0.93 for monkeys E and J, respectively; $p < 0.01$). Although presented location of the shapes had weak effects on the leverage in monkey E (data not shown), such effect would not affect the rest of the analyses because the location dependency is averaged out in any analysis that is based on a large number of samples. From these, we conclude that both monkeys learned the approximate weights of the shapes. We refer to these approximations as subjective weights.

Physiology

We recorded from single neurons in the ventral part of the lateral intraparietal area (LIPv) while the monkeys performed the sequential inference task in Figure 1. We selected all well-isolated neurons that exhibited spatially selective persistent activity during memory-guided saccades (n = 67; 38 in monkey E; 29 in monkey J). During recording, we mapped the response field (RF) of a recorded neuron and placed one choice target inside the RF (T_{in}) and the other choice target outside the RF (T_{out}). The targets were positioned symmetrically about the fixation point. The targets differed in color for monkey J, whereas both were red for monkey E, reflecting the two training strategies (see Discussion and Experimental Procedures).

Shortly after the presentation of the first shape, LIP activity began to reflect the cumulative evidence in favor of the choice associated with the neuron's RF. Figure 3A shows the evolution of the firing rate as a sequence of up to five shapes were presented. The trials are grouped into five quantiles based on the cumulative assigned logLR in each epoch. To avoid contamination of these averages with activity associated with saccade preparation, the firing rate averages exclude activity induced by the N^{*th} shape. Note that any one trial will tend to contribute to different quantiles in each epoch, and while almost all trials

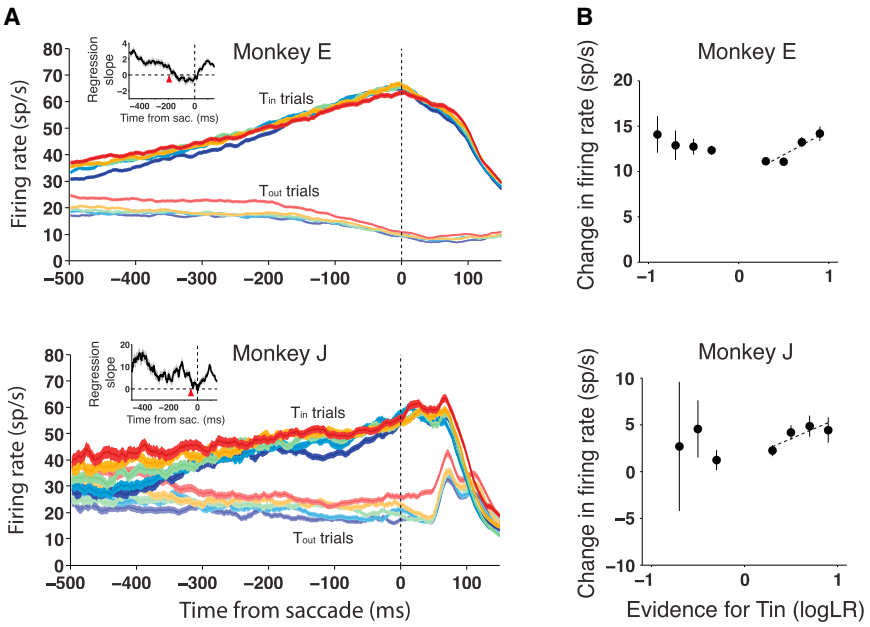

Figure 4. LIP Neural Activity Accompanying Decision Termination

(A) Average firing rates were obtained from data aligned to saccade initiation, grouped by choice (T_{in} or T_{out}). Within this grouping, the trials were sorted by quantile of the assigned cumulative logLR at the point in time τ_s+100 ms prior to each plotted time point (i.e., 100 ms after the responses begin to diverge in the leftmost panel of Figure 3A). Curve thickness depicts SEM. Insets show the diminishing effect of cumulative logLR on firing rate preceding saccade initiation to T_{in}. Points are regression slopes (firing rate as function of cumulative logLR); shading indicates SE. The red arrow marks the time when the cumulative logLR ceases to affect the neural response.

(B) Changes in LIP activity induced by the last effective shape (N^{*th}) when the monkeys chose T_{in}. The abscissa shows the logLR assigned to each of eight unique shapes, as in Figure 3B. Dashed lines are fits using the four shapes that support T_{in}. Monkey J never chose T_{in} when the N^{*th} shape supplied the strongest evidence for T_{out}. Error bars represent SEM.

For monkey J, only trials with the red target in the RF are shown. See also Figure S3 and Movies S1 and S2.

contribute to the responses to the first shape epoch, the number declines in the later epochs.

There are two salient observations evident in the data from both monkeys. First, shortly after appearance of the first shape, the firing rates vary as a function of the logLR assigned to the shapes. The effect is not evident immediately but begins after 200 ms or 130 ms (for monkeys E and J, respectively; Figure 3A, arrows). These long "sensory" latencies (τ_s) are consistent with other studies in which parafoveal information impacts the peripheral RF (Huk and Shadlen, 2005; Roitman and Shadlen, 2002), and they might reflect the different strategies employed by the two monkeys to solve the task (see below). Second, with each successive shape, the range of the cumulative logLR expands, and this is reflected in the separation of the firing rates, grouped by quantile of logLR. In addition, there are some features unique to each monkey. For monkey E, even when the evidence is neutral (middle quantile; green trace), there was a gradual increase in the firing rate as a function of time. Such an evidence-independent rise is thought to reflect a temporal cost, or "urgency," to complete the decision (Churchland et al., 2008; Drugowitsch et al., 2012; Hanks et al., 2014; Thura et al., 2012) as explained further below. For monkey J, the firing rates reflected cumulative logLR only on trials when the red target was displayed in the RF (Figure S3). For this reason, we will show physiological results separately for trials with the red target (Figures 3 and 4) and the green target (Figure S3) in the RF.

The quantile grouping helps to simplify the graphs, but it conceals a more refined relationship between the gradation of evidence and LIP firing rate. The insets in Figure 3A show that the firing rate varies approximately linearly with the cumulative logLR throughout its range. Each subplot shows the mean of the firing rate as a function of all unique values of cumulative logLR in each epoch. Notice that the correspondence between firing rate and cumulative logLR is roughly consistent for the first five to six shape epochs, which accounts for ~90% of the trials. We do not expect these relationships to be identical (owing to conditionalization based on RT; see Figure S4B), but the similarity provides support for the following conclusion: although the shapes appear for only 250 ms and then disappear, they continue to impact the firing rate by contributing to the cumulative evidence in later epochs. This claim is supported more formally by a regression analysis (Equation 4). For example, shapes presented in the first epoch have significant leverage on the firing rate through the later epochs (up to sixth and fifth epoch for monkeys E and J, respectively; $p < 0.05$).

During decision formation, each new shape adds a quantity of evidence and should therefore increment the LIP activity accordingly. We therefore estimated the change in firing rate (ΔFR) induced by each of the eight shapes (Equation 5). Figure 3B compares the eight ΔFR values, averaged across epochs, to the assigned logLR values (see also Figure S2). For both monkeys, shapes that were assigned positive logLR induced increments in the firing rate ($\Delta FR > 0$), and shapes that were assigned negative logLR induced decrements. For monkey E, the magnitudes of the decrements were small, but this is explained by a positive offset to all of the ΔFR values, consistent with an evidence-independent increase in firing rate as a function of time mentioned above (e.g., neutral quantile in Figure 3A).

The reaction-time task allows us to examine the neural events associated with termination of deliberation and commitment to a choice. In contrast to the divergence of responses associated with evidence accumulation, the activity near the end of the decision process exhibited convergence. Figure 4A shows the average firing rates aligned to the onset of eye movements to T_{in} or T_{out}. The responses are broken down further by the relative support for that choice. Note that relative to the time of saccade,

new shapes do not appear and disappear at regular intervals (unlike Figure 3). Nonetheless, we could infer the cumulative logLR that would influence the LIP firing rate after the sensory delay (τ_s) plus 100 ms rise time (evident in Figure 3A). Thus, we sorted the LIP firing rates at each time point (t) into five quantiles based on the cumulative logLR τ_s + 100 ms prior to each time point (i.e., $t - \tau_s$ – 100 ms) (colors, Figure 4A). The resulting graph allows us to visualize when, in relation to the saccade, variation in the cumulative evidence influenced the neural response.

For T_{in} choices, the diversity of neural activity associated with strength of evidence gradually dissipated near the time of the saccade. We interpret this as a sign of a threshold or bound that would terminate decisions in favor of T_{in}. Notably, responses converged to the same level of threshold for both correct and error T_{in} choices ($p > 0.2$), suggesting that error choices occurred because the LIP firing rate misrepresented the cumulative logLR. For T_{out} choices, the neural activity did not converge (monkey J) or the convergence was less pronounced (monkey E), suggesting that activity of other neurons, presumably those with the RFs aligned to T_{out}, precipitated termination of these decisions. The effect of cumulative evidence on firing rate can be captured at each time point by regressing the firing rate (associated with one choice) against the cumulative logLR (at τ_s + 100 ms prior to each time point; see above). For trials ending in T_{in} choices, this modulation is no longer detectable before saccade initiation (red arrows, Figure 4A insets; ~190 ms and ~50 ms before saccades for monkeys E and J, respectively, $p > 0.05$), consistent with a bound or threshold at ~50 sp/s. We interpret the time from the red arrows to saccade initiation as an average motor latency from decision termination to saccade initiation. While these latencies are supported by other analyses (e.g., decline in firing rate variance), we feel they are at best rough approximations (e.g., for monkey J, there is a competing nadir in the regression slope to the left of the red arrow). Nonetheless, the loss of evidence dependence is consistent with the idea that the decision process ends when the firing rate, averaged across a population of neurons, reaches a threshold level.

According to this hypothesis, the final shape that affects a decision for T_{in} should be associated with an increase in the firing rate. This prediction is supported by the analysis in Figure 4B, which shows ΔFR values for each shape on the trials when the shape was inferred to be the last effective shape leading to a T_{in} choice (i.e., the N^{*th} shape of the sequence). Not surprisingly, the last shape was more often one of the four assigned positive weights, and these continued to exert differential effects on the ΔFR (fitted lines, Figure 4B; weighted regression; $p < 0.05$). Shapes with negative weights are expected to terminate some trials, because they happen to give rise to an increase in firing rate, owing to noise, or because a decrement is compensated by the time-dependent urgency signal (monkey E). Interestingly, when a negatively weighted shape terminated the trial, the ΔFR also tended toward positive, on average (12.6 ± 0.4 and 1.7 ± 1.0 sp/s for monkeys E and J, respectively; $p < 10^{-100}$ and 0.09; t test). Monkey J provides a less compelling case possibly because this monkey terminated trials with a negatively weighted shape less often (492 compared to 2,837 trials for monkey E). In principle, the ΔFR associated with termination should never be negative, but this applies to the neural population, not to single trials from one neuron. These observations lend further support to the hypothesis that decisions terminate when the LIP firing rate reaches a threshold, which is detected, presumably, by neurons downstream (Heitz and Schall, 2012; Lo and Wang, 2006).

Finally, it is worth noting that the graded representation of evidence in Figure 4A, even when trials are grouped by choice, implies that the graded responses evident in Figure 3A do not arise as a consequence of averaging responses comprising different mixtures of T_{in} and T_{out} responses. Thus, taken together, Figures 3A and 4A imply that LIP neurons represent an evolving decision variable that will ultimately lead to a T_{in} or T_{out} choice.

Model

The observations from both behavioral and neural recordings suggest that decisions are made when a sufficient amount of evidence has been accumulated in favor of one of the alternatives. For example, Figure 2E indicates that, on average, deliberation ends when the absolute value of the cumulative logLR reaches a common level, regardless of the number of shapes it takes to achieve this level. Furthermore, the stereotyped firing rate before T_{in} choices also supports a stopping rule based on the neural representation of this cumulative evidence. However, there are also anomalous observations that cannot be overlooked. Across trials, there is considerable variation in the state of the cumulative evidence at termination. This is evident in the SDs depicted in the same Figure 2E. Indeed, the choice function displayed in Figure 2D spans a range of cumulative evidence levels, which at face value contradicts the premise that all decisions stop at or very near a positive or negative threshold. We hypothesized that these contradictory observations can be explained by the noisy representation of logLR by LIP neurons and a termination mechanism that operates on the LIP firing rate rather than directly on logLR. The following neurally constrained model supports the plausibility of this hypothesis. See Supplemental Experimental Procedures for mathematical details and supporting information.

We modeled the decision process as a race between two accumulators: one accumulating evidence in favor of Target A and the other accumulating evidence in favor of Target B. The race architecture allows us to apply parameters derived from neural recordings (Churchland et al., 2008; Mazurek et al., 2003; Usher and McClelland, 2001; Wang, 2002), as follows. Each shape produces an expected increase or decrease in the firing rate, estimated from the detrended average firing rates of the population of recorded neurons (Figure 3B and Equation S1). These expectations, $\langle \Delta r_i \rangle$, are assumed to be of opposite sign for the two accumulators (Figure 5B). We assume that each shape induces an actual change in firing rate, Δr_i, that is determined by this expectation plus unbiased noise (SD σ_ε; Figure 5B, orange distribution), which is assumed to be independent in the competing process. Sampled values of Δr_i scale a temporal impulse function, $x(t)$ (Equation S2), which also captures the delay and dynamics of the response to single shapes (e.g., Figure 3A). The noisy impulses are integrated and added to a time-dependent urgency signal, $u(t)$, which is derived from the neural recordings (Equation S4, Figure S2C). The process starts from an initial rate of v_0 (20 and 30 sp/s for monkeys E and J, Figure 3A) and

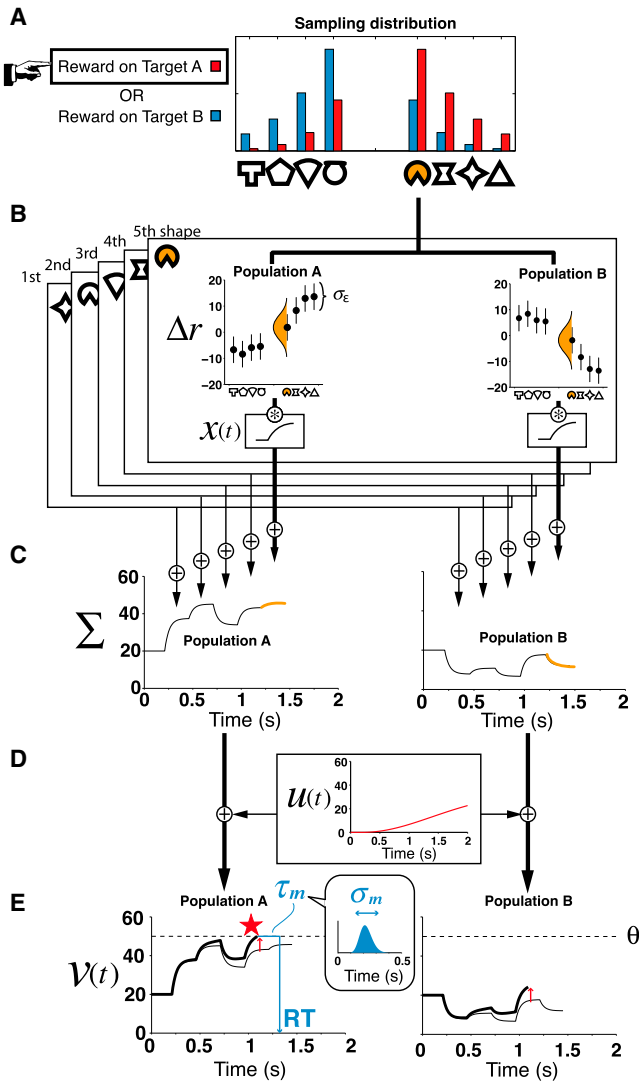

Figure 5. A Bounded Accumulation Model Informed by the Neural Recordings

Two competing neural populations with response fields at choice target A or B, respectively, accumulate a noisy representation of evidence bearing on the likelihood that choosing A or B will be rewarded. The panels establish the logic of the model simulation in stages via an example trial in which a correct decision for A is made with RT = 1.3 s, based on the evidence from $N^* = 4$ shapes. The fifth shape does not affect the decision in this example.

(A) At the beginning of each trial, a reward is assigned randomly to Target A or to B. The assignment determines the sampling distribution for the shapes in the trial. For example, if the reward were assigned to Target A, the shapes would be sampled from the distribution in red, perhaps giving rise to the sequence of shapes: star, pacman, wedge, hourglass, and pacman.

(B) The presentation of each shape gives rise to a change in firing rate in each of the racing accumulators, termed population A and B. The expected Δr for population A is taken from the detrended data (Equation S1); the expectations are the opposite sign for population B. These expectations are corrupted by unbiased noise, represented by the Gaussian distributions with standard deviation σ_ε (error bars). For example, the pacman presented as the fifth shape (orange shading) gives rise to a change in firing rate in population A, modeled as a random draw from the orange Gaussian distribution on the left. Population B undergoes a change modeled as an independent draw from the orange Gaussian on the right. The corrupted samples scale a dynamic template, $x(t)$ (bottom insets), which captures the sensory delay and rise time of the LIP response.

(C) The evidence-dependent activity, Σ, is the running sum of these scaled templates. The portion of the response that would be contributed by the fifth shape is shown as an orange curve.

(D) The time-dependent urgency signal, $u(t)$, is added to Σ in each neural population. This undoes the detrending step in (B).

(E) The urgency signal pushes both accumulations Σ (thin lines) upward to establish $v(t)$ (thick lines). When $v(t)$ in either population reaches the threshold (θ) (marked by a red star), the model chooses the corresponding target (Target A in this example). Notice that it is based only on the first four shapes. The RT occurs after a motor delay (τ_m), a random variable drawn from a Gamma distribution (blue).

See also Figure S2 and Supplemental Experimental Procedures.

terminates when one of the accumulators reaches a threshold firing rate level, θ (50 sp/s for both monkeys), thereby determining the choice and the time of decision termination on a trial. The measured reaction time is the decision time plus a gamma distributed delay, τ_m, with mean $\langle \tau_m \rangle$ adopted from the analysis in Figure 4A, and SD σ_m. We fit the model to approximate the choice and RT distributions for each monkey (Figure 6A; Equation S5) using just two free parameters: the noise terms associated with the evidence-induced increments in firing rate and the motor latency (σ_ε and σ_m; Table 1).

As shown in Figure 6, the model reproduces the key behavioral measurements. Importantly, the model resolves the anomalies mentioned above. It gives rise to the consistency of mean cumulative logLR irrespective of RT, suggesting a terminating bound, while at the same time explaining the variation in this same quantity across trials (Figures 6B and 6C). It does so by postulating that decisions terminate when a population firing rate reaches a threshold, and these firing rates are noisy reflections of the underlying evidence.

The model also explains a trend, evident in monkey J, that the magnitude of cumulative logLR increases with the number of accumulated shapes at the end of decision (Figure 6C). The intuition is as follows. Imagine a pair of trials that share the same exact sequence of shapes, and suppose further that in the absence of noise this sequence ought to lead to choice A after three shapes. If the shapes are represented by noisy quantities, then it is possible that one trial leads to choice A after three shapes, whereas a fourth shape is required on the other trial. The cumulative logLR will tend to be greater on this trial because it is highly likely that the fourth shape was assigned positive logLR. Such an effect would be mitigated by the urgency signal measured in monkey E.

Similar reasoning guides interpretation of another potentially misleading observation: shapes presented nearer to decision termination have greater leverage on choice than shapes presented earlier in the sequence (Figure 6D). Again, the explanation relies on the noisy representation of logLR. For a trial to extend through many shapes, it is likely that those shapes presented in earlier epochs changed the firing rate by a smaller magnitude than those presented in later epochs. As shown by the red curve (Figure 6D), the increased leverage of shapes presented in later epochs is the expected trend for bounded noisy evidence accumulation. It should not be construed as support for forgetting or "leaky accumulation." One could argue that the nonterminating,

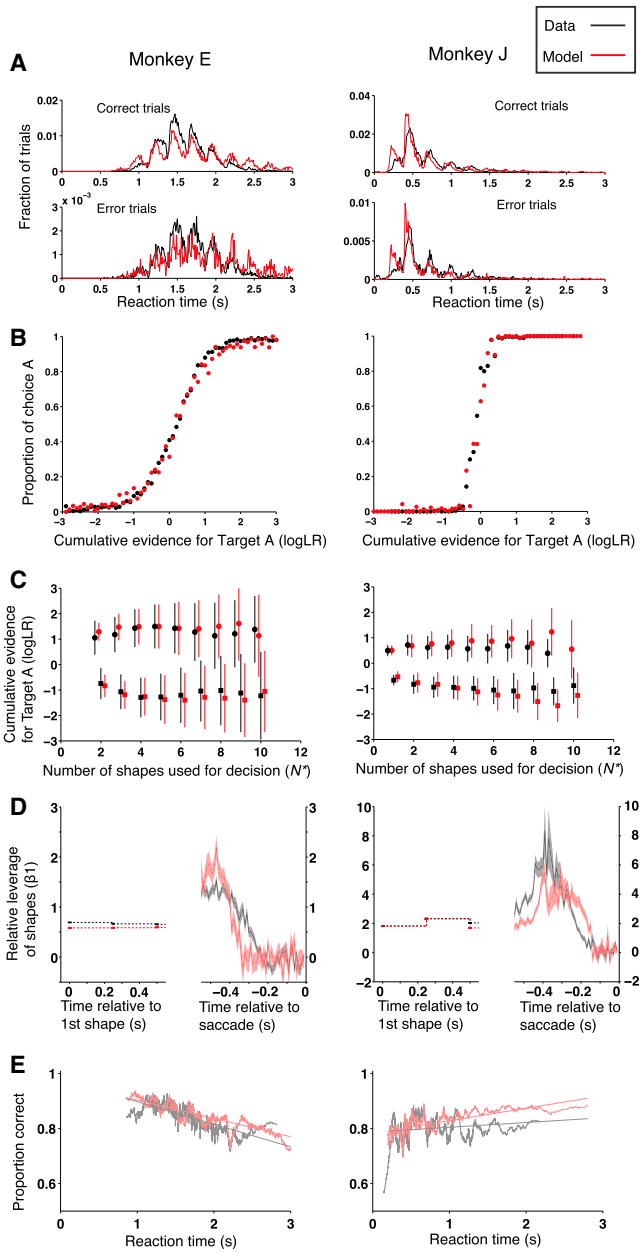

Figure 6. Comparison of the Model to Experimental Data

The data (black) and model predictions (red) for monkeys E and J are shown in the left and right columns of graphs, respectively. All but two parameters of the model are derived from the neural recordings (see Table 1). RT distributions in (A) were used to fit these parameters. Red traces in (B)–(E) are predictions.

(A) Reaction time distributions. Data are the same as in Figure 2A but are split by correct and error trials.

(B) Dependence of choice on the cumulative evidence supplied by the sequence of shapes. Data are the same as in Figure 2D.

(C) Mean and SD of the cumulative logLR at the end of decision. Data are the same as in Figure 2E.

(D) Influence of shapes on choice as a function of time (shading indicates SE). Data are same as in Figure 2B.

(E) Time-dependent accuracy (TDA). Data are the running proportions of correct choices sorted by RT. Values are plotted at the mean RT of the 300 trials constituting the running mean. The decreasing TDA for monkey E is explained by the evidence-independent "urgency" signal that is added to the representation of cumulative evidence. Lines are maximum likelihood fits to the data and model simulation using the individual trials (not running means) ordered by RT.

See also Figure S4.

lower bound in our model leaves open the potential for another source of information loss, but simulations that permit negative firing rates demonstrate that this is not a viable explanation of the pattern of weighting in Figure 6D (data not shown). Thus, we find no evidence for memory leak or sources of noise beyond those accompanying the evidence samples, consistent with a recent study in rat (Brunton et al., 2013).

The model also furnishes insight into the differences between the behavioral patterns exhibited by the two monkeys. Compared to monkey J, monkey E appears to derive smaller changes in firing rate—relative to noise—from the shapes (Table 1). However, this monkey compensates by applying a more conservative bound. That is, the difference in firing rate from beginning of evidence accumulation to the terminating bound is larger in monkey E, consistent with the longer RT, or equivalently, the larger number of sampled shapes, on average. This excursion also incorporates an evidence-independent, time-dependent "urgency" signal, $u(t)$ (Figure 5D), which reduces the amount of evidence required to terminate the decision, and thus explains the decline in accuracy as a function of RT (Figure 6E). Interestingly, these differences led monkey E to a higher overall accuracy rate (85% versus 80% correct) but to similar reward rates (0.183 versus 0.195 s^{-1}).

Finally, although it is no surprise that the neurally inspired model recapitulates the neural responses (Figure S4A), it explains an otherwise puzzling feature of these data. Recall that the relationship between firing rate and cumulative logLR is approximately constant for the first five shape epochs (Figure 3A). However, for trials with longer RT, the relationship tends to flatten (Figure S4B). This occurs without any leakage but is instead consistent with the accumulation of noisy evidence to a threshold. That is, the monkeys tended to use more shapes on those trials in which the evidence affected the firing rate by a smaller magnitude than average. The correspondence between model and physiology reassures us that the model architecture and estimates we obtained were sufficient to capture the main features of the data. Thus, the model supports the hypothesis that an approximation to the accumulation of logLR, as represented by neurons in area LIP, explains the pattern of choice and RT manifested by the two monkeys.

DISCUSSION

When a decision is based on samples of evidence of different reliability, then it is rational to give more weight to the more reliable cues (Ernst and Banks, 2002; Fetsch et al., 2012; Jacobs, 1999; Knill, 2007). If the evidence arrives sequentially and more is expected, then the decision maker might also decide when to terminate the process and commit to a choice. Here, the notion of rationality is less well prescribed, as it depends on the desired level of accuracy and the cost of time. In many instances, the speed-accuracy tradeoff is thought to be explained

Table 1. Model Parameters

	Fitted Parameters		Physiologically Derived Parameters			
	σ_ε	σ_m	τ_s	$< \tau_m >$	v_0	θ
Monkey E	5.9 sp/s	39 ms	200 ms	204 ms	20 sp/s	50 sp/s
Monkey J	8.3 sp/s	69 ms	130 ms	57 ms	30 sp/s	50 sp/s

by a common mechanism in which evidence is accumulated to a threshold level or bound. Previous studies of perceptual decision making have demonstrated a neural correlate of both evidence accumulation and termination in area LIP (Gold and Shadlen, 2007). In these studies, the reliability of evidence was statistically stationary during a trial, and as best we can tell, all samples of evidence during the course of a single decision were weighted identically. On the other hand, while many behavioral studies have demonstrated a rational combination of cues of different degrees of reliability, nearly all such studies present the cues simultaneously. The main exception is the study upon which the present experiment is based (but see Wyart et al., 2012).

Our previous study (Yang and Shadlen, 2007) demonstrated that monkeys were capable of combining evidence from a sequence of four shapes, like those used in the present experiment, by assigning greater weight to the more reliable cues. Although the LIP responses in that study appeared to encode the cumulative evidence from the sequence of shapes, this was the only support for the conjecture that the monkeys actually reasoned sequentially from the four samples. The monkey was required to make its choice only after all four shapes were visible on the display.

The present study enforced a sequential strategy in two ways. First, the task required integration in time because only one shape was visible at a time. Second, the monkey was free to terminate each trial when it had seen enough shapes. The behavioral data indicate that shapes from early epochs influenced the monkeys' choices, even on trials in which such shapes were followed by many more shapes (Figure 2B). Moreover, the speed and accuracy of the monkeys' decisions were well described by bounded evidence accumulation as shown in our modeling exercise. Thus, monkeys demonstrated a capacity to form decisions based on sequential analysis of evanescent cues of varying reliability.

A well-known strategy to solve this class of decisions, known as the sequential probability ratio test (SPRT) (Barnard, 1946; Good, 1979; Wald, 1947), would place termination criteria on the cumulative logLR supplied by the shapes. The level of these criteria determines the error rates and the distribution of the number of shapes supporting decisions, across trials. In fact SPRT is optimal in the following sense. Given a desired error rate, it ensures that the number of samples is minimized on average (Wald and Wolfowitz, 1948). A natural candidate for "desired error rate" is one that would maximize the rate of rewarded decisions over many trials (Gold and Shadlen, 2002; Simen et al., 2009). Neither of the monkeys achieved this optimal level, but they came close—monkeys E and J achieved a reward rate that is 86% and 92%, respectively, of the maximum possible. However, it is probably incorrect to assume that the monkeys aimed to maximize the rate of rewarded decisions.

As has been previously shown, decision time itself appears to be costly to monkeys. For example, in some perceptual decisions, when stimuli are presented for durations exceeding 0.5 s, subjects tend to ignore late information, as if performing a RT task (Kiani et al., 2008; Tsetsos et al., 2012). This time cost can be realized by collapsing the termination bounds as a function of time, or equivalently, by adding a monotonically increasing signal, termed urgency, to the accumulated evidence represented by competing accumulators (Churchland et al., 2008; Cisek et al., 2009; Ditterich, 2006; Drugowitsch et al., 2012; Hanks et al., 2014). In LIP activity, we observed the latter: the additive urgency. A behavioral signature of such a process is the declining accuracy as a function of RT for monkey E (Figure 6E). The neural correlate of this is the gradual rise in firing rate that is apparent even when the accumulated evidence favored neither choice (Figure 3A, green traces, monkey E).

Other explanations for the less than optimal performance are more informative about the underlying process. The most obvious is that neither monkey assigned the eight shapes the correct weights, although these anomalies were concentrated on the shapes that were shown least often. This led to systematic error in the choices, although the effect was modest. For example, if we assume an accumulation of weights like those shown in Figure 2F, without noise, and apply the optimal bound, the overall rate of reward would be reduced by only ~1%. This reduction is due to imperfect assignment of relative weights (Figure 2F) rather than the absolute values of weights. Therefore, the more important discrepancy, highlighted by the model, is that the neural representation of accumulated evidence is corrupted by noise. This introduces a diminution in the signal-to-noise ratio that can only be overcome by increasing integration times (e.g., a higher bound setting) at the cost of more time per trial. One might argue that the brain is incapable of achieving an optimal solution to this problem, but given limitations in training and proximity to optimal performance, we tend to take the rosier view that despite limitations of biology (e.g., noisy neurons) and learning, the brain can achieve a reasonable approximation to optimal cue combination.

We have mainly focused on aspects of the decision process that were shared by the two monkeys, but the model also sheds light on differences between the monkeys, which may have been induced by their training histories. We introduced a minimum delay to reward during training of monkey E in order to discourage very fast RT, which is the natural tendency of many monkeys on RT tasks (e.g., see Roitman and Shadlen, 2002). We suspect that this might have induced some procrastination between decision termination and saccade initiation (i.e., τ_m) in this monkey. In addition, monkey E could have increased its reward rate, albeit slightly (~1%), by accumulating evidence without bounds during the minimum delay. An approximation to this strategy might account for the time-dependent urgency signal evident only for monkey E, because the addition of $u(t)$ to both accumulators is equivalent to decreasing the bound height as a function of time. We do not know if this training difference also led monkey E to give less weight to the shapes or if instead the lower signal to noise led to a compensatory change in the termination criterion. What we can say, however, is that the

combination of bound and weight conferred monkey E with the better accuracy than monkey J.

Another difference between the monkeys can be attributed more clearly to a difference in their training history. Monkey J had been trained extensively on a version of the task in which the shapes were associated with colored targets. In an earlier report, there was a hint that its neural responses were more reliable when the red choice target was displayed in the RF (see Figure S8 in Yang and Shadlen, 2007). This trend evolved during training on the RT task such that the monkey now appears to use only one population of LIP neurons on this task. That is, when the green target was in the RF, although the monkey performed the task equally well, the neurons did not modulate their activity during the evidence accumulation phase of the task (Figure S3A). This calls for a qualification of the model in Figure 5. For this monkey, we have no neural correlate of a terminating threshold for green target choices. One possibility is that such a process exists in another decision-related brain area that is less associated with eye movement planning per se (e.g., dorsolateral prefrontal cortex; Brody et al., 2003; Kim and Shadlen, 1999; Romo et al., 1999; Wallis et al., 2001). Another possibility is that monkey J performed the task by deciding for or against the red choice target. In that case, the termination rule for green might be a lower bound on the firing rate of LIP neurons that support red. Indeed, this lower bound might be sensed by the neurons with the green target in their RF, since they exhibit presaccadic responses before T_{in} choices (Figure S3B). However, the neural activity depicted in Figure 4A—that is, the neurons with the red target in RF—does not provide compelling evidence for a lower bound on T_{out} trials (i.e., green choices): the firing rates continue to reflect the accumulated evidence up to saccade initiation.

We considered and rejected a variety of alternative models to explain the monkeys' behavior. First, the monkeys did not base their choices or decision times on the occurrence of particular shapes. With this strategy, the subjective weight should be zero for all shapes except for those particular ones, whereas each unique shape was assigned a non-zero weight with an appropriate sign and rank order (Figure 2F). Second, the monkeys did not rely on evidence conferred in a particular epoch during a trial because shapes impacted the choice whether presented early or late (Figure 2B). Third, decision times were not dictated by some type of deadline that is independent of the state of the cumulative evidence. This idea would not explain the consistent level of evidence at termination (Figure 2E; see Figure S1B for further explanation).

That said, the model we propose is undoubtedly incomplete. We have already mentioned the absence of a signature of decision termination on green target choices (monkey J). More importantly, we do not know how the monkeys learn to associate differential weights to the 8 shapes, and it is not clear how the appearance of a shape leads to an increment or decrement in the LIP firing rate that is of the appropriate magnitude for that shape. Comparison to the well-studied random dot motion task is instructive. In that task, neurons in extrastriate area MT (V5) represent the momentary sensory evidence, and the accumulation is thought to be of the difference in firing rates of such neurons tuned to opposite directions. The noisy difference signal is itself proportional to units of logLR (Gold and Shadlen, 2001). We assume that the shapes are differentiated in extrastriate cortical areas in the ventral stream (DiCarlo et al., 2012; Tanaka, 1996), but it is hard to conceive of an analogous difference between pro- and anti-pentagon neurons. Instead, it has been suggested that LIP neurons acquire evidence-dependent activity by learning appropriate synaptic weights applied to shape-selective inputs (Soltani and Wang, 2010; Rombouts et al., 2012).

What seems to be required is an operation resembling memory retrieval in which the shape is associated with a quantity that bears on another process. In memory retrieval, this quantity bears on a decision about satisfaction of a match (e.g., similitude) (Ratcliff, 1978; Ratcliff and McKoon, 2008; Shadlen and Kiani, 2013), whereas in our experiment it bears on the relative merit of a left/right or red/green choice. Perhaps this insight bears on the success of sequential sampling models to a variety of cognitive tasks, including memory retrieval, that do not appear to involve any obvious need to integrate independent samples of evidence as a function of time.

EXPERIMENTAL PROCEDURES

Two male rhesus monkeys (*Macaca mulatta*) were implanted with a head fixation device and a recording chamber above the intraparietal sulcus. Throughout training and recording sessions, eye positions were recorded by an infrared video-tracking system (EyeLink, SR Research; sampling rate, 1,000 Hz). Timing of task-related events was controlled by a real-time computer data acquisition system (Hays et al., 1982). Visual stimuli were displayed on a CRT monitor (frame rate 75 Hz) controlled by MATLAB (MathWorks) and the Psychophysics Toolbox (Brainard, 1997). All animal procedures complied with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, and were approved by the University of Washington Animal Care Committee.

Task

Monkeys were trained to perform a choice-reaction time task in which they evaluated a sequence of shapes to choose the better of two saccade targets. The monkeys were trained extensively to interpret eight unique shapes as evidence bearing the likelihood that one or the other choice target would furnish a reward (Figure 1B). Monkey E learned that half of the shapes favored the left or right target, whereas monkey J learned that half of the shapes favored the red or green target. For ease of presentation we refer to these targets as Targets A and B, and in both situations one or the other target was placed in the RF of the LIP neuron under study.

After acquiring fixation on a central fixation point (FP), two choice targets appeared in the periphery equidistant from the FP. After a random delay (0.2–0.5 s, mean 0.3 s for monkey E; 0.5–3 s, mean 0.9 s, for monkey J), a sequence of highly visible shapes appeared every 250 ms centered on a vertex of an invisible 3° × 3° grid centered on the FP. The shapes were approximately 1.2° × 1.2°, high contrast line art with equal perimeter (Figure 1B). The shapes were displayed sequentially, until the monkey initiated a saccade to one of the choice targets. Successive shapes were not shown in the same grid location. The computer randomized which target would be rewarded on each trial, and sampled the eight shapes accordingly (Figure 1B). Thus, on any one trial, any of the eight shapes could be displayed, but four were more likely, depending on whether the reward was assigned to Target A or B.

After the first shape's onset, the monkey was allowed to make a saccadic eye movement to one of the targets to indicate its choice. The time from the first shape's onset to the time of saccade initiation defines the reaction time (RT) in each trial. Monkeys received a liquid reward for all correct choices so long as the gaze remained on the choice target for 0.2 s (saccade validation). Monkey J received this reward 1.3 s after saccade validation. For monkey E, the interval to reward depended inversely on RT—a minimum of 1.8 s from onset of the first shape (e.g., immediately after validation for all RT > 1.6 s). Similar techniques have been used previously to counter monkeys' natural

tendency to respond quickly on choice-reaction time tasks (Hanks et al., 2014; Roitman and Shadlen, 2002). Upon saccade initiation, no more shapes were shown, and the screen was blanked after the saccade hold period. The next trial followed an inter-trial interval (T_{ITI}) defined as the time from saccade initiation to the display of the FP for the next trial. For monkey J, T_{ITI} was 2.1 s for correct trials and 3.9 s for error trials. For monkey E, T_{ITI} for correct trials was adjusted based on RT: T_{ITI} was 2.1 + (1.8 – RT) s if RT was shorter than 1.8 s, and 2.1 s otherwise. T_{ITI} for error trials was 5.4 s.

Targets A and B refer to the left and right choice target for monkey E and the red and green choice target for monkey J. For monkey J, either the red or green target was randomly designated as T_{in} in each trial. We introduced the spatial version of the task for monkey E after preliminary analyses indicated that half of the trials from monkey J were uninformative (see Figure S3). For monkey J, we included two extra "trump" shapes that the animal had been trained on previously. These shapes, though rarely shown (~2% of trials), would guarantee reward at one or the other choice target. However, the monkey did not always terminate decisions upon seeing these shapes, consistent with our previous inference that the monkey did not learn the significance of these shapes (Figure 1C and Figure S2B in Yang and Shadlen, 2007). Trials incorporating these "trump" shapes are not included in this report. We mention it here because the differences in training might bear on the interpretation of the differences in results from the two monkeys.

Both monkeys were initially trained extensively on a four-shape version of the task (>200,000 trials) (see Yang and Shadlen, 2007). For the RT version of the task, monkey E was trained for ~20,000 trials (~2 months) and monkey J was trained for ~111,000 trials (~6 months). We commenced neural recording when the decision speed and accuracy achieved stability. Monkeys E and J performed ~44,000 trials (~5 months) and ~93,000 trials (~4 months), respectively, during the period of recording (including practice days when we did not record neural data).

Data Acquisition and Neuron Selection

Quartz-platinum/tungsten electrodes (Thomas Rec./Alpha-Omega with 1–3 MΩ impedance) were used for recording. Recorded signals from the electrode were amplified and bandpass filtered (150 Hz–8 kHz for monkey E; 0.1 Hz–10 kHz for monkey J) before action potentials (spikes) were detected by a dual voltage-time window discriminator (Plexon and Bak Electronics).

We recorded from neurons in the ventral division of the lateral intraparietal area (LIPv) (Lewis and Van Essen, 2000) in the right and left hemisphere for monkeys E and J, respectively. Recording sites were targeted using a postoperative MRI displaying the recording chamber and grid. We registered these images with the standard MRI supplied with the CARET software (Van Essen, 2002) and targeted the posterior third of the flat-map representation of LIPv, where, in our experience, clusters of neurons with appropriate RF locations and spatially selective persistent activity are abundant (Patel et al., 2014). We then identified LIPv using physiological criteria and selected isolated single units that showed a robust and spatially selective persistent activity while the monkeys waited to execute a saccadic eye movement toward a remembered target location during the delay period of the memory-guided saccade task (Hikosaka and Wurtz, 1983). This property was common in the region of the intraparietal sulcus ~4–8 mm below the cortical surface. The criterion was qualitative, but it was typically confirmed by interleaving memory-guided saccades during the main experiment. Of 196 units encountered in putative LIPv, 104 cells exhibited robust persistent activity. Of these, 67 cells were held for sufficient trials (>80) for inclusion in the dataset (38 and 29 cells for monkeys E and J, respectively).

Analysis of Behavioral Data

We analyzed 27,201 and 17,994 trials accompanying the neural recordings from monkeys E and J, respectively. We applied several logistic regression analyses to estimate the leverage of the shapes on the probability that the monkey would choose Target A [*P(A)*], as a function of a sum of leverages, *Q*, which might affect *P(A)* monotonically:

$$P(A) = \frac{1}{1 + 10^{-Q}}. \qquad \text{(Equation 1)}$$

To determine whether shapes that appeared in a specific time window affected the monkey's choices (Figure 2B), we considered all the other shapes shown in the trial as confounders:

$$Q = \beta_1 w_t + \beta_2 W \qquad \text{(Equation 2)}$$

where w_t is the assigned logLR of shapes that appeared during the time window of interest, W is the cumulative logLR of the rest of the shapes that appeared during the trial. β_1 and β_2 are the fitted coefficients, where β_1 quantifies the subjective leverage of shapes in the specific time window after accounting for the influence of other shapes. To estimate the leverage of shapes presented at the beginning of the trial, we obtained w_t from logLR of the k^{th} shape (k = 1, 2, …) (Figure 2B, left). To estimate the leverage of shapes presented at time points, t, relative to the saccade, we obtained w_t only from trials in which a shape appeared within ±10 ms of t (Figure 2B, right). For both estimates, we pooled trials across all RT. Standard errors (shading in Figure 2B) are uncorrected for multiple observations and data overlap.

We also performed a variant of this analysis that incorporated a constant (bias term) at each time point. It is difficult to justify an independent bias at each time point, but it is reassuring that the results are consistent with the one we report based on Equation 2 (data not shown).

To estimate the subjective weight that the monkey assigned to each of the eight shapes (Figure 2F), we fit eight coefficients:

$$Q = \sum_{i=1}^{8} \beta_i n_i \qquad \text{(Equation 3)}$$

where n_i is the count of each unique shape during a trial up to the N^{*th} shape. The fitted logistic coefficients (β_i) quantify the leverage of each unique shape in units of logLR.

For the time-dependent accuracy functions (Figure 6E), we constructed running "proportion correct" from trials (or simulated trials) sorted by RT. The lines fit to the data and model predictions used individual trials (not running proportions). These lines and all logistic models were fit to the data as generalized linear models (GLMs) assuming binomial error.

Analysis of Physiological Data

Owing to differences in the training and behavior, all analyses of neural responses were conducted separately for the two monkeys. For monkey J, we include only trials in which the red choice target was in the neuron's RF (8,928 trials). As shown in Figure S3, when T_{in} was green, the neurons did not modulate their firing rate before the peri-saccadic epoch.

We calculated the firing rate of LIP neurons by convolving spike trains with a causal kernel: $\alpha(t) = t/\tau_K^2 \cdot \exp(-t/\tau_K),\ \forall t \geq 0$, where τ_k is 10 ms. We formed population averages from these traces. To establish the epoch in which a shape influences the LIP response, we examined the time course of the response to the first shape. We first identified the time, τ_s, when the responses begin to diverge as a function of evidence. We estimated this delay by dividing all trials into four groups based on the first shape's logLR ([–0.9, –0.7], [–0.5, –0.3], [0.3, 0.5], [0.7, 0.9]), and measuring when the response curves of the four groups started to diverge after the first shape onset. Since shapes were presented every 250 ms, the epoch associated with the k^{th} shape in a sequence extends from $\tau_s + 250(k-1)$ to $\tau_s + 250k$ ms. As is evident from Figure 3A, it takes ~100 ms for the response to approximate a steady level. Therefore, we used the last 150 ms of this epoch to measure the average firing rate (i.e., $[\tau_s + 100,\ \tau_s + 250]$ ms after presentation of the k^{th} shape; Figure 3A gray bars on the abscissa).

We quantified the leverage of each shape in a sequence (j = 1 .. k) on the firing rate in the k^{th} epoch using linear regression:

$$FR_k = \beta_0 + \sum_{j=1}^{k} \beta_j w_j \qquad \text{(Equation 4)}$$

where w_j is the assigned logLR of the j^{th} shape in the sequence. To avoid contamination of activity associated with saccade preparation, FR_k excludes activity induced by the N^{*th} shape. In the example cited in the Results, the null hypothesis is H_0: $\beta_1 = 0$ (weighted regression, t test).

The change in the firing rate induced by the k^{th} shape was estimated from the firing rates relative to the one preceding it:

$$\Delta FR_{i,k} = FR_{i,k} - FR_{i,k-1} \quad \text{(Equation 5)}$$

where $FR_{i,k}$ is the average firing rate induced by shape i (i = 1 .. 8). For the first shape epoch, the baseline firing rate (FR_0) was estimated from the mean firing rate in the epoch from onset of the first shape to τ_s. We computed the average change in firing rate for each unique shape i by averaging across k epochs, $\Delta FR_i = \langle \Delta FR_{i,k} \rangle_k$, where $\langle \cdots \rangle_k$ denotes the mean across epochs. The ΔFR_i are displayed in Figure 3B. To estimate $\Delta FR_{i,N^*}$ induced by the N^{*th} shape (Figure 4B), we excluded data after saccade initiation and trials in which N* = 1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and two movies and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2015.01.007.

AUTHOR CONTRIBUTIONS

T.Y. and M.N.S. designed the experiment. S.K. and T.Y. recorded the physiological data. S.K. and M.N.S. developed the model. S.K., T.Y., and M.N.S. analyzed the data and wrote the paper.

ACKNOWLEDGMENTS

We thank J. Palmer, E. Fetz, E. Shea-Brown, A. Pasupathy, S. Shushruth, and L. Woloszyn for comments on this manuscript and useful suggestions and A. Boulet, K. Ahl, K. Morrisroe, M. McKinley, and J. Wang for technical assistance. This work was supported by the U.S. National Institutes of Health (EY011378, RR000166, P30EY01730) and the Howard Hughes Medical Institute. S.K. was supported by a predoctoral fellowship from the Nakajima Foundation.

Received: July 31, 2014
Revised: November 8, 2014
Accepted: January 6, 2015
Published: February 5, 2015

REFERENCES

Barnard, G.A. (1946). Sequential tests in industrial statistics. Supple. J. Roy. Stat. Soc. *8*, 1–26.

Brainard, D.H. (1997). The psychophysics toolbox. Spat. Vis. *10*, 433–436.

Brody, C.D., Hernández, A., Zainos, A., and Romo, R. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. Cereb. Cortex *13*, 1196–1207.

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.

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

Cisek, P., Puskas, G.A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. J. Neurosci. *29*, 11560–11571.

DiCarlo, J.J., Zoccolan, D., and Rust, N.C. (2012). How does the brain solve visual object recognition? Neuron *73*, 415–434.

Ding, L., and Gold, J.I. (2012). Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. Cereb. Cortex *22*, 1052–1067.

Ditterich, J. (2006). Evidence for time-variant decision making. Eur. J. Neurosci. *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. J. Neurosci. *32*, 3612–3628.

Ernst, M.O., and Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. Nature *415*, 429–433.

Fetsch, C.R., Pouget, A., DeAngelis, G.C., and Angelaki, D.E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. Nat. Neurosci. *15*, 146–154.

Gold, J.I., and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. Trends Cogn. Sci. *5*, 10–16.

Gold, J.I., and Shadlen, M.N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. Neuron *36*, 299–308.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Good, I.J. (1979). Studies in the history of probability and statistics. XXXVII A.M. Turing's statistical work in World War II. Biometrika *66*, 393–396.

Hanks, T., Kiani, R., and Shadlen, M.N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. eLife *3*, 02260.

Hays, A.V., Richmond, B.J., and Optican, L.M. (1982). A UNIX-based multiple-process system for real-time data acquisition and control. WESCON Conf. Proc. *2*, 1–10.

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

Hikosaka, O., and Wurtz, R.H. (1983). Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. J. Neurophysiol. *49*, 1268–1284.

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. J. Neurosci. *25*, 10420–10436.

Jacobs, R.A. (1999). Optimal integration of texture and motion cues to depth. Vision Res. *39*, 3621–3629.

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. J. Neurosci. *28*, 3017–3029.

Kim, J.N., and Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat. Neurosci. *2*, 176–185.

Knill, D.C. (2007). Robust cue integration: a Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. J. Vis. *7*, 1–24.

Lewis, J.W., and Van Essen, D.C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. J. Comp. Neurol. *428*, 112–137.

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

Mazurek, M.E., Roitman, J.D., Ditterich, J., and Shadlen, M.N. (2003). A role for neural integrators in perceptual decision making. Cereb. Cortex *13*, 1257–1269.

O'Connell, R.G., Dockree, P.M., and Kelly, S.P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. Nat. Neurosci. *15*, 1729–1735.

Patel, G.H., Kaplan, D.M., and Snyder, L.H. (2014). Topographic organization in the brain: searching for general principles. Trends Cogn. Sci. *18*, 351–363.

Pouget, A., Beck, J.M., Ma, W.J., and Latham, P.E. (2013). Probabilistic brains: knowns and unknowns. Nat. Neurosci. *16*, 1170–1178.

Ratcliff, R. (1978). Theory of Memory Retrieval. Psychol. Rev. *85*, 59–108.

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

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

Rombouts, J.O., Bohte, S.M., and Roelfsema, P.R. (2012). Neurally plausible reinforcement learning of working memory tasks. F. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger, eds. Advances in Neural Information Processing (NIPS 2012) *25*, 1880–1888.

Romo, R., Brody, C.D., Hernández, A., and Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. Nature *399*, 470–473.

Shadlen, M.N., and Kiani, R. (2013). Decision making as a window on cognition. Neuron *80*, 791–806.

Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. *86*, 1916–1936.

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J. Neurosci. *16*, 1486–1510.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., and Cohen, J.D. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. J. Exp. Psychol. Hum. Percept. Perform. *35*, 1865–1897.

Soltani, A., and Wang, X.J. (2010). Synaptic computation underlying probabilistic inference. Nat. Neurosci. *13*, 112–119.

Tanaka, K. (1996). Inferotemporal cortex and object vision. Annu. Rev. Neurosci. *19*, 109–139.

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

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. Front. Neurosci. *6*, 79.

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

Van Essen, D.C. (2002). Windows on the brain: the emerging role of atlases and databases in neuroscience. Curr. Opin. Neurobiol. *12*, 574–579.

Wald, A. (1947). Sequential Analysis. (Wiley).

Wald, A., and Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. Ann. Math. Stat. *19*, 326–339.

Wallis, J.D., Anderson, K.C., and Miller, E.K. (2001). Single neurons in prefrontal cortex encode abstract rules. Nature *411*, 953–956.

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

Wyart, V., de Gardelle, V., Scholl, J., and Summerfield, C. (2012). Rhythmic fluctuations in evidence accumulation during decision making in the human brain. Neuron *76*, 847–858.

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

Neuron, Volume *85*

Supplemental Information

A Neural Implementation of Wald's Sequential Probability Ratio Test

Shinichiro Kira, Tianming Yang, and Michael N. Shadlen

Table of Contents

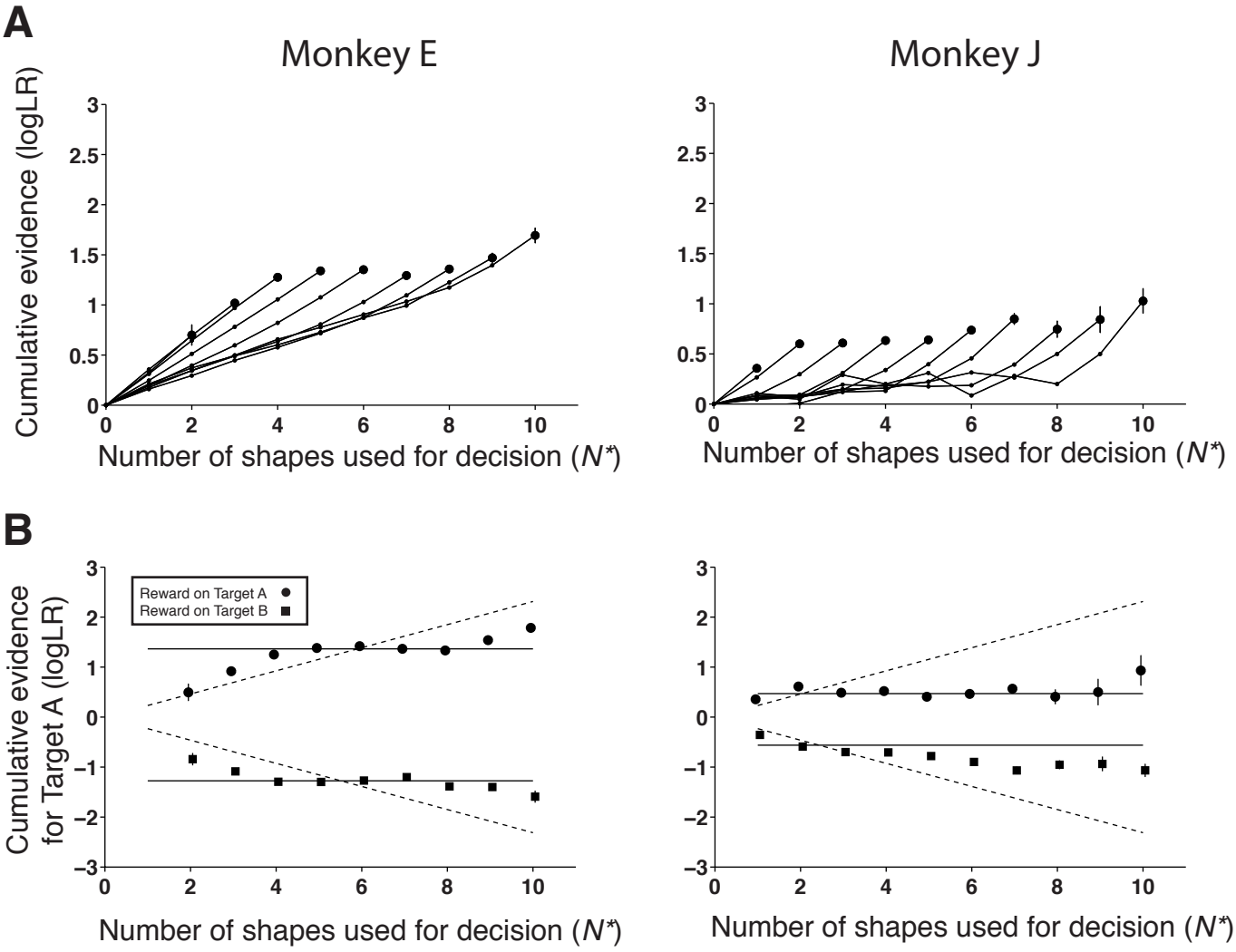

Figure S1 [related to Figure 2]. Evolution of cumulative logLR is consistent with a bounded evidence accumulation.
(A) Each trace connects the mean cumulative logLR values leading up to the final effective shape (*N**). The analysis includes both correct and error trials, and combines data for both choices (A and B) by reversing the sign of the assigned logLR for trials in which reward was assigned to Target B. Error bars at *N** are s.e.m. (cf. Figure 2E, where we show standard deviations in order to illustrate the variability across single decisions.) Notice that the mean cumulative evidence in early epochs tends to be weaker on trials leading to longer RT (i.e., larger *N**).
(B) The main alternatives to bounded evidence integration would posit separate processes for decision termination and evidence accumulation. Such models predict that the cumulative evidence presented on correct trials is the number of shapes presented multiplied by the expectation of the logLR (dashed lines); there are no degrees of freedom. According to the simplest bounded evidence accumulation model (e.g., SPRT), the cumulative evidence is the same for all *N** (solid lines); there is one degree of freedom that would explain the level. The analysis includes both correct and error trials. Error bars are s.e.m. Both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) favor bounded evidence accumulation (ΔAIC = -2041 and -1771 for monkeys E and J, respectively; ΔBIC = -2038 and -1768). AIC and BIC for each model is

$$AIC = 2k + \Lambda$$

$$BIC = k\left(\ln(n) - \ln(2\pi)\right) + \Lambda$$

where *k* is the degrees of freedom, *n* is the number of data points, and

$$\Lambda = -2\sum_i \log P(y_i \mid Model)$$

where the sigma sums the log likelihood of observing the set of sample means (y_i; black circles) given the model. Likelihoods assume Gaussian error (with s.d. approximated by s.e.m.).

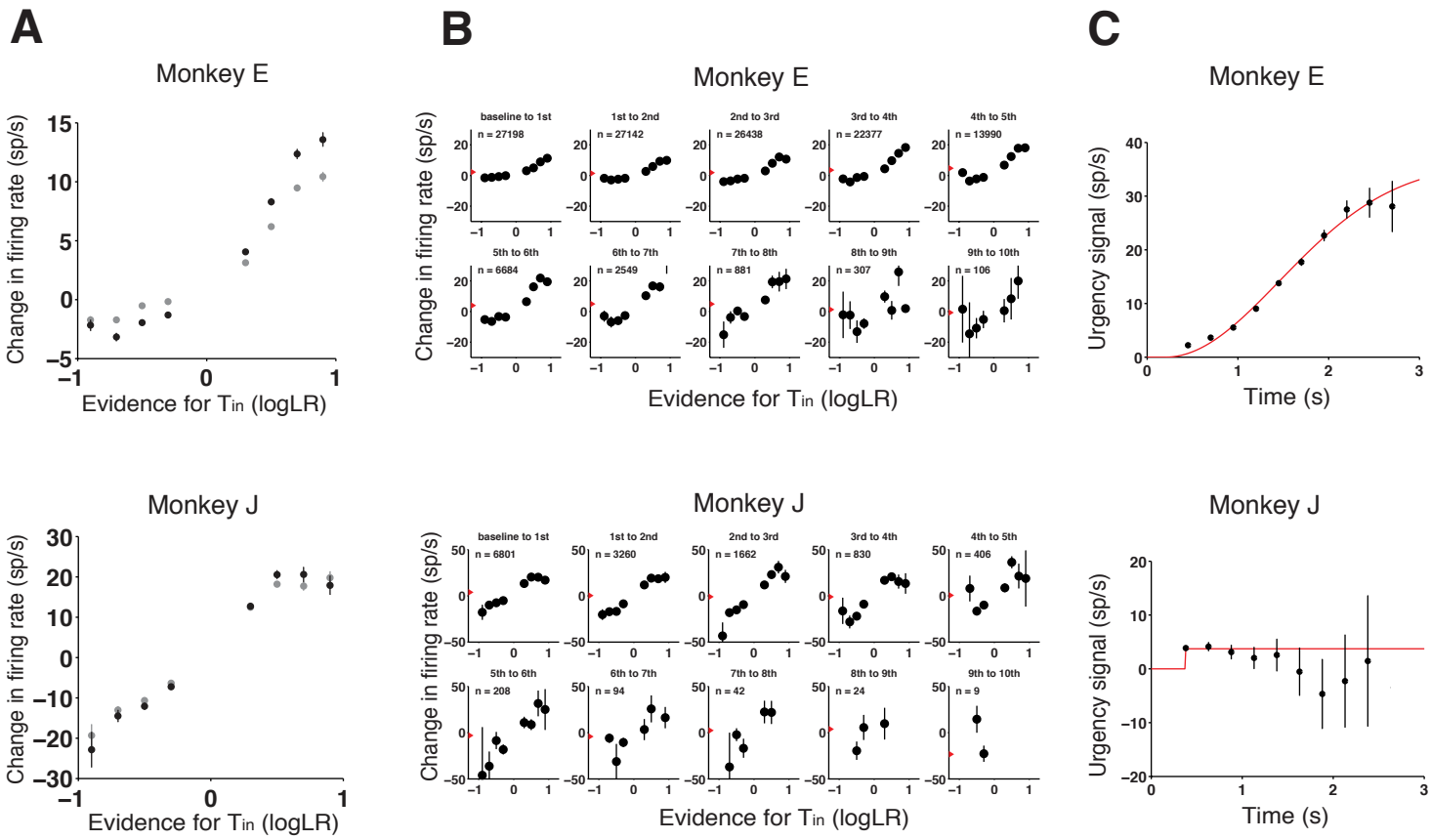

Figure S2 [related to Figures 3 and 5]. Shape-induced change in the LIP firing rate (ΔFR) in each epoch and derivation of the urgency signal.
(A) Comparison of the change in firing rates (ΔFR) estimated using a 150 ms average and from the firing rates at the very end of each shape epoch. We used the former approach for data analyses, as described in Experimental Procedures. For the model, we desired a scalar value to multiply a dynamic template, $x(t)$, (Figure 5B). This is the average of smoothed firing rates (convolved with $\alpha(t)$; see Experimental Procedures) at $\tau_S + 250$ ms after each shape. The ΔFR estimates (black) are the averages of the ΔFR_k, pooled across shape epochs. Note that the magnitudes of these estimates are not identical to the values obtained from the 150 ms average (gray; same as Figure 3B), due to evolution of the response. The black points shown here were detrended (i.e., mean subtracted) to produce the Δr values displayed in Figure 5B. Error bars are s.e.m.
(B) The change induced by the k^{th} shape is the difference between the firing rate at the very end of k^{th} shape epoch and the one preceding it (or baseline for $k = 1$). Each data point shows mean ± s.e.m. of the change in a similar fashion to Figure 3B. The number of trials contributing to each panel is indicated at the upper left corner of each panel. The number decreases in accordance with the survival function associated with the RT histograms like those depicted in Figure 2C (excluding the N^{*th} shape). Arrowheads on the left of each graph show the mean ΔFR.
(C) Cumulative sum of the mean ΔFR is shown as a function of time (black; mean±s.e.m.). The urgency signal, $u(t)$, was estimated by fitting the discrete data points with a smooth curve (Equation S4) or a step function (for monkeys E and J, respectively; see Supplemental Experimental Procedures).

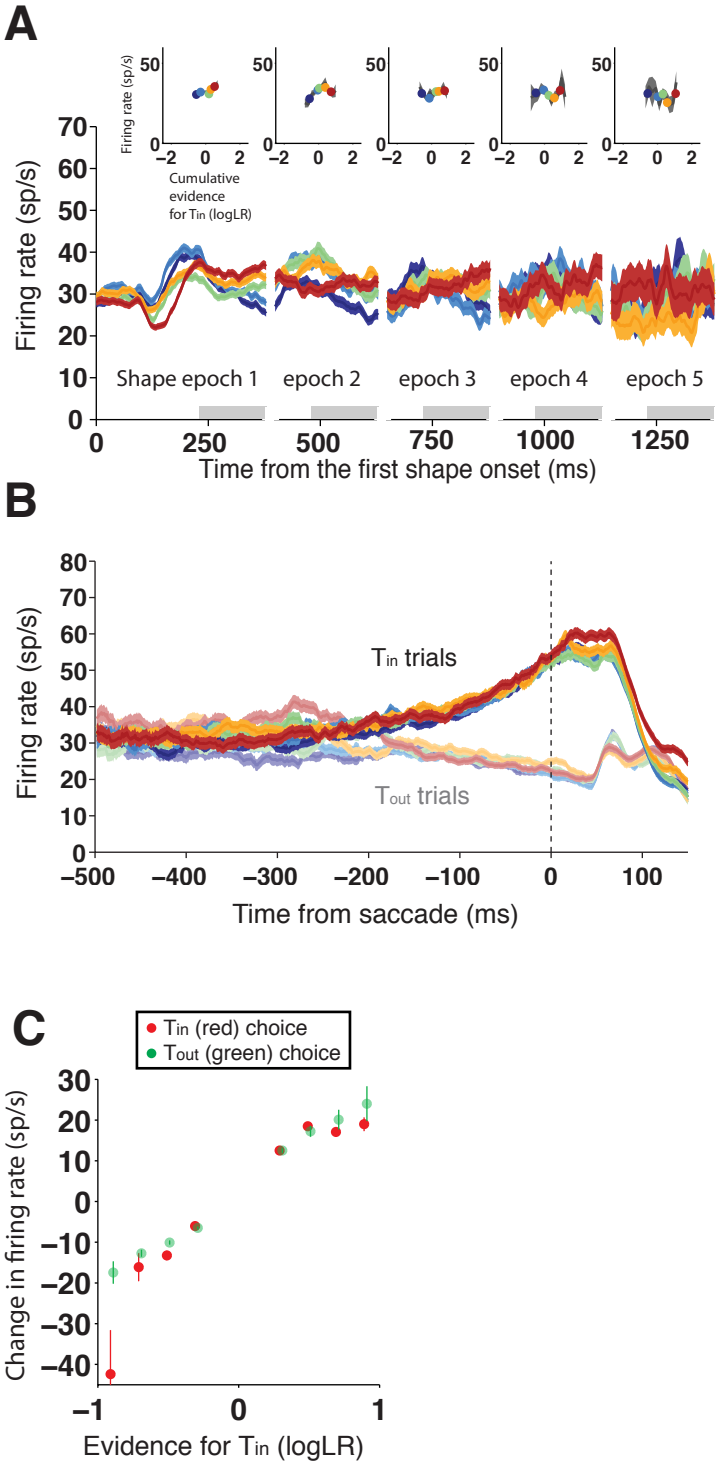

Figure S3 [related to Figures 3 and 4]. LIP neural activity did not reflect evidence accumulation when the green choice target was displayed in the RF. Thus monkey J appears to have decided for or against the red choice target.
(A) Neural activity was largely uninformative when the green choice target was displayed in the RF. Conventions are the same as in Figure 3A (curve thickness depicts s.e.m.). The firing rate did not reflect the cumulative evidence in epochs 2-5 ($p>0.05$, Bonferroni corrected), and the weak effect in epoch 1 ($p<0.05$) is explained by trials with short RT ($N^{*} \leq 2$).
(B) Neural activity was informative about the saccade (same trials as in panel A). Average firing rates accompanying decision termination are aligned to saccade initiation and grouped by choice (T_{in} or T_{out}) as in Figure 4A (curve thickness depicts s.e.m.). The selectivity arises in association with the final shapes that affected the decision (i.e., N^{*} and possibly $N^{*}-1$), but there is no representation of the cumulative evidence leading to the choice.
(C) Neurons with the red target in RF were informative about all choices (red or green). The graph comprises the same data as in Figure 3B (monkey J), split by the monkey's choice on the trial. The effect of the 8 shapes on the firing rate was similar whether or not the monkey ultimately chose red or green ($p>0.05$, Bonferroni corrected). Error bars are s.e.m.

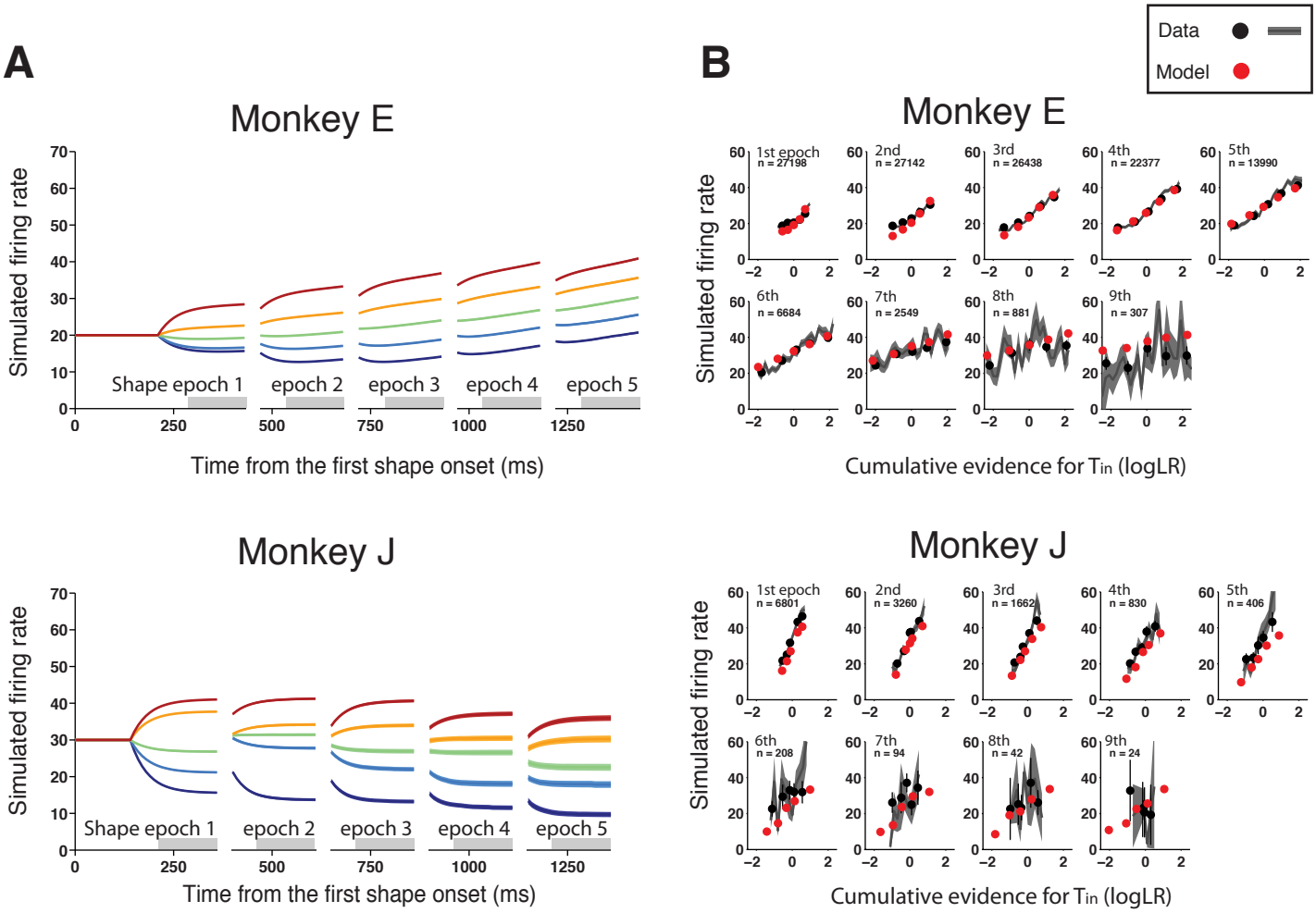

Figure S4 [related to Figure 6]. Model simulations recapitulate key observations from the LIP neural recordings obtained in the sequential inference task. Conformance to the physiology is not surprising because the model is informed by parameters derived from the neural recordings.

(A) Model firing rates reflect the cumulative evidence from the sequence of shapes. Compare to the average firing rates shown in Figure 3A. The temporal latency, τ_s, and the dynamic template, $x(t)$, are evident in the responses to the first shape. The five curves correspond to five quantiles grouped by the cumulative logLR. Note that different trials contribute to the quantiles in each shape epoch. Graphs are based on 10,000 simulated trials.

(B) The insets in Figure 3A are reproduced with changes to the plotting convention in order to facilitate comparison to model (data black; model red) and extension to later epochs. Note that <10% of the data contribute to epochs beyond $N^* = 5$ and 4 for monkeys E and J, respectively. Error bars and thickness of shaded gray traces indicate s.e.m., and the number (n) on the upper left corner of each panel indicates the number of trials that contribute to gray traces.

Movie S1 [related to Figures 3 and 4]. Example trial (T_{in} choice) with neural recording.
The movie depicts a trials in which monkey E chose the left target (T_{in}) after viewing 7 shapes. The movie is played at half speed. Eye position is depicted by the yellow dot. After acquiring fixation at the central white dot, two red choice targets appear. The left target is in the RF of the neuron. After a delay, shapes appear sequentially until the monkey initiates its saccade. Action potentials can be heard as clicks. They are also displayed as tick marks above the graph, which shows the cumulative logLR associated with each new shape (horizontal white line segments), displaced by 200 ms relative to the spikes in order to compensate for the sensory delay (τ_S; see main text). They are thus aligned approximately to the spikes induced by each new shape. The neutral level of evidence (logLR = 0) is shown as a horizontal yellow line. The eye position and lower graph were not displayed to the monkey; nor were the audio clicks. Only the fixation spot, targets, and shapes were visible to the monkey. In this trial, the cumulative logLR meandered around zero for the first few shapes and then favored T_{in}. Reflecting this pattern, the firing rate increased with the rise in the cumulative logLR until the monkey chose T_{in}.

Movie S2 [related to Figures 3 and 4]. Example trial (T_{out} choice) with neural recording.
The recording is from the same LIP neuron as in Movie S1. In this trial, the cumulative logLR initially favored T_{in} but then reversed to favor T_{out}. Reflecting this pattern, the firing rate increased initially but then diminished until the monkey chose T_{out}.

Supplemental Experimental Procedures

Bounded accumulation model

The model presented in Figure 5 applies the bounded evidence accumulation framework to explain choice and RT. Our strategy was to constrain the model as best we could, using neural data, subject to a variety of simplifications in the interest of parsimony. The models were evaluated separately for the two monkeys. The basic structure is a race between two accumulation processes. One of these uses the average changes in the observed firing rate differences (similar to Figure 3B; see Figure S2A for further details), 'detrended' by their weighted mean

$$\Delta r_i = \Delta FR_i - \left\langle \Delta FR_i \right\rangle_{i=1..8} \quad \text{(Equation S1)}$$

This detrending was necessary to remove the effect of the evidence-independent contribution to ΔFR_i, which we restore after introducing response dynamics below. The larger Δr_i are associated with shapes that furnish stronger evidence for the choice target in the neuron's RF (i.e., T_{in}), whereas the more negative Δr_i are associated with shapes that furnish stronger evidence for T_{out}. For the competing process, we used the same 8 values of Δr_i, but changed their sign to reverse the association between shape and choice target (Figure 5B). We recognize that this is an oversimplification (we have recorded from only a few neurons with ipsilateral RFs), but it allows us to constrain the model with neural responses instead of leaving Δr_i for the competing process as free parameters.

These Δr_i values can be thought of as the expected momentary evidence supplied by the shape. We assume that the actual change induced by each shape on each trial is both noisy and

dynamic. Thus, we add an unbiased sample of a Normal distribution $\varepsilon \in N(0,\sigma_\varepsilon)$ to the expectation, Δr_i, and allow the sum to scale a temporal impulse function, $x(t)$, illustrated in Figure 5C. We approximate the template as the integral of an exponential decay

$$x(t) = \begin{cases} 1-\exp(-[t-\tau_s]/\tau_\Delta) & \text{for } t \geq \tau_s \\ 0 & \text{otherwise} \end{cases} \qquad \text{(Equation S2)}$$

where τ_Δ = 50 ms and τ_s is the "sensory" delay described in Results. We used the same τ_Δ for the two monkeys.

For each simulated trial, the model firing rate, $v(t)$, is described by an accumulation of noisily scaled impulses added to a time-dependent, evidence-independent baseline, termed the urgency signal, $u(t)$.

$$v_j(t) = v_0 + u(t) + \sum_k \left(\delta\left[t - 250(k-1)\right] * \left[(\Delta r_{i,j,k} + \varepsilon_{j,k}) \cdot x(t) \right] \right) \qquad \text{(Equation S3)}$$

where i signifies the shape identity (as above), j indicates the accumulator (population A or B in Figure 5), k is epoch (as above), * denotes convolution, and $\varepsilon_{j,k}$ are samples of unbiased noise. The $\Delta r_{i,j,k}$ are conditionally independent—that is, given reward assignment to Target A or B, the shapes are sampled independently from the corresponding distribution j (= A or B) in Figure 1B throughout the trial. The $\varepsilon_{j,k}$ are also independent and identically distributed Gaussian random variables with mean zero and standard deviation σ_ε. In this equation, the value of $\Delta r_{i,j,k}$ plus an independent random value ($\varepsilon_{j,k}$) is updated every 250 ms. In principle, one might expect that σ_ε

could be measured from the firing rate variability of the recorded neurons. However, single neurons are more variable than a population of neurons, and the degree of reliability in the population depends on factors we did not measure, such as the degree of correlated variability among neurons (Shadlen et al., 1996; Zohary et al., 1994). Therefore, we treat σ_ε as a free parameter.

We approximated a continuous urgency function, $u(t)$, from neural data as follows. In each epoch, we calculated the weighted mean of the ΔFR with respect to the 8 possible shapes, $\Delta FR_k = \left\langle \Delta FR_{i,k} \right\rangle_{i=1..8}$ (Figure S2B, red arrows). We approximated the cumulative function of these evidence-independent increments with a scaled, cumulative Rayleigh distribution, which is continuous in time (Figures 5D and S2C, red traces):

$$u(t) = \kappa\left[1 - \exp\left(-(t-\tau_s)^2 / 2\sigma_u{}^2\right)\right] \qquad for\ t \geq \tau_s \qquad \text{(Equation S4)}$$

where κ and σ_u are free parameters. For monkey J, a step function approximated the urgency signal as well as a Rayleigh cumulative distribution function (Figure S2C). The series of steps may seem overly baroque, but they merely serve to replace the discrete measurements of ΔFR in each epoch with a continuous function of time. Note that $u(t)$ effectively restores the average ΔFR_k, removed in the detrending step (Equation S1), in a manner that varies as a function of time (i.e., shape epoch).

Three boundary conditions are informed by the neural recordings. The firing rate before the first shape-induced change (v_0, Equation S3) was approximated by the mean firing rate in these intervals (20 and 30 sp/s for monkeys E and J, respectively). We implemented a non-absorbing (i.e., nonterminating; also known as reflecting) lower bound at zero for both stochastic processes

so that $v_j(t)$ does not take a negative value. Implementation of the reflecting bound is supported by physiology because the T_{in}-favoring shapes can induce positive ΔFR even when accumulated evidence strongly favored T_{out}, and the magnitude of the change is comparable to ΔFR without any conditionalization (data not shown). The process stops when one of the rates exceeds a threshold (θ). The value of θ was based on an estimate of the firing rate at a point of convergence preceding T_{in} choices (Figure 4A, insets), obtained by regressing the firing rate at time t against the cumulative logLR at τ_s+100 ms prior to that time point. A shape appearing at this latency would give rise to a change in firing rate that is nearing its maximum at time t. The analysis leads to an estimate of when shapes fail to influence the firing rate (arrows, Figure 4A insets; $<\tau_m>$ in the model). For both monkeys, this time point establishes the threshold firing rate for terminating decisions (θ = 50 sp/s for both monkeys), and the expected motor delay, $<\tau_m>$, from the threshold crossing to the saccade initiation. We are least secure about the estimates of $<\tau_m>$, in part because of the discrepancy between monkeys and in part because the analysis depicted in Figure 4A does not identify a unique termination point. Since variation in $<\tau_m>$ would induce small changes in the estimated θ, we repeated the modeling exercise using a range of values of $<\tau_m>$ (hence θ) and confirmed that the results are not altered appreciably with variation in θ associated with other candidate values of $<\tau_m>$.

Our uncertainty about $<\tau_m>$ is mitigated somewhat because we model the motor delay on single trials as a random variable: gamma distributed with mean $<\tau_m>$ and standard deviation σ_m, the second free parameter in the model (Figure 5E). The predicted RT distribution is the convolution of this gamma distribution with the predicted distribution of decision times. Note that the decision time already incorporates the sensory latency and the rise time of the temporal impulse.

To fit the free parameters of the model, σ_ε and σ_m, we used Monte Carlo methods to simulate trials and maximized the likelihood of observing the RT, given predicted RT distributions for correct and error trials: $P_{cor}(RT \mid \sigma_\varepsilon, \sigma_m)$ and $P_{err}(RT \mid \sigma_\varepsilon, \sigma_m)$ (Figure 6A, red histograms; total probability = 1). These distributions were obtained by simulating 10,000 trials, using the same set of random number seeds (one for each trial). For i^{th} trial in each iteration (i.e., each choice of σ_ε and σ_m), the identical sequence of shapes was presented up to RT to furnish the log likelihood,

$$\log L = \sum_j \left(m_j \log P_{cor}(RT_j \mid \sigma_\varepsilon, \sigma_m) + n_j \log P_{err}(RT_j \mid \sigma_\varepsilon, \sigma_m) \right) \qquad \text{(Equation S5)}$$

where j denotes the index for time bins, m_j and n_j indicate the observed number of correct and error trials, respectively, in a given time bin. Because the probability distributions were estimated using Monte Carlo methods, we used histogram approximations (bin size = 10 ms) to avoid singularities that would arise if $P(RT) = 0$. The sequences of shapes used in the simulation were different from those in the experiment.

Reward rate

We compared the monkeys' rate of reward acquisition to levels achievable by the SPRT and variants thereof, tailored to our task. Under SPRT, the ideal observer accumulates evidence

$$W_n = \sum_{k=1}^{n} w_{i,k} \qquad \text{(Equation S6)}$$

where i represents the identity of unique shapes (i = 1 .. 8), k indicates the order in the sequential presentation, and $w_{i,k}$ is the assigned weight to the corresponding shape, and terminates according to the rule

$$\begin{cases} W_n > \Theta & : \text{choose A} \\ W_n < -\Theta & : \text{choose B} \\ -\Theta \le W_n \le \Theta & : \text{acquire a new sample} \end{cases} \qquad \text{(Equation S7)}$$

Upon termination at the k^{th} shape, we assume a minimum processing time of 250 ms per shape plus a fixed non-decision time T_{nd}, during which shapes failed to influence the choice (270 ms and 180 ms for monkeys E and J, respectively; see Results). The time per trial is the sum of RT plus the inter-trial interval plus a pre-stimulus latency (T_{pre}) from fixation to the onset of the first shape (0.3 and 0.9 s, on average, for monkeys E and J, respectively). The reward rate is

$$\dot{R} = \frac{P(c \mid \Theta)}{\langle RT \rangle + \langle T_{ITI} \rangle + \langle T_{pre} \rangle} \qquad \text{(Equation S8)}$$

where $P(c \mid \Theta)$ is the probability of a correct choice given the bound height Θ. Note that the expectation of T_{ITI} depends on $P(c \mid \Theta)$ because of the error penalty, and it depends on the distribution of RT because of the minimum delay to reward (monkey E). The optimal reward rate was calculated by maximizing over Θ. We also mention a variant of SPRT that would accommodate the minimum delay to reward (monkey E) and places Θ at ±∞ until most of the minimum delay to reward has passed (i.e., 1600 – $\langle \tau_m \rangle$ ms).

To evaluate the effect of imperfect learning of assigned weights on performance, the weights were replaced by the corresponding subjective weight (Figure 2F; β_i in Equation 3). To evaluate the effect of noise, we corrupted the subjective weights by Gaussian noise (mean = 0) with standard deviation η, chosen to equate the noise level at the expectation of logLR (given the reward is assigned to Target A) to the fitted noise parameter, σ_ε, relative to the expectation of the Δr_i when the correct choice is T_{in}. That is,

$$\eta = \left\langle w_i \right\rangle_{i,A} \frac{\sigma_\varepsilon}{\left\langle \Delta r_i \right\rangle_{i,T_{in}}} \qquad \text{(Equation S9)}$$

We recognize that these reward rates are only approximations to the optimal values. To achieve the latter, we would be required to optimize over policy including the possibility of non-stationary bounds, as shown by Drugowitsch et al., 2012.

Supplemental references

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

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J. Neurosci. 16, 1486-1510.

Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature 370, 140-143.